

A Simple Theory of the Survey Response: Answering Questions versus Revealing Preferences

Author(s): John Zaller and Stanley Feldman

Source: *American Journal of Political Science*, Vol. 36, No. 3 (Aug., 1992), pp. 579-616

Published by: Midwest Political Science Association

Stable URL: <https://www.jstor.org/stable/2111583>

Accessed: 03-08-2018 18:48 UTC

---

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact [support@jstor.org](mailto:support@jstor.org).

Your use of the JSTOR archive indicates your acceptance of the Terms & Conditions of Use, available at <https://about.jstor.org/terms>



JSTOR

*Midwest Political Science Association* is collaborating with JSTOR to digitize, preserve and extend access to *American Journal of Political Science*

# *A Simple Theory of the Survey Response: Answering Questions versus Revealing Preferences\**

John Zaller, *University of California, Los Angeles*

Stanley Feldman, *State University of New York at Stony Brook*

Opinion research is beset by two major types of “artifactual” variance: huge amounts of over-time response instability and the common tendency for seemingly trivial changes in questionnaire form to affect the expression of attitudes. We propose a simple model that converts this anomalous “error variance” into sources of substantive insight into the nature of public opinion. The model abandons the conventional but implausible notion that most people possess opinions at the level of specificity of typical survey items—and instead assumes that most people are internally conflicted over most political issues—and that most respond to survey questions on the basis of whatever ideas are at the top of their heads at the moment of answering. Numerous empirical regularities are shown to be consistent with these assumptions.

Virtually all public opinion research proceeds on the assumption that citizens possess reasonably well formed attitudes on major political issues and that surveys are passive measures of these attitudes. The standard view is that when survey respondents say they favor *X* they are simply describing a preexisting state of feeling favorably toward *X*.

Accumulating evidence on the vagaries of mass political attitudes, however, has made this view increasingly dubious. If, as is well known, people are asked the same question in a series of interviews, their attitude reports are highly changeable. Many, as much evidence also shows, react strongly to the context in which questions are asked, to the order in which options are presented, and to wholly nonsubstantive changes in question wording. These phenomena are more than methodological curiosities; they raise serious doubts about what public opinion surveys measure.

In view of this, we propose a new understanding of the mass survey response. Most citizens, we argue, simply do not possess preformed attitudes at the level of specificity demanded in surveys. Rather, they carry around in their heads a mix of only partially consistent ideas and considerations. When ques-

\*Originally prepared for delivery at the fifth annual meeting of the Political Methodology Society. The data used in the paper were collected by the Board of Overseers of the National Election Studies under a grant from the National Science Foundation. Collecting these data made unusually heavy demands on the staff of the Institute for Social Research at the University of Michigan, where the study was conducted. In this connection we are grateful to ZoAnne Blackburn, Steve Pinney, and Santa Traugott. We also thank Barbara Geddes, Shanto Iyengar, Don Kinder, Matt Lyons, and Douglas Rivers for their helpful comments on various drafts of the paper. Finally, we would like to thank Hiroaki Minato for invaluable assistance in preparing the data for analysis. Neither these individuals nor institutions bear any responsibility for any errors of judgment or fact that may appear in this paper.

*American Journal of Political Science*, Vol. 36, No. 3, August 1992, Pp. 579–616  
© 1992 by the University of Texas Press, P.O. Box 7819, Austin, TX 78713

tioned, they call to mind a sample of these ideas, including an oversample of ideas made salient by the questionnaire and other recent events, and use them to choose among the options offered. But their choices do not, in most cases, reflect anything that can be described as true attitudes; rather, they reflect the thoughts that are most accessible in memory at the moment of response.

A model based on these claims can, as we show, provide a far better account of the existing evidence on political attitudes, including such related matters as attitude consistency and the effects of political awareness, than can currently dominant models of the survey response.

We begin with a review of existing models of public opinion and then outline an alternative model and draw upon a range of new and existing evidence to demonstrate its value.

### Limits of Existing Theories

#### *Response instability*

One of the most unsettling findings of opinion research has been the discovery of a large component of randomness in most people's answers to survey questions. If the same people are asked the same question in repeated interviews, only about half give the same answers. The data in Table 1, based on interviews of the same persons six months apart, illustrate the problem. As can be seen from the entries on the main diagonals, only 45% to 55% gave the same answer both times, even though about 30% could have done so by chance alone.<sup>1</sup> The amount of response instability differs from one issue to another (see Feldman 1989), but the cases shown in Table 1 are fairly typical.

In his famous paper "The Nature of Belief Systems in Mass Publics," Converse (1964) argued that response instability is due mainly to individuals who lack meaningful attitudes but nevertheless indulge interviewers by politely choosing between the response options put in front of them—but choosing in an almost random fashion. "Large portions of an electorate," he suggested, "simply do not have meaningful beliefs, even on issues that have formed the basis for intense political controversy among elites for substantial periods of time" (1964, 245).

This conclusion has been strongly challenged by scholars who contend that, although people's "survey responses" fluctuate greatly, citizens have underlying "true attitudes" that are overwhelmingly stable (Achen 1975, 1983; Dean and Moran 1977; Erikson 1979; Feldman 1989). The fluctuations that appear in people's overt survey responses are attributed to "measurement error," where

<sup>1</sup>Given data from only two points in time, it is impossible to distinguish systematic attitude change from random fluctuation. However, analysis of data from three- and five-wave panels strongly suggests that almost all response instability represents random rather than systematic change (Feldman 1989).

**Table 1. Response Stability over Repeated Interviews:  
Two Examples**

| American Relations with Russia<br>(Corner Percentaging) |           |        |         |        |
|---|-----------|--------|---------|--------|
| Attitudes in January 1980                               |           |        |         |        |
|   | Cooperate | Middle | Tougher | Unsure |
| June 1980:  |           |        |         |        |
| Cooperate   | 25%       | 8      | 8       | 2      |
| Middle  | 7         | 4      | 5       | 1      |
| Tougher   | 5         | 5      | 17      | 2      |
| Unsure  | 4         | 2      | 3       | 4      |
| <i>N</i> =  | (338)     | (153)  | (266)   | (74)   |

  

| Level of Government Services<br>(Corner Percentaging) |       |        |           |        |
|---|-------|--------|-----------|--------|
| Attitudes in January 1980                             |       |        |           |        |
|   | Cut   | Middle | Keep Same | Unsure |
| June 1980:  |       |        |           |        |
| Cut   | 24%   | 6      | 5         | 6      |
| Middle  | 8     | 4      | 2         | 2      |
| Keep Same   | 5     | 4      | 15        | 3      |
| Unsure  | 8     | 2      | 3         | 6      |
| <i>N</i> =  | (362) | (122)  | (208)     | (138)  |

*Note:* The exact questions were: "Some people feel it is important for us to try very hard to get along with Russia. Others feel it is a big mistake to try too hard to get along with Russia. Where would you place yourself on this scale, or haven't you thought about this?" Respondents were then asked to place themselves on a seven-point scale. In this table, points 1, 2, and 3 have been counted as "cooperate"; 4 is counted as middle; 5, 6, and 7 have been counted as "tougher."

The second question was: "Some people think the government should provide fewer services, even in areas such as health and education, in order to reduce spending. Other people feel it is important for the government to continue the services it now provides even if it means no reduction in spending. Where would you place yourself on this scale, or haven't you thought about this?" Item was recoded as above.

*Source:* National Election Studies, 1980 Panel Survey.

such error is said to stem from the inherent difficulty of mapping one's attitudes onto the unavoidably vague language of survey questions.

Both approaches to response instability have critical deficiencies. Converse's thesis, which takes any instability as evidence of a "nonattitude," was an extreme claim intended to characterize attitudes only on highly abstract issues.

On more typical issues, as Converse and Markus (1979) argue, people's attitudes may be more or less "crystallized" and are, as a result of this, more or less stable. But this only raises the question of how crystallization can be measured apart from its supposed effect on response stability. Since no one has ever said, attitude crystallization remains, as Krosnick and Schuman (1988) have pointed out, more a metaphor than a testable theory of attitude stability.

The newer "measurement error" theory of response instability appears equally underspecified at its theoretical core. When, as all estimates agree, *measurement "error" typically constitutes one-half or more of the variance of typical attitude items*, one naturally wonders what exactly this "error" consists of and how it has been generated. Yet we presently know so little about these questions that the term remains essentially an alternative name for "unexplained variance."

### *Response Effects*

In addition to the *random response variance* that has been attributed to measurement error, there exists *systematic variance* from artifactual "response effects." Consider a well-known Cold War experiment on attitudes toward Soviet journalists. In a split-half sample, 37% of respondents were willing to allow communist reporters in the United States. Yet when, in the other half-sample, respondents were first asked whether U.S. reporters should be allowed in Russia (which most favored), the percentage agreeing to allow Russian reporters here doubled to 73%.

There are numerous other findings of this type: people are less likely to describe themselves as interested in politics just after they have been asked about obscure issues (Bishop, Oldendick, and Tuchfarber 1984); people's attitudes toward abortion are affected by the kinds of items (concerning, e.g., religion or women's rights) that precede it (Tourangeau and Rasinski 1988; Tourangeau et al. 1989); people give quite different answers to open-ended questions than to questions that ask them to choose among a series of prespecified options (Schuman and Scott 1987). Seemingly irrelevant features of survey design do not, as might be suspected, affect only unsophisticated people who might be suspected of having nonattitudes; they ensnare all types of respondents (Krosnick and Schuman 1988; see also Bishop 1990).

The literature on response effects thus makes it clear that survey questions do not *simply* measure public opinion. They also shape and channel it by the manner in which they frame issues, order the alternatives, and otherwise set the context of the question. This has led researchers to a conclusion that seems indisputable but that is fundamentally at odds with the assumptions of most political scientists about the nature of political attitudes: namely, people do not merely *reveal* preexisting attitudes on surveys; to some considerable extent, people are using the questionnaire to decide what their "attitudes" are (Bishop, Oldendick, and Tuchfarber 1984; Zaller 1984; Feldman 1990).

Psychologists make this argument most directly. Tourangeau and Rasinski

(1988), for example, argue that responses to attitude questions can be understood as the outcome of a question-answering process in which people (1) decide what the issue is; (2) canvas their minds for relevant thoughts; (3) combine ideas into a coherent attitude; and (4) map the resulting attitude onto available response options. Because, as they maintain, features of the interview process can affect each of these steps, the questionnaire can also readily affect what gets reported as public opinion.

More recently, Wilson and Hodges (1991) have proposed a model in which attitudes are “temporary constructs” that are made up at the moment of response on the basis of ideas in a large but internally conflicted “data base.” In perhaps the most ambitious attempt to deal with the question-answering process (and much else), Wyer and Srull (1989) offer an information-processing model having more than three dozen elements.

### *What Needs to Be Done*

Despite the evidence from psychologists and survey methodologists, public opinion researchers largely ignore both the longstanding problem of massive over-time response instability and the newer findings on questionnaire effects. Moreover, many of those who recognize the problems make what amount essentially to patch-ups of the traditional view. In the case of response effects, the patch-up consists of trying to prevent the problem from becoming conspicuous; this is done by, for example, making sure to keep question order constant across time series surveys or, in some computer-assisted surveys, randomizing question order across respondents. In the case of response instability, the patch-up consists of statistical corrections for measurement error, corrections that create the comfortable illusion that fixed “true attitudes” exist beneath the enormous surface noise.

The challenge, then, is to devise a theory that accommodates both response instability and response effects and that is crafted to the kinds of problems and data facing analysts of public opinion. This is what we attempt to do in this paper. The theory we propose is, we admit, simpler than would be necessary to explain all of the findings that psychologists have now documented. But a theory sufficiently complex to do this would have little value to most political scientists, and our aim is, above all, to reach this group—convincing it that the conventional understanding of public opinion is unworkable and that a practical alternative is available.

### **An Alternative Model of the Survey Response**

According to conventional attitude theory, individuals choose whichever prespecified option comes closest to their own position. But if, as we contend, people typically do not have fixed positions on issues, how do they make their choices?

Since most survey research takes the “true attitude” as its primitive unit of

analysis, little attention has been devoted to this question. If, however, one turns to studies that employ depth interviews, one finds much useful evidence. Among the best of these is Hochschild's (1981) study of attitudes toward equality, *What's Fair?* From her interviews with 28 persons, Hochschild found that people would, if asked to do so, readily answer fixed-choice questions, but that given the opportunity to talk, "people do not make simple statements; they shade, modulate, deny, retract, or just grind to a halt in frustration. These manifestations of uncertainty are just as meaningful and interesting as the definitive statements of a belief system" (238).

Hochschild particularly emphasizes the ambivalence of many of her respondents. This ambivalence frequently leads them to contradict themselves—which is to say, *to give temporally unstable responses in the course of a single conversation*. Consider this account of the attitudes of one of her subjects toward government income guarantees: "Vincent Sartori cannot decide whether or not the government should guarantee incomes, because he cannot decide how much weight to give to the value of productivity. He believes that the rich are mostly undeserving and . . . yet he is angry at 'welfare cheats' who refuse to work. . . . Caught between his desire for equality and his knowledge of existing injustice, on the one hand, and his fear that a guaranteed income will benefit even shirkers, on the other, he remains ambivalent about policies toward the poor" (252).

Current attitude models seem quite irrelevant to these observations. The reason for Sartori's vacillation is not, as students of Converse might say, that he has no opinion on this question, nor is it that, as users of measurement error models might say, Sartori has a "true attitude" that Hochschild is unable to measure reliably. It is rather that Sartori has conflicting opinions, or at least conflicting *considerations*, that lead him to give different responses at different times, depending on how he thinks about the issue.

It is easy to object to the limitations on rigor inherent in depth interviews such as Hochschild's. Nonetheless, we are persuaded that the basic point about ambivalence—that *individuals possess multiple and often conflicting opinions toward important issues*—represents an important insight. Much psychological research reinforces this view. Memory researchers, for example, have shown that people store huge amounts of information in their long-term memories, but can retrieve and use only a fraction of it at one time. The particular material they do recall depends on a combination of chance and recency of activation. Hence, people make quite different judgments and belief statements, depending on the information they happen to recall from long-term memory (Raaijmakers and Shiffrin 1981; Wyer and Hartwick 1984).

Another research tradition, mainly concerned with social cognition, focuses on the organization of ideas in the mind. A central concept in much of this research is the "schema," a term that has been adapted from cognitive psychology. A schema is a cognitive structure that organizes prior information and experience

around a central value or idea and that guides the interpretation of new information and experience.

A critical point about schemas is that people typically have several of them available for understanding any given phenomena. For example, an individual being introduced to a "professor" would react quite differently if the new person were instead described as "a mother of four." That is, different associations would come to mind, different qualities of the person would be noticed, different conclusions would be drawn from the person's mannerisms, and so forth. In short, the perceiver's attitude toward the person would be different. Thus, Tesser (1978), in statements that represent the dominant thrust of much cognitive psychology and that nicely capture a central feature of the model we propose, writes: "An attitude at a particular point in time is the result of a constructive process. . . . And, *there is not a single attitude toward an object* but, rather, any number of attitudes depending on the number of schemas available for thinking about the objects" (297–98). And "persons do not have a *single* feeling or evaluation of an object. Feelings vary depending upon the particular cognitive schema we 'tune in'" (307).

These studies tend both to corroborate Hochschild's insights concerning ambivalence and to undermine the conventional political science assumption (which is at the heart of both Converse's black-and-white model and Achen's "true attitude" measurement error model) that it is normal for individuals to have a single, coherent attitude on issues. In view of this, our model will follow Hochschild in assuming that people carry around in their heads a mix of more or less consistent "considerations," where a consideration is defined as a *reason for favoring one side of an issue rather than another*. (E.g., a person who thinks about "Pentagon waste" while deciding a question about defense spending has raised a consideration that may well control her decision on that issue.)

The first axiom of our model may now be stated as:

**AXIOM 1:** The ambivalence axiom. Most people possess opposing considerations on most issues, that is, considerations that might lead them to decide the issue either way.

We emphasize that the concept of consideration, as used in this axiom, is not just another word for schema. First, it is cast in the language of everyday political discourse (see Kelley 1983), as befits a term intended for political rather than psychological analysis. Second, it makes no reference to mental structures or operations, such as the interpretation of raw sensory input, that are central to the concept of schema.

Our next problem is to decide how individuals transform the diverse considerations in their heads into closed-ended responses. One possibility, as Taylor and Fiske (1978) suggest, is that individuals make choices "off the top of the



head” on the basis of the first idea that comes to mind. Thus, people may make social judgments by seizing on “a single, sufficient and salient explanation . . . often the first satisfactory one that comes along. . . . [I]nstead of employing base rate or consensus information logically, people are more often influenced by a single, colorful piece of case history evidence. . . . Instead of reviewing all the evidence that bears upon a particular problem, people frequently use the information which is most salient or available to them, that is, that which is most easily brought to mind” (251). Tversky and Kahneman’s (1982) well-known work on framing effects reinforces the view that individuals often are overly influenced by a single, dominant consideration.

At the same time, much data in both political science (Campbell et al. 1960) and cognitive psychology (Anderson 1974) indicate that on other occasions, individuals reach decisions by averaging across a range of competing ideas. Thus, Kelley (1983) shows that voters seem to decide which presidential candidate to support by summing up all of their “likes” and “dislikes” about each party and presidential candidate and choosing the one with the highest net total.

The axioms we propose allow individuals to respond to survey questions on the basis of either one or many considerations, depending on how many happen to be readily accessible in memory at the moment the question is posed:

**AXIOM 2:** The response axiom. Individuals answer survey questions by averaging across the considerations that happen to be salient at the moment of response, where saliency is determined by the *accessibility* axiom.

**AXIOM 3:** The accessibility axiom. The accessibility of any given consideration depends on a stochastic sampling process, where considerations that have been recently thought about are somewhat more likely to be sampled.<sup>2</sup>

For the case in which a person devotes great thought and attention to an issue, Axiom 3 implies that there may be multiple considerations salient in memory at the moment of answering questions about the issue and hence many considerations to be averaged across. But a person who rarely thinks about an issue and who is confronted by an interview situation that requires a succession of quick answers (Feldman 1990) may have only one consideration immediately available in memory, in which case the averaging rule reduces to answering on the basis of a single “top-of-the-head” consideration, as suggested by Taylor and Fiske.

These three axioms, although spare and informal, can be used both to organize much existing research and to generate testable new hypotheses about the nature of the mass survey response, as we shall now begin to show.

<sup>2</sup>The empirical warrant for this axiom is extremely strong (see Bargh et al. 1986; Higgins and King 1981; Bodenhausen and Wyer 1987).

### Data

Since we base much of our analysis on data from the 1987 Pilot Study of the National Election Studies (NES), it is worth pausing briefly to describe this study. The survey attempted to measure, *inter alia*, the “considerations” that underlie people’s responses to standard closed-ended survey items. The study was conducted in two waves a month apart; 457 persons were interviewed in the May wave and 360 in the June wave. All had previously participated in the 1986 National Election Study. Other technical details of the study are available through the NES at the University of Michigan.

The basic method was to ask people a closed-ended policy item and then to ask them to talk in their own words about the issues it raised. The closed-ended items were telephone versions of the standard NES items on job guarantees, aid to blacks, and government services and spending. In form A, respondents were asked the open-ended probes immediately after answering the given closed-ended policy item. The exact form of the “retrospective” open-ended probes was:

Still thinking about the question you just answered, I’d like you to tell me what ideas came to mind as you were answering that question. Exactly what things went through your mind. (Any others?)

This probe was designed to elicit a “memory dump” of the considerations immediately salient in people’s minds. Prior work by Ericsson and Simon (1984) shows that such probes can work effectively if asked immediately after a given task has been carried out.

In form B, interviewers read the items in the usual way, but, without waiting for an answer, they asked respondents to give their reactions to the principal idea elements in the question. For the job guarantees question, the probes were as follows:

Before telling me how you feel about this, could you tell me what kinds of things come to mind when you think about *government making sure that every person has a good standard of living*? (Any others?)

Now, what comes to mind when you think about *letting each person get ahead on their own*? (Any others?)

Immediately following these probes, the interviewer reread the original closed-ended question and took the person’s reply to it. (Full question wording is in the appendix.)

The two types of probes are clearly not equivalent. The “retrospective” probes, which were posed after people had answered the question in the normal way, were designed to find out what exactly was on people’s minds at the moment

of response. The “prospective” or “stop-and-think”<sup>3</sup> probes, on the other hand, were designed to induce people to search their memories more carefully than they ordinarily would for pertinent considerations. Note that the stop-and-think probes do not raise new ideas or push the respondent in a particular direction; they simply require the respondent to say explicitly what meaning he or she attaches to the defining phrases of the question.

Respondents were randomly assigned to question form and answered the same type of the question in each wave of the study. The three test items and associated open-ended probes appeared near the end of each wave of the survey. Interviewers wrote down as faithfully as possible all responses to the open-ended probes, including incidental side comments (e.g., “This is a tough one”). The transcribed comments were subjected to an elaborate classification scheme, with as many as four comments coded for each probe.<sup>4</sup> Respondents on the stop-and-think side averaged about 3.7 codable comments per policy item, with almost all respondents offering at least one codable comment. The average on the retrospective side was 2.9.

All comments, including side comments, were rated on several variables by staff coders at the Institute for Social Research at the University of Michigan. Because the coding project was considered a difficult one, only experienced coders were used. The most important variable was “directional thrust of comment,” which indicated which side of the issue, if any, the remark favored. Although this variable noted ambivalence, confusion, and nonissue concerns, 75% of comments had a clear directional thrust. The other key coding classification was “frame of reference,” a variable that included more than 140 categories and tried to capture the substantive content of each remark. The frame codes referred to general principles (e.g., equality, the role of government), individualism and the work ethic, the fairness of the economic system, particular groups (e.g., blacks, the elderly), personal experience, and particular government programs. (Further information about these codes may be found in Table 5 and associated text; all codes are fully described in the study codebook.)

These data are not without limitations. The most obvious is coder reliability. About a tenth of all interviews were double-coded, and, although exact reliability data are not available, the coding supervisor reported a “difference” between coders on 10% to 15% of all cases. This difference rate was regarded as normal for material of this type,<sup>5</sup> but it is higher than one would hope for. In addition, 10% of remarks were so unclear that they could not be assigned a directional thrust. A final limitation is the difficulty in confidently distinguishing one “considera-

<sup>3</sup>This apt designation is the invention of Kathleen Knight.

<sup>4</sup>For the aid to blacks item, there were up to six probes: three questions, each followed by a query for “any others?” As many as four remarks were coded in connection with each of these six probes. On the other items, there were two initial probes, each with follow-up probes.

<sup>5</sup>Personal communication from Steve Pinney, who supervised our project at ISR.

tion” from another. When two remarks have clearly opposing thrusts, this is no problem. But people sometimes offer a series of remarks on the same side of an issue. Do such remarks represent separate considerations, or just elaborations on a single idea? Even a person listening to the interviews as they occurred, as we did, would sometimes be uncertain; coders working from an imperfect transcript would experience greater uncertainty.

Even in retrospect, we are not sure how these problems could have been significantly ameliorated in the context of a mass survey. We thus feel the Pilot Study data are the best that are readily attainable for directly examining the ideational underpinnings of mass attitudes, but we admit that they are, nonetheless, far from perfect.

### Tests of the Model

#### *Preliminary Check of the Ambivalence Axiom*

We begin assessment of the model by making a plausibility check of the axiom that claims that most individuals hold multiple, conflicting ideas on most issues. Our data give us three ways to measure the extent of ambivalence in the public, as follows:

1. *A count of the number of opposing remarks by each person that can be paired against one another.* If, for example, a respondent makes two comments with a liberal thrust and two with a conservative thrust, his score on the conflict scale is two. If he makes three (or more) on one side of the issue and only two on the other, the conflict score is still two, since the number of opposing comments that can be paired remains two.

2. *A count of the times people spontaneously express ambivalence or difficulty in making up their minds.* A special code was created to capture such remarks; it reads as follows: “Mention indicates ambivalence, conflict (e.g., ‘I see merit in both sides’; ‘That’s a tough question’; ‘Depends’; ‘Both are valid points’).”

3. *A count of the number of times people make “two-sided comments.”* Included in the frame of reference codes are special “star codes” that indicate a directional thrust to the comment but also ambivalence with respect to that direction. Star codes apply to cases in which respondents had a preference but were clearly paying some attention to the other side of the issue. Instructions to coders for use of star codes read as follows:

*A star code is used only for cases in which there is a single thought or comment that encompasses two opposing elements (e.g., “Although I think X, I nevertheless favor Y.”* Star codes are used for comments in which *R* sees two sides to an issue.

Examples of star codes are “people should try to get ahead on their own, but government should help when necessary” and respondent “admits problem(s) with any program or type of program, but insists it is worthwhile anyway.” A count of the star-coded remarks may thus be considered a measure of ambivalence.

From these three measures we created a fourth: a summary of the indices on which a person scored +1 or higher. Because conflict and ambivalence are equally consequential whether they occur within the course of one interview or across separate interviews, all indices are calculated across both waves of the survey.

Frequency distributions on these four measures are shown in Table 2. As can be seen, each measure captures substantial amounts of ambivalence, a result that is consistent with the first axiom. Even on the more conservative evidence of the retrospective probes, which involve only one query in each wave,<sup>6</sup> the summary measure indicates that 36% to 48% of respondents are to some degree ambivalent on these three issues. And this is surely an understatement. What the retrospective probes capture, as explained, is the reason the person has answered the item as he just has; they cannot capture anything like the full range of ideas in the person’s head. However, the prospective probes were designed to tap a wider range of the ideas in people’s minds; on evidence from them, roughly 75% of respondents are at least somewhat conflicted on the three issues.

These results provide clear initial support for the model’s first axiom, the ambivalence axiom. Since the other two axioms cannot be tested directly, we turn now to an examination of the deductive implications of the three axioms taken all together.

### *First Deductions from the Model*

We begin with deductions from the model that are entirely straightforward and perhaps uninteresting and proceed to more useful and important ones. If, as the accessibility axiom claims, the accessibility of a given consideration depends on the amount of thought devoted to an issue, we should find that people who are, in general, more politically involved have more considerations at the top of their heads and available for use in answering survey questions. This is the first of 18 deductions from the model that we make and test (Deduction 1). (To keep track of the deductions, each will be numbered in parentheses, as here.) Despite some indication of nonmonotonicity in the data, Table 3 essentially confirms this expectation.<sup>7</sup>

Similarly, we would expect persons who have greater interest in an issue to

<sup>6</sup>More probes were made, but since they were directive, we make no use of them.

<sup>7</sup>The nonmonotonicity is probably real, as a more sophisticated version of our model leads one to expect (see Zaller, in press, chap. 8).

**Table 2. Expressions of Conflict and Ambivalence on Political Issues**

|       | Retrospective Probes<br>Conflicting Considerations |          |               | Stop-and-Think Probes<br>Conflicting Considerations |          |               |
|-------|--|----------|---------------|---|----------|---------------|
|       | Jobs   | Services | Aid to Blacks | Jobs  | Services | Aid to Blacks |
| Count |  |          |               |   |          |               |
| 0     | 73.9% <sup>a</sup>                                 | 57.8     | 73.4          | 36.9%   | 30.7     | 29.0          |
| 1     | 22.6   | 33.6     | 22.6          | 27.3  | 29.0     | 21.6          |
| 2     | 3.5  | 5.2      | 4.0           | 22.2  | 21.6     | 25.0          |
| 3 +   | 0.0  | 3.4      | 0.0           | 13.1  | 18.7     | 24.4          |
|       | Expressions of Ambivalence                         |          |               | Expressions of Ambivalence                          |          |               |
| 0     | 76.9   | 83.5     | 78.8          | 63.2  | 72.2     | 71.1          |
| 1 +   | 23.1   | 15.5     | 27.8          | 36.8  | 27.8     | 28.9          |
|       | Two-Sided Remarks (Star Codes)                     |          |               | Two-Sided Remarks (Star Codes)                      |          |               |
| 0     | 75.0   | 91.7     | 81.4          | 64.9  | 85.2     | 72.3          |
| 1 +   | 25.0   | 8.3      | 18.6          | 35.1  | 14.8     | 27.7          |
|       | Total Indications                                  |          |               | Total Indications                                   |          |               |
| 0     | 60.2   | 51.4     | 63.6          | 25.9  | 24.3     | 24.7          |
| 1     | 15.7   | 37.6     | 16.9          | 39.1  | 50.9     | 44.0          |
| 2     | 18.5   | 7.3      | 12.7          | 9.8   | 14.8     | 12.0          |
| 3     | 5.6  | 3.7      | 6.8           | 25.3  | 10.1     | 19.3          |

*Note:* <sup>a</sup>Measures are described in text.

*Source:* 1987 NES Pilot Study.

**Table 3. Effect of Political Awareness on Volume of Open-Ended Comments**

|                     | Level of Political Awareness |     |        |     |      |
|---------------------|------------------------------|-----|--------|-----|------|
|                     | Low                          |     | Medium |     | High |
| Standard of living  | 2.3                          | 2.9 | 3.5    | 3.5 | 3.2  |
| Government services | 2.3                          | 3.2 | 4.3    | 4.6 | 4.7  |
| Aid to blacks       | 3.3                          | 3.9 | 4.4    | 5.2 | 4.4  |
| <i>N</i> =          | 44                           | 53  | 54     | 38  | 46   |

*Note:* Cell entries are the average number of discrete substantive remarks in the given cell. These data are derived from stop-and-think probes, but results from retrospective probes show the same pattern. The awareness measure is described in the appendix.

*Source:* 1987 NES Pilot Study.

have, all else equal, more thoughts about that issue readily accessible in memory than other persons (Deduction 2). Since the Pilot Study did not directly ask respondents how important or interesting each of the policy issues was to them, our ability to test this expectation is limited. We did, however, find that blacks raised more considerations than whites in connection with the aid to minorities item ( $p < .01$ ). Government employees also had slightly more to say than other persons about the government services item ( $p = .07$ ). Unemployed persons, however, were not more likely to raise considerations pertinent to the job guarantees item.

Axioms 2 and 3 claim people answer survey questions by averaging across whatever considerations are salient in memory. If this is so, we should find strong correlations between measures of people's thoughts as they answer a survey item and the direction of decision on the item itself (Deduction 3). Thus if, for example, a person makes two remarks that favor the liberal side of the issue and one that favors the conservative side, we would expect that the person would, on average, take the liberal side of the issue. Although this inference may seem hardly worth testing, it is by no means obvious that it can be confirmed. Social psychologists, working in the domain of social cognition, have turned up cases in which the direction of people's open-ended thoughts is uncorrelated, or even negatively correlated, with evaluations of the given issue.

As Hastie and Park (1986) have contended in an influential essay, the surprising noncorrelations occur because people typically do not *construct* attitude statements from ideas they can retrieve from memory as they are questioned. Instead, people *recall* attitudes formed at an earlier time. Thus, they maintain, there is no necessary correlation between top-of-the-head ideas and attitude statements (see Lodge, McGraw, and Stroh 1989).

Notwithstanding this, our data indicate substantial correlations between the ideas most accessible to individuals at the moment of response and the response given. To show this, we created additive indices of people's open-ended remarks, coded for directional thrust. We then correlated these indices with responses to the closed-ended items, as shown in Table 4. On the stop-and-think side, correlations between the indices and their associated dichotomous item in each wave of the survey averaged about .40. When an index of all remarks over both waves of the survey is correlated with a scale that consists of responses to the closed-ended items from the two waves, the correlations average about .50. In the other half of the study, the correlations between individuals' retrospective remarks and their closed-ended responses in the same wave of the survey averaged .70. When retrospective remarks and items were summed and correlated across waves of the survey, the correlations averaged .80. (Much of the difference in correlations between the stop-and-think side and the retrospective side appears to be due to error in the coding of the closed-ended responses, which, as discussed below, seems to have been higher on the stop-and-think side.) Given that the closed-

Table 4. Relationship between Available Thoughts and Closed-Ended Items

|  | Wave 1       | Wave 2       | Combined     |
|--|--------------|--------------|--------------|
| Correlations with Remarks Made Just <i>before</i><br>Answering Closed-Ended Question |              |              |              |
| Job guarantees   | .39<br>(212) | .39<br>(161) | .50<br>(173) |
| Government services  | .31<br>(187) | .36<br>(153) | .41<br>(165) |
| Aid to blacks  | .57<br>(220) | .48<br>(165) | .63<br>(166) |
| Correlations with Remarks Made Just <i>after</i><br>Answering Closed-Ended Question  |              |              |              |
| Job guarantees   | .79<br>(126) | .70<br>(123) | .79<br>(105) |
| Government services  | .79<br>(137) | .70<br>(105) | .78<br>(106) |
| Aid to blacks  | .67<br>(144) | .83<br>(114) | .83<br>(112) |

*Note:* Cell entry is Pearson correlation between open-ended remarks made just before (or just after) answering closed item and scores on the closed item. First column shows correlations from first wave of survey; second column shows correlations from second wave; third column shows correlations between remarks from both waves and combined item scores on both waves.

*Source:* 1987 NES Pilot Study.

ended items in these tests are essentially dichotomies,<sup>8</sup> and the fragile nature of our open-ended data, these are sizable correlations.<sup>9</sup>

We should add that the findings of correlations between top-of-the-head thoughts and attitude reports are not, taken alone, clinching evidence that the former have caused the latter. Rather, the correlations—which, as we have indicated, could not be taken for granted in light of past research—simply represent one of more than a dozen cases in which we have been able to develop evidence that is consistent with the three-axiom model.

<sup>8</sup>A few respondents volunteered “it depends” middle responses, which were accepted.

<sup>9</sup>The substantial magnitudes of these correlations do not, however, show that closed-ended survey responses are, after all, perfectly valid indicators of the considerations in people’s heads. As we note immediately below, different considerations are salient at different times, leading people to make different attitude reports at different times. What Table 4 shows is only that closed-ended attitude reports are reliable indicators of the considerations that are salient *at the moment of making a response*.



### *Explaining Response Instability*

Response instability over repeated interviews is, as we have indicated, one of the most important and disturbing empirical regularities associated with the mass survey response. In this section, we attempt to use our model to explain this instability. We begin with simple illustrations of our approach and then proceed to more systematic analysis.

When asked in the May interview about the proper level of government services, one respondent, identified as a teacher, emphatically favored higher levels of services and spending. The country was facing an educational crisis, the teacher said, and more expenditures for education were drastically needed. Any cuts in federal services or spending would inevitably reduce the already inadequate funds available for education. Just a month later, however, the same individual favored cuts in government spending. Government was too big and had to be cut back. There was no reference to the educational crisis that had preoccupied this individual just a few weeks earlier.<sup>10</sup>

Researchers have long known that different people can answer identical questions as if they concerned different topics. What the vignette of the vacillating teacher shows is that *the same person can answer the same question at different times as if it involved different topics*. This can happen, according to the model, because the considerations that determine people's survey answers vary across interviews. Thus, people can give strongly felt, contradictory survey responses without either changing their mix of feelings on the issue or consciously experiencing any ambivalence or conflict—if the particular considerations that determine their survey responses have shifted.

Our data were collected with a specific view to detecting and measuring such shifts. Table 5 presents closed-ended summaries of these data for four respondents.<sup>11</sup> Note that these data are from the stop-and-think side, in which respondents were encouraged to think about issues somewhat more fully than they ordinarily would.

Consider respondent A. His first reaction to a guaranteed standard of living was that it was inconsistent with American ideals; he was also bothered by the unfairness of supporting those who refuse to work. Yet he worried about letting individuals get ahead on their own, saying that some people need special help and that society has an obligation to help the needy. In the second interview, however, there was no sign of this ambivalence. Respondent A gave six reasons why individuals ought to get ahead on their own, including a restatement of his

<sup>10</sup>We would like to present verbatim transcriptions of what this and other respondents said. However, the Human Subjects committee at the University of Michigan has determined that such use of the raw protocols would be an invasion of the respondents' right to privacy.

<sup>11</sup>The selections were subject to the constraint that we wanted two respondents who were stable on the closed-ended items and two who were unstable.

**Table 5. Content of Open-Ended Responses on Job Guarantees Question**

| <i>Could you tell me what kinds of things come to mind when you think about . . .</i> |  |   |  |
|---|--|---|--|
| First Wave  |  | Second Wave   |  |
| <i>. . . government making sure that every person has a good standard of living?</i>  | <i>. . . letting each person get ahead on their own?</i> | <i>. . . government making sure that each person has a good standard of living?</i> | <i>. . . letting each person get ahead on their own?</i> |
| <i>Respondent A:</i>  |  | (Stable Conservative on Fixed Response)   |  |
| 110. Idea is un-American (Con.)   | 180. Some people need help (Lib.)                        | 158. Tax burden too great (Con.)  | 140. Individualism/work ethic (Con.)                     |
|   | 161. Duty to help the needy (Lib.)                       | 110. Idea is un-American (Con.)   | 133. Equal opportunity exists for all (Con.)             |
|   | 136. Unfair if some don't work (Con.)                    | 346. Program/food stamps (Con.)   | 344. Program/education (Con.)                            |
|   | 145. All should make it alone but some need help (ETU)   |   |  |
| <i>Respondent B:</i>  |  | (Stable Conservative on Fixed Response)   |  |
| 158. Tax burden too great (Con.)  | 344. Program/education (Lib.)                            | 344. Program/education (Lib.)   | 140. Individualism/work ethic (Con.)                     |
| 156. Gov't. red tape (Con.)   | 344. Program/education (Lib.)                            | 219. Group ref. to middle class (Con.)  | 142. Shiftless people deserve fate (Con.)                |
| 144. Value of competition   | 110. Idea is un-American (Con.)                          | 152. Limited gov't. (Con.)  | 147. Motivation to work (Con.)                           |
|   |  | 140. Individualism/work ethic (Con.)  |  |
| <i>Respondent C:</i>  |  | (Unstable, Conservative to Liberal)   |  |
| 140. Individualism/work ethic (Con.)  | 140. Individualism/work ethic (Con.)                     | 151. Gov't. must insure equal opp. (Lib.)   | 140. Individualism/work ethic (Con.)                     |
| 161. Duty to help the needy (Lib.)  |  |   |  |
| <i>Respondent D:</i>  |  | (Unstable, Liberal to Conservative)   |  |
| 344. Program/education (Lib.)   | 140. Individualism/work ethic (ETU)                      | 150. Idea of welfare state (Lib.)   | 111. Fairness of Amer. system (Con.)                     |

Table 5 (continued)

| Could you tell me what kinds of things come to mind when you think about . . . |   |  |   |
|--|---|--|---|
| First Wave   |   | Second Wave  |   |
| . . . government making sure that every person has a good standard of living?  | . . . letting each person get ahead on their own? | . . . government making sure that each person has a good standard of living? | . . . letting each person get ahead on their own? |
| 357. Program/housing (Lib.)  |   | 348. Program/health (Lib.)   |   |
| 356. Program/welfare (Lib.)  |   | 143. Work & welfare (Lib.)   |   |

*Note:* A full description of these codes can be found in the ICPSR codebook for the 1987 Pilot Study. Each remark is identified in parentheses as having a liberal, conservative, or uncertain directional thrust (ETU indicates that the evaluational thrust of the remark was unclear). Evaluational thrust was coded independently of substantive content of remark.

feeling that job guarantees are un-American, without raising any opposing considerations.

Respondent A, who opposed government job guarantees in both of his closed-ended attitude reports, is thus less stable in his reaction to the guaranteed standard of living question than his stable closed-ended responses would suggest. He went from being an ambivalent conservative on this issue to being a confident conservative. Given the ambivalence of the initial interview, one would not be surprised if, over many encounters with this question or slightly rephrased versions of the question, he occasionally saw the central issue as aid to the needy rather than un-American ideals and, on this basis, expressed support for the social welfare option. Certainly he has real impulses in that direction.

Now consider respondent C. Although she is ambivalent at both interviews on the basis of similar considerations, she changes her closed-ended response from conservative to liberal. One can imagine that she may have, in effect, tossed a mental coin in deciding how to answer the fixed question—not because, as Converse’s nonattitudes thesis suggests, she had no meaningful reaction to the issue (clearly, she did), and not because, as the measurement error models suggest, she was not quite sure what the question asked (she saw it the same way at both interviews), but because she was undecided between largely stable but conflicting impulses.

These cases show that even if, as we contend, people base their attitude reports on the ideas that are most immediately salient to them, it is by no means easy to explain over-time response instability. In the most straightforward case, instability arises from changes in the considerations that are immediately salient

at the time of making an attitude report, as in the case of the vacillating teacher. But other, quite different scenarios are also possible, as we have seen. With these complexities in mind, let us proceed to a more systematic examination of response instability.

If, as the model claims, individuals possess competing considerations on most issues, and if they answer on the basis of whatever ideas happen to be at the top of their minds at the moment of response, one would expect a fair amount of over-time instability in people's attitude reports (Deduction 4). The reason is that the consideration(s) that are stochastically accessible at one interview might not be so prominent at the next. This inference is strongly supported by a mass of existing evidence (e.g., Table 1).

The model not only anticipates response instability, but also expects it to have a definite structure. Suppose that 80% of the considerations in one person's head induce her toward a liberal response on a given issue, while 20% induce her toward a conservative response; and suppose that for a second person, these proportions are reversed. If each based her survey responses on a one-element sample from the distribution of considerations in her head, each would exhibit response instability over time, but over the long run, the first person would be liberal 80% of the time and the second would be conservative 80% of the time. Thus, citizens would have central tendencies that are stable over time, but their attitude statements would fluctuate greatly around these central tendencies (Deduction 5). This is, in fact, exactly the pattern that has been obtained repeatedly by researchers in the "measurement error" tradition (Achen 1975; Erikson 1979; Judd and Milburn 1980; Judd, Milburn, and Krosnick 1981; Feldman 1989; Zaller 1990; see, however, Krosnick 1988).

If, as shown in Table 3, more politically aware persons have a larger number of considerations at the top of their head and accessible for use in answering questions, they should, all else being equal, exhibit greater stability in their closed-ended responses. The reason is that attitude reports formed from an average of many considerations will be a more reliable indicator of the underlying population of considerations than an average based on just one or two considerations (Deduction 6).

Although initial research failed to confirm this expectation, Feldman (1989) and Zaller (1990) have more recently shown in separate data sets that political awareness is, in fact, associated with a reduction in the chance variation associated with people's attitude reports. (The difference between the initial and later tests of this expectation is that the recent work uses tests of political information as the measure of political awareness.)

By parallel logic, people should be more stable in their responses to closed-ended policy items concerning doorstep issues—that is, issues so close to everyday concerns that most people routinely give some thought to them (Deduction 7).

Evidence tending to support Deduction 7 may be found in the discussion of racial issues in Converse (1964), the discussion of moral issues in Converse and Markus (1979), and generally in Feldman (1989).

Another implication of our model is that greater ambivalence ought to be associated with higher levels of response instability (Deduction 8). Since, as we just saw, some individuals who exhibit no apparent ambivalence within a single interview may nonetheless be quite conflicted, it is essential that, in testing this implication of the model, we employ a measure of ambivalence that spans both interviews. Accordingly, we have built a measure of the extent to which an individual's considerations consistently favor one or the other side of a given issue across both waves of the survey. We constructed this measure by means of the following formula:

$$\left| \frac{\sum (\text{liberal remarks}) - \sum (\text{conservative remarks})}{\sum (\text{liberal}) + \sum (\text{conservative}) + \sum (\text{ambivalent})} \right|.$$

A score of one on this measure would indicate that the person's remarks were either all liberal in their thrust or all conservative, while a score of zero would indicate that the person had made an equal number of liberal and conservative remarks. We expect higher stability on the closed-ended items for cases in which all of the person's remarks run in the same direction. The data in Table 6 support this expectation. In five of six trials, this measure was associated with a statistically significant increase in response stability; in the sixth case, the relationship achieves marginal statistical significance ( $p = .07$ ).

The results are not, however, as strong as they might be. If people formed responses by averaging across accessible considerations, stability rates should vary between 50% (for people evenly divided in their considerations) to 100% (for people with perfectly consistent considerations). Particularly on the stop-and-think side, the data significantly depart from this expectation. How can this be explained?

The most likely reason for the shortfall from expectations is coding error. As reported earlier, random check-coding indicated that coders disagreed on the coding of 10% to 15% of all open-ended remarks, and such miscoding obviously impairs our ability to determine which respondents should be perfectly stable and which perfectly random.

To test this explanation for the imperfect results in Table 6, we developed an item-level measure of coding error and correlated it with the magnitude of the stability-consistency relationships shown in Table 6. The correlation was, as expected, quite high.<sup>12</sup> Moreover, coding error was consistently higher on the stop-

<sup>12</sup>To estimate item-level error rates, we used item-level percentages of remarks that the coders found to be uncodable. The rate of uncodable remarks varied from a low of about 5% for one item to a high of nearly 20% for another. We expected that the magnitude of the stability-consistency relation-

**Table 6. Response Stability and Consistency of Considerations**

|   | Job<br>Guarantees     | Government<br>Services | Aid to<br>Blacks |
|---|-----------------------|------------------------|------------------|
| Retrospective Considerations              |                       |                        |                  |
| <i>Consistency<br/>of considerations:</i> | <i>N</i>              | <i>N</i>               | <i>N</i>         |
| .00                                       | .50 (7)               | .59 (11)               | .57 (7)          |
| .01 to .50                                | .80 (20)              | .70 (25)               | .71 (19)         |
| .51 to .99                                | .77 (15)              | .78 (16)               | .80 (15)         |
| 1.00                                      | .91 (63)              | .87 (54)               | .96 (71)         |
|   | ( $p < .01$ )         | ( $p < .02$ )          | ( $p < .01$ )    |
| Stop-and-Think Considerations             |                       |                        |                  |
| <i>Consistency<br/>of considerations:</i> | <i>N</i>              | <i>N</i>               | <i>N</i>         |
| .00                                       | .63 <sup>a</sup> (16) | .54 (14)               | .57 (14)         |
| .01 to .50                                | .68 (74)              | .77 (63)               | .83 (66)         |
| .51 to .99                                | .73 (37)              | .80 (50)               | .84 (44)         |
| 1.00                                      | .88 (45)              | .73 (37)               | .88 (41)         |
|   | ( $p < .02$ )         | ( $p < .07$ )          | ( $p < .01$ )    |

*Note:* Cell entries are proportion stable in their responses to closed-ended items from wave 1 to wave 2. Measure of consistency is described in text. *P*-values are based on uncollapsed measure.

*Source:* 1987 NES Pilot Study.

and-think side, which explains why the shortfall from expectations is higher in that form of the survey.<sup>13</sup>

Within the limits of the data, then, the results in Table 6 are probably about as strong as could be expected. We obtained sizable stability effects at conventional levels of statistical significance in five of six trials and were close to significance on the sixth, despite significant measurement error, a small sample, and abnormally low base rates of instability owing to the relatively short time between reinterviews.

One additional point needs to be made. The measure of consistency in

ships in Table 6, as summarized by unstandardized regression slopes, would be greatest for those items with the lowest rates of uncodable remarks. This was strongly the case ( $r^2 = 0.93$  on a logarithmic fit;  $df = 5$ ;  $p < .05$ ). Further details of this test are reported in Zaller (in press, chap. 4).

<sup>13</sup>It is, we might add, not surprising that error rates are higher on the stop-and-think side. It would seem easier, a priori, to determine the directional thrust of a remark if the remark has been made in justification of an opinion just rendered, as on the retrospective side, than if it has been given as part of a discussion of what a given phrase means to the respondent, as on the stop-and-think side.

Table 6 was calculated over both waves of the survey in order to capture consistency of considerations both within interviews and across them. The consistency measure can, however, be calculated within a single interview. When this is done, we come up with a strong, clear finding: the one-wave measure of consistency of considerations has almost no capacity to predict over-time response instability (in either form).

This finding indicates that the conflict most responsible for response instability is conflict that occurs across rather than within interviews and that respondents are often unaware of their conflict as they answer questions. Thus, the vacillating teacher exhibited no conflict over government services within either interview, but substantial conflict across interviews. Most likely, once the teacher began to view the government services item through the prism of either “bloated government” or “education crisis,” he or she fell into a mindset that blocked thinking about the other point of view.<sup>14</sup> This suggestion fits nicely with our model, which holds that people answer survey questions hastily and on the basis of incomplete memory searches.

### *Explaining “Response Effects”*

The model can also explain several important response effects in mass surveys, that is, cases in which seemingly irrelevant features of question order or design affects the responses given. In this section we survey a variety of these response effects.

Consider first “order effects,” such as the effect of a prior question on support for the rights of communist reporters, as discussed earlier. If, as the model claims, people are normally ambivalent on issues but answer on the basis of whatever ideas are most accessible at the moment of answering, raising new considerations in immediate proximity to a question should be able to affect the answers given by making different considerations salient (Deduction 9).

The intrusion of unexpected or novel considerations into the question-answering process would not, however, be expected to affect all respondents equally. Some people may possess considerations that are so consistent in support of one side of an issue that the admission of one competing consideration should have no effect. Others, however, may be deeply ambivalent on the issue—that is, may possess a roughly even balance of pro and con considerations. These are the persons who should be most strongly affected by artificial changes in question order (Deduction 10).

Tourangeau et al. (1989) have reported support for this expectation. People who reported that they had mixed feelings about an issue (and who also said the

<sup>14</sup>Tourangeau and Rasinski (1988) distinguish between “interpretation” of a question and “retrieval” of information relevant to answering it. In terms of this distinction, our results suggest that interpretation is the more important source of response instability.

issue was very important to them) were quite susceptible to question order effects—or what the authors call “carryover effects”—while other persons were not susceptible at all. The carryover effects in the vulnerable groups (where the groups varied from issue to issue) ranged from lows of 4 and 8 percentage points on abortion and welfare, to highs of 34 and 36 percentage points on aid to the Contra rebels and defense spending, with an average carryover of 19 percentage points.

Another important type of response effect is what may be called the “endorsement effect.” In this case, reference to a political figure or group systematically alters the public’s responses to a given policy dilemma. For example, Mueller (1973) found that questions that indicated that President Johnson favored a particular policy option led to greater support for it. Similarly, references to communism systematically increased support for U.S. involvement in Korea and Vietnam. This anticommunism endorsement effect, according to Mueller, suggested “somewhat conflicting observations. On the one hand, support for the war was clearly tied to the anti-Communist spirit in America at the time. To generate a kind of war fever, one merely had to toss the words, ‘Communist invasion,’ into the discussion. On the other hand, the Communist element was not entirely built into the response to the war because Americans had to be reminded of it before their anti-Communism was fully activated” (1973, 48). The tendency of people to base attitude reports on the ideas that are most immediately salient to them, as specified in Axioms 2 and 3, well explains such effects (Deduction 11).

Although most response effects are considered “methodological artifacts,” they are sometimes given substantive names or interpretations. Consider these:

1. *Race of interviewer effects.* Shortly after a 1986 *New York Times* poll found that President Reagan’s approval rate among blacks was 37%, a *Washington Post* poll estimated that black approval of Reagan was only 23%. The difference was traced to the fact that, while the *Times* followed normal interview procedures, the *Post* used black interviewers who informed their black respondents that they would be participating in a study of the attitudes of black Americans. As Sussman (1986) maintains, the likely effect of this was to induce black respondents to “think black” in their evaluations of Reagan’s performance (the first item on the survey).<sup>15</sup>

<sup>15</sup>This and other race-of-interviewer effects might be interpreted as “social desirability” effects. If by social desirability effects one means cases in which people consciously misreport attitudes in order to avoid embarrassment, a social desirability interpretation is at variance with our argument. But if, as seems equally consistent with available evidence, social desirability is taken to mean cases in which people are unsure what their attitudes are but are influenced by the immediate context to give greater weight to a particular consideration, then social desirability effects represent simply another type of situation in which people make attitude reports on the basis of the ideas most immediately salient to them.



2. *Reference group effects.* In a classic social psychology experiment, Charters and Newcomb (1958) found that Roman Catholics were more likely to state attitudes that were consistent with church doctrine (e.g., on birth control) if, just prior to questioning, their religion was made salient to them. This effect was taken as evidence of the importance of “reference groups” on attitudes.

3. *“Priming effects” of television news.* According to Iyengar and Kinder (1987) and Iyengar (1991), television news often functions to “prime” certain ideas, thereby making them more accessible for use in evaluating presidential performance, deciding between candidates in elections, and assessing the nation’s most important problems. In this way, TV news is said to affect attitude reports without permanently changing individuals’ underlying attitudes.

4. *“Framing effects” of question wording and question order.* If the considerations people use in answering questions can be primed by TV news, they can also be primed by how questions are worded or framed. For example, Lau and Sears (1983) and Lau, Sears, and Jessor (1990) have experimentally shown that questions about personal financial status are more strongly correlated with evaluations of incumbent politicians when the questions are asked in close proximity. The reason, presumably, is that a consideration used in answering one question remains available for answering subsequent questions, thereby inducing a correlation.<sup>16</sup>

The mechanism responsible for each of these effects appears to be a tendency for people to answer questions at least partly on the basis of ideas that have been made momentarily salient to them. As such, they may be counted as response effects for which our model of the survey response gives an explanation (Deductions 12, 13, 14, 15).

Milburn (1987) has documented yet another way in which question order effects can occur. He found that asking respondents to “tell me everything that comes to mind when you think of a Liberal (Conservative)” caused subsequent attitude reports to be more ideologically consistent with one another than were the attitude reports of a control group. Milburn obtained this effect, however, only for persons who had either liberal or conservative leanings to begin with. Price (1991) has recently replicated these findings in an experiment conducted on college students and in nonexperimental data from national samples. In both cases, Price found that simply asking persons to place themselves on a liberal-conservative rating scale was sufficient to increase the ideological consistency of responses to subsequent policy items—but, again, only for persons who possessed a clear ideological leaning (i.e., were not centrists or unable to place themselves on the self-rating scale).

<sup>16</sup>There is disagreement on whether these effects have artificially inflated estimates of “pocket-book voting” in NES surveys (see Lewis-Beck 1985; Lau, Sears, and Jessor 1990), but no disagreement that framing effects could, in principle, do so.

This type of question order effect is also explainable from the model. Having had their ideological orientations made salient to them just prior to answering policy items, those respondents who possess such orientations are more likely to rely on them as a consideration in formulating responses to subsequent policy questions, thereby making those responses more strongly correlated with their ideological positions and hence also more ideologically consistent with one another (Deduction 16).

This section has examined a variety of empirical regularities—question order effects, endorsement effects, race-of-interviewer effects, reference group effects, priming effects, question frame effects, and the effect of making ideology salient. Within conventional attitude theory, some of these empirical regularities are taken as substantive findings and some as methodological artifacts. We have shown, however, that all may be explained as manifestations of a common theoretical mechanism, namely, the normal tendency of people to respond to survey questions on the basis of the ideas that happen to be, for whatever reason, immediately salient to them.

#### *Explaining the Effects of Extra Thought*

Survey responses, as conceived here, are not “attitudes” per se; they are unreliable indicators of the mix of considerations in the person’s mind—unreliable because, among other things, people normally answer without retrieving from memory all relevant considerations. If, however, people could be artificially induced to retrieve a larger than normal number of considerations, it should improve the reliability of their responses to closed-ended items.

Our intent in designing the stop-and-think probes was to create such an inducement. By requiring individuals to discuss the elements of a question before answering it, we were inducing them to call to mind and take account of a wider range of ideas than they normally would. We therefore expected that responses following the stop-and-think treatment would be, all else equal, more reliable indicators of the set of underlying considerations than responses made in the standard way, that is, in the retrospective condition (Deduction 17).

Unfortunately, our ability to test this expectation is compromised by an artifact. Because of the use of an explicit “no interest” option in the retrospective condition but not in the stop-and-think condition, more respondents failed in the retrospective condition to respond to the issue item. Low-awareness persons were most affected by this question difference; their no opinion rate averaged 38% in the retrospective condition but only 4% in the stop-and-think condition.<sup>17</sup> This means that retrospective respondents, especially less-aware ones, are a more *selected* group and would, for this reason alone, be expected to be more ideologically consistent than their stop-and-think counterparts. This artifact runs against

<sup>17</sup>By contrast, the difference in no opinion rates across forms averaged 13% (16% vs. 3%) in the highest information quintile.

the grain of the anticipated stop-and-think effect, making it more difficult to demonstrate the effect, particularly for less-aware respondents.

We developed two tests of the expectation of increased response reliability in the stop-and-think condition. In the first test, we expected a measure of social welfare ideology (see appendix) to be more strongly correlated with the target items (jobs, government services, and aid to blacks) in the stop-and-think condition than in the retrospective condition. We used the following interactive regression model to test this expectation, where Form refers to question form:

$$\text{Item} = b_0 + b_1 \times \text{Form} + b_2 \times \text{Ideo.} + b_3 \times \text{Form} \times \text{Ideo.}$$

When we estimated this model for respondents who scored in the upper 40% of our measure of political awareness, we found that the critical coefficient,  $b_3$ , ran in the expected direction for all three items, but achieved statistical significance in only one case. To increase the statistical power of the interaction test in our small sample—the number of cases in each test averaged about 140—we reestimated the model under the constraint that all coefficients be equal across the three items. The results, shown in the top panel of Table 7, confirmed expectations: the effect of ideology is twice as large in the stop-and-think condition, a difference that is statistically significant.

We also estimated the model, under the same constraints, for respondents scoring in the bottom 40% of the awareness measure. Here we found that, as also

**Table 7. The Effect of Stop-and-Think on Ideological Consistency**

|                 | Low Awareness   | High Awareness |
|-----------------|-----------------|----------------|
| Intercept       | 0.01            | – .20          |
| Ideology        | 0.89<br>(.21)   | 0.62<br>(.14)  |
| Form            | – 0.36<br>(.21) | 0.00<br>(0.20) |
| Ideology × Form | – 0.24<br>(.27) | 0.62<br>(.26)  |
| N =             | 434             | 437            |

*Note:* Model is shown in text; model was estimated simultaneously across the jobs, services, and minority aid items. Data are taken from wave 1 of the survey, except in cases in which respondents answered “don’t know” on wave 1, in which case their wave 2 responses, if any, were used. Test was conducted for respondents scoring in the bottom 40% and top 40% of awareness scale. Measures are described in the appendix.

*Source:* 1987 NES Pilot Study.

shown in Table 7, the stop-and-think test not only failed to increase consistency, but might actually have reduced it. We shall return to this apparent reversal in a moment. But first, we have a second test of Deduction 17. If extra thought induces more reliable attitude reports, it should enhance not only correlations with ideology but also the over-time stability of these responses. As can be seen in Table 8, however, the data completely fail to support this expectation. In fact, less-aware people exhibit *less* consistency in the stop-and-think condition, while more-aware ones show no effect.

It is essential to evaluate these results in light of the selection artifact we have described. The gain in reliability among highly informed persons in Table 7 runs against the grain of the artifact and so is especially likely to be real. The null

**Table 8. The Effect of Stop-and-Think  
on Test-Retest Correlations**

|                     | Retrospective | Stop-and-Think |
|---------------------|---------------|----------------|
| Low Awareness       |               |                |
| Job guarantees      | .68<br>(40)   | .45<br>(62)    |
| Government services | .56<br>(41)   | .43<br>(58)    |
| Minority aid        | .79<br>(53)   | .53<br>(57)    |
| Middle Awareness    |               |                |
| Job guarantees      | .64<br>(29)   | .41<br>(60)    |
| Government services | .48<br>(31)   | .38<br>(51)    |
| Minority aid        | .81<br>(33)   | .51<br>(58)    |
| High Awareness      |               |                |
| Job guarantees      | .70<br>(39)   | .55<br>(51)    |
| Government services | .61<br>(37)   | .58<br>(48)    |
| Minority aid        | .70<br>(32)   | .86<br>(45)    |

*Note:* Cell entries are test-retest Pearson correlations. Numbers are in parentheses.

*Source:* 1987 NES Pilot Study.

findings and apparent reversals, because they might have been caused by the artifact, are more suspect. But might not the reversals also represent real effects of the stop-and-think treatment? Indeed, they might. In an impressive series of experiments, Tim Wilson and colleagues (Wilson et al. 1989; Wilson and Hodges 1991) have shown that, contrary to our model, asking people to articulate the reasons for their attitudes consistently *reduces* the predictive reliability of attitude reports, especially for persons less knowledgeable about the given attitude object.

The explanation for the disruptive effects of thought, as Wilson et al. maintain, is that asking people to think about the *reasons* for their attitudes causes them to sample ideas that are too heavily weighted in the direction of cognitive reactions to the attitude object rather than affective ones. Attitude reports that are based on this unrepresentative sample are, as they conclude, less reliable than reports based on the ideas that are otherwise most accessible in memory.

Note that this argument accepts the central assertion of our model, which is that people's attitude reports (and also, as the Wilson et al. studies show, behaviors) reflect the ideas that are at the top of the head at the moment of decision rather than any deeper type of "true attitude." In fact, it is precisely because attitude reports depend on immediately salient ideas that extra thought, in bringing a biased sample of ideas to the top of the head, proves disruptive.

The argument of Wilson et al. about oversampling cognitions at the expense of feelings could explain the unexpected results we obtained in Tables 7 and 8. Sniderman, Brody, and Tetlock (1991) have amassed considerable evidence that less-aware persons are more likely to base attitude reports on feelings, whereas more-aware ones tend to respond on the basis of ideological principle. It would follow from this that the stop-and-think treatment, with its relatively cognitive flavor, should prove most disruptive to the attitude reports of the less aware—which is what we have found.

Since this post hoc explanation runs in the direction of the selection artifact, we cannot be sure that it is correct. It nonetheless makes the useful point that failure to confirm (Deduction 17) may be due to a thought manipulation that failed to mimic people's natural thought processes. If so, a more carefully crafted manipulation might yet produce the reliability gain that our model anticipates.<sup>18</sup>

### Summary

The model we have proposed is, like all models, a simplification of what actually occurs. In addition, it has less formal precision than conventional

<sup>18</sup>There are, in fact, psychological studies showing that thought manipulations that stress either the affective or cognitive aspects of situations produce predictably different effects, disrupting attitude reports in some situations but not others (see Millar and Tesser 1986). For further discussion of these issues as they relate to the model, see Zaller (forthcoming, chap. 5).

measurement error models (see Achen 1975) and pays less attention to mental processes than do the attitude models of psychologists. But, despite its limitations, the simple three-axiom model is, we believe, uniquely sensitive to the wide range of empirical regularities associated with the mass survey response. Our hope, therefore, is that as others criticize our model and propose alternatives, they keep their theorizing on a sufficiently broad plane that it can accommodate the range of empirical regularities we have examined.

The empirical phenomena for which our model offers an explanation, which we summarize in Table 9, may be grouped under three general headings:

1. *Dependence of attitude reports on probabilistic memory search.* Because attitude reports are based on memory searches that are both probabilistic and incomplete, attitude reports tend to be (1) unstable over time; (2) centered on the mean of the underlying considerations; and (3) correlated with the outcomes of memory searches (Deductions 3–5). This is also why people who are more conflicted in their underlying considerations are more unstable in their closed-ended survey responses (Deduction 8).

2. *Effects of ideas recently made salient.* The notion that individuals' survey responses can be deflected in the direction of ideas made recently salient has been used to explain question order effects, endorsement effects, race-of-interviewer effects, reference group effects, question framing effects, and TV news priming effects (Deductions 9–16).

3. *Effects of thought on attitude reports.* The notion that thinking about an issue, as gauged by general levels of political awareness, enables people to recall a larger number of considerations and hence to make more reliable responses has been used to explain why more politically aware persons exhibit greater response stability and why the public as a whole is more stable on "doorstep" issues (Deductions 6, 7). It also explains why more politically aware persons, and persons especially concerned about an issue, are able to recall more thoughts relevant to it (Deductions 1, 2). Finally, the notion that greater thought makes attitude reports more reliable has been invoked, with only limited success, to explain the effects of extra thought at the moment of responding to an issue (Deduction 17).

Although it is easy to imagine alternative explanations for many of the particular phenomena we have explained with our model, it would not, we think, be easy to develop an alternative that has the range and simplicity of our model. Yet our model does, as we have admitted, represent a simplification of a process that must be much more complicated. What sorts of simplifications have we made?

One of the most important involves the issue of "on-line" processing. As indicated earlier, Hastie and Park (1986), among others, argue that people often use a "judgment operator" to update continuously their attitudes "on-line" as they acquire new information. People are said to store their updated attitudes in

**Table 9. List of Empirical Phenomena That the Model Claims to Explain**

- 
1. People who are, in general, more politically aware have more considerations at the top of their heads and available for use in answering survey questions.
  2. People who have greater interest in an issue should have, all else equal, more thoughts about that issue readily accessible in memory than other persons.
  3. There should be strong correlations between the ideas at the top of people's minds as they answer survey items and their decisions on the items themselves.
  4. There should exist a fair amount of over-time instability in people's attitude reports.
  5. Opinions that are subject to repeated measurement should have central tendencies that are stable over time, but should fluctuate around these central tendencies.
  6. The attitude reports of politically aware persons should exhibit greater over-time stability than those of less-aware persons.
  7. People should be more stable in their responses to closed-ended policy items concerning doorstep issues—that is, issues so close to everyday concerns that most people routinely pay some attention to them.
  8. Greater ambivalence ought to be associated with higher levels of response instability.
  9. Raising new considerations in immediate proximity to a question should affect the answers given by making different considerations salient.
  10. People who are ambivalent on an issue should be most affected by manipulations that raise new considerations in immediate proximity to a question about the issue.
  11. Inserting the name of a prominent politician or group into a question should affect the public's responses to the question (the "endorsement effect").
  12. The race of an interviewer should at least sometimes affect the responses to questions which he or she asks.
  13. Manipulations that raise the salience of a reference group can affect responses to questions on which the reference group has a well-known position.
  14. News reports can "prime" certain ideas, thereby making them more accessible for use in formulating attitude statements on related subjects (the "priming effect").
  15. Question order can "prime" certain ideas, thereby inducing correlations with proximate related items.
  16. Inducing individuals to think about their ideological orientation in close proximity to questions having ideological content can "prime" ideology for use in answering those questions.
  17. Inducing people to think more carefully about an issue before stating an opinion should enhance the reliability of the opinion report. (Not confirmed.)
- 

long-term memory and retrieve them as required, rather than, as in our model, create attitude statements on the spot as they confront each new survey question.

We are sympathetic to on-line information processing and believe that some form of it eventually must be included in models of political attitudes. Yet even the strongest advocates of on-line processing so far have found no evidence of such processing in the formation of policy attitudes (as against the formation of evaluations of persons). And this is scarcely surprising. The issues about which

citizens must answer survey questions are too numerous, too multidimensional, and, in most cases, too obscure for it to be feasible to engage in on-line processing of all relevant information.

### Discussion

In closing we would like to consider three broad questions about our model. Is it plausible? What are its implications for future opinion research? What are its normative implications for the role of mass opinion in a democracy?

#### *Substantive Plausibility of the Model*

Of the various claims we have made, the ambivalence axiom (along with its implication that individuals normally do not have a single, fixed, and firm attitude on issues but instead have many, potentially opposing “considerations”) is perhaps the least familiar and hence the least intuitive. It also flatly contradicts the dominant academic theories of political attitudes. Achen (1975) assumes that all respondents have “true attitudes” and maintains that these true attitudes, although appearing to vacillate because of measurement error, are overwhelmingly stable. Converse (1964) likewise assumes that, although many respondents have nonattitudes, those who do have attitudes have perfectly fixed and stable ones.

Despite the breadth of its appeal, the conventional ideal of fixed and stable “true attitudes” is well past due for strong questioning. One need only consider that, in both public opinion surveys and in elite debate, policy questions are typically framed in terms of *summary judgments* (whether abortion should be permitted or not; whether school children should be bused to promote racial integration or not) so that the making of these judgments *requires an aggregation of one's feelings across frequently diverse concerns*. There is absolutely no reason to suppose that a person must feel consistently about each of the elements that he or she aggregates across when making such summary judgments. Thus, someone who supports a woman's right to control her reproductive schedule need not also feel comfortable with aborting fetuses (Scott 1989); advocates of government aid to the needy need not feel comfortable with big government (Feldman and Zaller 1992).

The heart of our argument is that for most people, most of the time, there is no need to reconcile or even to recognize their contradictory reactions to events and issues. Each represents a genuine feeling, capable of coexisting with opposing feelings and, depending on its salience in the person's mind, controlling responses to survey questions. Analysts of public opinion long have been aware that few citizens are “ideologically consistent” in their responses to different issues (Converse 1964). Our argument is that many citizens are equally inconsistent in their reactions to different aspects of the same issue, where an “issue” is simply any bundle of concerns that a pollster asks respondents to aggregate across in the course of answering a question.

Although this line of argument may seem unobjectionable, it is easy to over-



look its implications. The most important is that individuals typically do not develop “true attitudes” of the type that opinion analysts routinely assume, but possess a series of autonomous and often inconsistent reactions to the questions asked by pollsters. Or, to put it another way, most opinions on most issues have both a central tendency and a variance.

We emphasize, however, that nothing in our model denies that some persons may develop wholly consistent sets of considerations with respect to some issues. Indeed, there is no doubt that some people, including many political activists and others who might be driven toward cognitive consistency, do exactly this. Such people are well captured by conventional “true attitude” models. But these people can be accommodated equally well by our model, which, although holding that *most* individuals are to some extent ambivalent, allows *some* individuals to be unambivalent. Thus, our model accommodates both the majority of persons who are, as Table 2 indicates, inconsistent in their reactions to diverse aspects of issues and the minority who are stable and consistent.

#### *Empirical Potential of the Model*

The value of our model for improving research on public opinion lies in three directions. One involves opening up the microfoundations of attitude reports to empirical scrutiny and analysis. For example, researchers often refer to attitudes that are more or less “crystallized,” more or less ideological, or in other ways heterogeneous across people and issues (Rivers 1988; Sniderman, Brody, and Tetlock 1991). But since the “attitude” is always the primitive term of analysis, they have had no way of directly verifying these imputations. Within our model, however, a crystallized attitude might be one that is found to be based on a larger or more homogeneous set of underlying considerations; similarly, an ideological attitude might be one based on abstract or principled considerations. The empirical consequences of these measured differences in attitudinal microfoundations could then be investigated.

Another direction for future research is in the area of communication and persuasion. Studies of political persuasion typically assume that attitude change involves a conversion experience in which one crystallized attitude structure replaces another. Our model, however, raises different possibilities, most notably “persuasion by framing.” The idea here is that if an elite or some other persuader discusses an issue so as to highlight one set of considerations rather than another, the public would be expected to respond by basing its opinion statements on the considerations thus made salient. Edelman’s (1977) and Bennett’s (1980) work on the effects on mass opinion of symbolic issue manipulation by elites, Kinder and Sanders’s (1990) work on the effects of different question frames, and Popkin’s (1991) observations on how candidates use issues to mobilize voter support all affirm the importance of manipulating mass opinion by manipulating the considerations that are salient to the public.

Traditional persuasion remains possible within our conception of attitudes, but it assumes a somewhat different form. As Zaller (1984, forthcoming) maintains, attitude change within our framework must not be understood as an all-or-nothing shift in a "true attitude," but as an adjustment in the mix of considerations relating to an issue.

Finally, the notion of a fundamentally ambivalent public can help develop more effective linkages between public opinion and the policymaking process. For example, V. O. Key (1961) found in *Public Opinion and American Democracy* that a sizable fraction of the electorate wanted stronger social welfare protections while also favoring tax cuts "even if it means putting off some important things that need to be done." In a comment much in the spirit of our model, Key observed that both attitudes, though appearing contradictory, were real. As he wrote, "a simple calculus of self-interest makes simultaneous support of tax reduction and expansion of social welfare activities entirely consistent. . . . For the system as a whole, however, this type of opinion combination is irrational and creates problems in program making. . . . The balance of forces drives policymakers back toward concealed and indirect taxation, which may be regressive in its incidence" (168).

A similar point has been made by Free and Cantril (1967), who in their study *The Political Beliefs of Americans* found that the public exhibited a combination of philosophical opposition to welfare state policies and operational support for virtually the full range of welfare programs. This "schizoid" belief pattern, Free and Cantril argued, made it extremely difficult for government to engage in rational policy-making.

Thus, we feel that a conception that emphasizes that "opinion" on an issue is generally a range of reactions rather than a single "true attitude" will be more fruitful for investigating links between mass opinion and the political process.

### *Normative Implications of the Model*

The normative implications of Converse's nonattitudes thesis are extremely bleak. In the limiting case of a public without attitudes—a claim that Converse did not actually make—self-government makes little sense. As Achen (1975, 1,227) put it, "Democratic theory loses its starting point." In contrast, the implication of Achen's empirical investigation was relatively optimistic: most members of the public have true attitudes that are almost perfectly stable. The disturbingly high levels of response change discovered by Converse, Achen argued, are due to the vague questions of survey researchers rather than to the vague minds of citizen respondents.

Our position falls somewhere between the Converse and Achen positions. We agree with Converse that there is a great deal of uncertainty, tentativeness, and incomprehension in the typical mass survey response. The problem, we further agree, is much deeper than vague questions. And yet, with Achen, we reject

the premise of Converse's black-and-white model, which is that most response fluctuation is due to essentially random guessing by people who have no meaningful opinions.

Our claim is that even when people exhibit high levels of response instability, the opinions they express may still be based on real considerations. Even when these considerations turn out to be transitory, the opinion statements they generate are not, for that reason, necessarily lacking in authenticity.

This argument extends to the interpretation of aggregate survey results. Thus, if 55% of Americans report that they approve of the way George Bush is doing his job as president, it should not be taken as evidence that a majority of the public is unequivocally supportive of the president. Rather, it should be taken to mean that 55% of Americans are *on balance* positive toward Bush's job performance—even though the particular 55% who express approval will naturally change from one survey to the next, depending on the cross-cutting accidents of many different people's memory searches (Page and Shapiro 1992).

There is, then, no inconsistency between our assertions of individual-level ambivalence and instability, on the one hand, and belief in the meaningfulness of aggregate-level poll results, on the other—provided one reads polls as revealing a *balance of considerations* rather than as counts of people's "true attitudes."

This conception of an ambivalent public may fall short of our ideal of what public opinion ought to be like, as this ideal is expressed in political oratory and democratic mythology. But if democracy is possible in a country that both glorifies economic individualism and demands the welfare state; that professes to cherish equality and practices racial discrimination; that insists on both higher levels of government services and lower taxes; and that hates Congress but re-elects congressional incumbents at extremely high rates—then it is also possible under our understanding of mass public opinion.

*Manuscript submitted 2 August 1991*

*Final manuscript received 4 December 1991*

## APPENDIX

### *Attitude Items*

*Job guarantees.* Some people think that the government in Washington should see to it that every person has a job and a good standard of living. Others think that government should just let each person get ahead on their own.

Before telling me how you feel about this, could you tell me what kinds of things come to mind when you think about *government making sure that every person has a good standard of living?* (Any others)

Now, what kinds of things come to mind when you think about *letting each person get ahead on their own?* (Any others)

*Government services.* Some people think the government should provide fewer services, even in areas such as health and education, in order to reduce spending. Other people feel it is important for the government to provide many more services even if it means an increase in spending.

Before telling me how you feel about this, could you tell me what kinds of things come to mind when you think about *fewer government services*? (Any others)

Now, what kinds of things come to mind when you think about *increases in government services*? (Any others)

*Aid to blacks.* Some people feel the government in Washington should make every effort to improve the social and economic position of blacks. Others feel that government should not make any special effort to help blacks because they should help themselves.

Before telling me how you feel about this, could you tell me what kinds of things come to mind when you think about "the social and economic position of blacks?" (Any others)

What comes to mind when you think about *efforts to improve the social and economic position of blacks*? (Any others)

And (what comes to mind) when you hear the phrase *blacks should help themselves*? (Any others)

#### *Political Awareness Scale*

A 19-point scale, having an alpha reliability of .85, as follows: v635 to v642, v202 (up to four points). Two recognition tests, v242, v244 (one item each). Four comparative candidate location items: v730 and v731; v749 and v750; v811 and v812; v831 and v832; plus v723 (a noncomparative location test).

#### *Social Welfare Ideology Scale*

The social welfare ideology scale consists of 14 items concerning individualism and equality, plus two measures of ideological self-designation, as follows: v620 to v622, v624, v626, v701 to v706, v2176, v2178, v2179, plus v722 and v1010.

## REFERENCES

- Achen, Christopher. 1975. "Mass Political Attitudes and the Survey Response." *American Political Science Review* 69:1218-31.
- . 1983. "Toward Theories of Data." Presented at the annual convention of the American Political Science Association, Chicago.
- Anderson, Norman. 1974. "Information Integration Theory: A Brief Survey." In *Advances in Mathematical Psychology*, vol. 2, ed. David Krantz et al.
- Bargh, John, R. N. Bond, W. J. Lombardy, and M. E. Tofa. 1986. "The Additive Nature of Chronic and Temporary Construct Accessibility." *Journal of Personality and Social Psychology* 50: 869-78.
- Bennett, Lance W. 1980. *Public Opinion in American Politics*. New York: Harcourt Brace Jovanovich.
- Bishop, George. 1990. "Political Involvement and Response Effects." *Public Opinion Quarterly* 53:209-18.
- Bishop, George, Robert W. Oldendick, and A. J. Tuchfarber. 1984. "What Must My Interest in Politics Be If I Just Told You 'I Don't Know'?" *Public Opinion Quarterly* 48:510-19.
- Bodenhauser, Gallen V., and Robert S. Wyer. 1987. "Social Cognition and Social Reality." In *Social Information Processing and Survey Methodology*, ed. Hans J. Hippler, Norbert Schwarz, and Seymour Sudman. New York: Springer-Verlag.

- Campbell, Angus, Philip E. Converse, Donald Stokes, and Warren Miller. 1960. *The American Voter*. New York: Wiley.
- Charters, W. W., Jr., and Theodore Newcomb. 1958. "Some Attitudinal Effects of Experimentally Increased Salience of Group Membership." In *Readings in Social Psychology*, 3d ed., ed. E. Macoby, Theodore Newcomb, and E. Hartley. New York: Holt, Rinehart.
- Converse, Philip E. 1964. "The Nature of Belief Systems in Mass Publics." In *Ideology and Discontent*, ed. David Apter. New York: Free Press.
- Converse, Philip E., and Gregory Markus. 1979. "Plus ça Change . . . : The New CPS Panel Study." *American Political Science Review* 73:32–49.
- Dean, Gillian, and Thomas Moran. 1977. "Measuring Mass Political Attitudes: Change and Unreliability." *Political Methodology* 4:383–401.
- Edelman, Murray. 1977. *Political Language*. New York: Academic Press.
- Ericsson, K. Anders, and Herbert Simon. 1984. *Protocol Analysis*. Cambridge: MIT Press.
- Erikson, Robert. 1979. "The SRC Panel Data and Mass Political Attitudes." *British Journal of Political Science* 9:16–49.
- Feldman, Stanley. 1989. "Measuring Issue Preferences: The Problem of Response Instability." *Political Analysis* 1:25–60.
- . 1990. "Theory, Measurement, and the Survey Response." Presented at the annual meeting of the American Political Science Association, Atlanta.
- Feldman, Stanley, and John Zaller. 1992. "Political Culture of Ambivalence: Ideological Responses to the Welfare State." *American Journal of Political Science*, 36: 268–307.
- Free, Lloyd, and Hadley Cantril. 1967. *The Political Beliefs of Americans*. New York: Free Press.
- Hastie, Reid, and Bernadette Park. 1986. "The Relationship between Memory and Judgment Depends on Whether the Judgment Task Is On-line or Memory Based." *Psychological Review* 93: 258–68.
- Hastie, Reid, and Nancy Pennington. 1989. "Notes on the Distinction between Memory-based and On-line Judgments." In *On-line Cognitive Person Perception*, ed. John Bassili. Hillsdale, NJ: Erlbaum.
- Higgins, E. Tory, and Gary A. King. 1981. "Accessibility of Social Constructs: Information-Processing Consequences of Individual and Contextual Variability." In *Personality, Cognition, and Social Interaction*, ed. Nancy Cantor and John F. Kihlstrom. Hillsdale, NJ: Erlbaum.
- Hippler, Hans J., and Norbert Schwarz. 1987. "Response Effects in Surveys." In *Social Information Processing and Survey Methodology*, ed. Hans J. Hippler, Norbert Schwarz, and Seymour Sudman. New York: Springer-Verlag.
- Hochschild, Jennifer. 1981. *What's Fair? Americans' Attitudes toward Distributive Justice*. Cambridge: Harvard University Press.
- Iyengar, Shanto. 1991. *Is Anyone Responsible?* Chicago: University of Chicago Press.
- Iyengar, Shanto, and Donald Kinder. 1987. *News That Matters*. Chicago: University of Chicago Press.
- Judd, Charles, and Cynthia Lusk. 1984. "Knowledge Structures and Evaluative Judgments: Effects of Structural Variables on Attitude Extremity." *Journal of Personality and Social Psychology* 47:331–38.
- Judd, Charles, and Michael Milburn. 1980. "The Structure of Attitude Systems in the General Public: A Comparison of a Structural Equations System." *American Sociological Review* 45:627–43.
- Judd, Charles, Michael Milburn, and Jon Krosnick. 1981. "Political Involvement and Attitude Structure in the General Public." *American Sociological Review* 46:660–69.
- Kelley, Stanley, Jr. 1983. *Interpreting Elections*. Princeton: Princeton University Press.
- Key, V. O., Jr. 1961. *Public Opinion and American Democracy*. New York: Knopf.
- Kinder, Donald, and Lynn Sanders. 1990. "Mimicking Political Debate with Survey Questions: The Case of White Opinion on Affirmative Action for Blacks." *Social Cognition* 8:73–103.

- Krosnick, Jon. 1988. "Attitude Importance and Attitude Change." *Journal of Experimental Social Psychology* 24:240–55.
- Krosnick, Jon, and Howard Schuman. 1988. "Attitude Intensity, Importance, and Certainty and Susceptibility to Response Effects." *Journal of Personality and Social Psychology* 54:940–52.
- Lane, Robert. 1962. *Political Ideology*. New York: Free Press.
- Lau, Richard, and David O. Sears. 1983. "Inducing Apparently Self-interested Political Preferences." *American Journal of Political Science* 27:233–52.
- Lau, Richard, David O. Sears, and Tom Jessor. 1990. "Fact or Artifact Revisited: Survey Instrument Effects and Pocketbook Politics." *Political Behavior* 12:217–42.
- Lewis-Beck, Michael. 1985. "Pocketbook Voting in U.S. National Elections: Fact or Artifact?" *American Journal of Political Science* 29:348–56.
- Lodge, Milton, Kathleen McGraw, and Patrick Stroh. 1989. "An Impression-Driven Model of Candidate Evaluation." *American Political Science Review* 83:399–420.
- Loftus, Elizabeth F., Stephen E. Feinberg, and Judith M. Tanur. 1985. "Cognitive Psychology Meets the National Survey." *American Psychologist* 40:175–80.
- Milburn, Michael. 1987. "Ideological Self-schemata and Schematically-Induced Attitude Consistency." *Journal of Experimental Social Psychology* 23:383–98.
- Millar, Murray, and Abraham Tesser. 1986. "Effects of Affective and Cognitive Focus on Attitude-Behavior Relationship." *Journal of Personality and Social Psychology* 51:270–76.
- . 1989. "The Effects of Affective-Cognitive Consistency and Thought on the Attitude-Behavior Relation." *Journal of Experimental Social Psychology* 25:189–202.
- Mueller, John. 1973. *War, Presidents, and Public Opinion*. New York: Wiley.
- Page, Benjamin, and Robert Shapiro. 1992. *The Rational Public*. Chicago: University of Chicago Press.
- Popkin, Sam. 1991. *The Reasoning Voter*. Cambridge: Cambridge University Press.
- Price, Vincent. 1991. "Priming Ideological Leanings: The Impact of Salience and Commitment." Report to the National Election Studies Board of Overseers, Ann Arbor, Michigan.
- Raaijmakers, J., and Richard Shiffrin. 1981. "Search of Associated Memory." *Psychological Review* 88:93–134.
- Rivers, H. Douglas. 1988. "Heterogeneity in Models of Voter Choice." *American Journal of Political Science*.
- Schuman, Howard, and Stanley Presser. 1981. *Questions and Answers in Attitude Surveys*. New York: Academic Press.
- Schuman, Howard, and Jacqueline Scott. 1987. "Problems in the Use of Survey Questions to Measure Public Opinion." *Science* 236: 957–59.
- Scott, Jacqueline. 1989. "Conflicting Beliefs about Abortion: Legal Approval and Moral Doubts." *Social Psychology Quarterly* 52: 319–26.
- Sears, David O., and Richard Lau. 1983. "Inducing Apparently Self-interested Political Responses." *American Journal of Political Science* 27:223–52.
- Sniderman, Paul, Richard Brody, and Philip Tetlock. 1991. *Reasoning and Choice*. Cambridge: Cambridge University Press.
- Sussman, Barry. 1986. "Do Blacks Approve of Reagan? It Depends on Who's Asking." *Washington Post* weekly ed. 10 February.
- Taylor, Shelley E., and Susan Fiske. 1978. "Salience, Attention, and Attribution: Top of the Head Phenomena." In *Advances in Social Psychology*, ed. Leonard Berkowitz. New York: Academic Press.
- Tesser, Abraham. 1978. "Self-generated Attitude Change." In *Advances in Social Psychology*, ed. Leonard Berkowitz. New York: Academic Press.
- Tourangeau, Roger, and Kenneth Rasinski. 1988. "Cognitive Processes Underlying Context Effects in Attitude Measurement." *Psychological Bulletin* 103:299–314.

- Tourangeau, Roger, Kenneth Rasinski, Norman Bradburn, and Roy D'Andrade. 1989. "Carryover Effects in Attitude Surveys." *Public Opinion Quarterly* 53:495–524.
- Tversky, Amos, and Daniel Kahneman. 1982. "The Framing of Decisions and the Psychology of Choice." In *Question Framing and Response Consistency*, ed. Robin Hogarth. San Francisco: Jossey-Bass.
- Wilson, Timothy, Dana S. Dunn, Dolores Kraft, and D. Lisle. 1989. "The Disruptive Effects of Explaining Why We Feel the Way We Do." In *Advances in Social Psychology*, vol. 19, ed. Leonard Berkowitz. Orlando, FL: Academic Press.
- Wilson, Timothy, and Sarah Hodges. 1991. "Attitudes as Temporary Constructs." In *The Construction of Social Judgment*, ed. Abraham Tesser and L. Martin. Hillsdale, NJ: Erlbaum.
- Wyer, Robert, and Jon Hartwick. 1984. "The Recall and Use of Belief Statements as Bases for Judgments." *Journal of Experimental Social Psychology* 20:65–85.
- Wyer, Robert, and Thomas Srull. 1989. *Memory and Cognition in Its Social Context*. Hillsdale, NJ: Erlbaum.
- Zaller, John. 1984. "Toward a Theory of the Survey Response." Presented at the annual meeting of the American Political Science Association, Washington, DC.
- . 1990. "Political Awareness, Elite Opinion Leadership, and the Mass Survey Response." *Social Cognition* 8:125–53.
- . 1992. *The Nature and Origins of Mass Opinion*. Cambridge: Cambridge University Press. Forthcoming.