SSOAR

Open Access Repository
www.ssoar.info

Cognition, communication, and survey measurement: some implications for contingent valuation surveys

Schwarz, Norbert

Veröffentlichungsversion / Published Version
Arbeitspapier / working paper

Zur Verfügung gestellt in Kooperation mit / provided in cooperation with:
GESIS - Leibniz-Institut für Sozialwissenschaften

Empfohlene Zitierung / Suggested Citation:
Schwarz, N. (1994). *Cognition, communication, and survey measurement: some implications for contingent valuation surveys.* (ZUMA-Arbeitsbericht, 1994/09). Mannheim: Zentrum für Umfragen, Methoden und Analysen -ZUMA-. https://nbn-resolving.org/urn:nbn:de:0168-ssoar-70186

Nutzungsbedingungen:
Dieser Text wird unter einer Deposit-Lizenz (Keine Weiterverbreitung - keine Bearbeitung) zur Verfügung gestellt. Gewährt wird ein nicht exklusives, nicht übertragbares, persönliches und beschränktes Recht auf Nutzung dieses Dokuments. Dieses Dokument ist ausschließlich für den persönlichen, nicht-kommerziellen Gebrauch bestimmt. Auf sämtlichen Kopien dieses Dokuments müssen alle Urheberrechtshinweise und sonstigen Hinweise auf gesetzlichen Schutz beibehalten werden. Sie dürfen dieses Dokument nicht in irgendeiner Weise abändern, noch dürfen Sie dieses Dokument für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, aufführen, vertreiben oder anderweitig nutzen.
Mit der Verwendung dieses Dokuments erkennen Sie die Nutzungsbedingungen an.

Terms of use:
This document is made available under Deposit Licence (No Redistribution - no modifications). We grant a non-exclusive, non-transferable, individual and limited right to using this document. This document is solely intended for your personal, non-commercial use. All of the copies of this documents must retain all copyright information and other information regarding legal protection. You are not allowed to alter this document in any way, to copy it for public or commercial purposes, to exhibit the document in public, to perform, distribute or otherwise use the document in public.
By using this particular document, you accept the above-stated conditions of use.

gesis Leibniz-Institut für Sozialwissenschaften

Mitglied der Leibniz-Gemeinschaft

ZUMA-Arbeitsbericht 94/09

Cognition, Communication, and Survey Measurement: Some Implications for Contingent Valuation Surveys

Norbert Schwarz

August 1994

ZUMA
Quadrat B2,1
Postfach 12 21 55
D-68072 Mannheim

Telefon: (0621) 1246 - 0
Telefax: (0621) 1246 - 100
E-mail: Name@zuma-Mannheim.de

Cognition, Communication, and Survey Measurement: Some Implications for Contingent Valuation Surveys

Norbert Schwarz

University of Michigan

Paper prepared for *"Determining the Value of Non-Marketed Goods: Economic, Psychological, and Policy Relevant Aspects of Contingent Valuation Methods"*, conference at the Werner Reimers Stiftung, Bad Homburg, Germany, July 1994 (W. Pommerehne & N. Schwarz, co-conveners). This paper was completed while the author was a guest professor at ZUMA. Address correspondence to Norbert Schwarz, Institute for Social Research, University of Michigan, Ann Arbor, MI 48106-1248, USA.

Cognition, Communication, and Survey Measurement: Some Implications for Contingent Valuation Surveys

Contingent valuation (CV) surveys provide a methodological approach for determining the value of non-marketed goods or services (see Mitchell & Carson, 1989, for a discussion of their theoretical and methodological rationale). In a nutshell, respondents are provided with a detailed description of the to-be-valued good or service and are asked to report what they would be willing to pay for it. The value inferred from their responses is contingent upon the nature of the constructed hypothetical market and the good or service described in the scenario, hence the term "contingent valuation". CV surveys are primarily used to assess the passive use values of natural resources and their results are used in cost-benefit analysis and natural resource damage assessment. The first CV study has been published in 1963 and more than 1,400 studies have been documented since.

Despite the frequent use of CV in natural resource valuation, there are no generally agreed upon guidelines for the design of CV studies. Moreover, the use of CV methodology in assessing the damage inflicted by the Exxon Valdez oil spill in Alaska (see Carson et al., 1992) spurred a considerable amount of research, mainly sponsored by Exxon, that questioned the validity of CV findings (see the contributions in Hausmann, 1993). As a result, the U.S. National Oceanic and Atmospheric Administration (NOAA) and the U.S. Department of the Interior are currently developing regulations to guide the future use of CV surveys for purposes of cost-benefit analysis and damage assessment. The crucial role of cost-benefit analysis in U.S. environmental policy (see Kopp & Pease, 1994), on the one hand, and the monetary implications of estimated passive use values in natural resource damage litigation on the other hand, contributed to a heated debate that far exceeds any other

controversy about the methodology of social science research. In the present paper, I address the key issues raised in this debate by drawing on recent psychological research into the cognitive and communicative processes underlying survey measurement.

Challenges to CV Research

Recent theoretical and methodological discussions of CV surveys have focused on the observation that "CV responses are sensitive to methodological factors that, from the standpoint of economic theory, are irrelevant, as well as insensitive to some theoretically relevant factors" (Schkade & Payne, 1993, p. 273). Underlying much of this discussion are the implicit assumptions of what Fischhoff (1991) called a "philosophy of articulated values." According to this meta-theoretical orientation -- which is usually attributed to economists, but is shared by many survey researchers as well (see Strack & Schwarz, 1992, for a discussion of meta-theories in survey research) -- people have well-formed preferences about (nearly) any topic, which they can retrieve from memory and report when asked. Although some economists have rejected this assumption as a caricature of economic theorizing (see the discussion contributions in Hausman, 1993), it has, more or less explicitly, provided the standard for measuring the performance of CV studies.

An alternative approach is based on what Fischhoff (1991) called the "philosophy of basic values". This orientation, which is dominant in psychology and decision making, assumes that people have well-defined values only for very few, and very familiar, topics of high importance. In most cases, they have to construct their preferences when asked. Importantly, these constructions depend on the information considered at that point in time, rendering them highly context dependent. From a psychological point of view, this context dependency is a normal feature of human judgment that is observed in daily behavior and

real economic decision making (see Payne, Bettman, & Johnson, 1992; Thaler, 1992) as well CV responses.

Given their different meta-theoretical assumptions, both approaches differ in their evaluation of the context dependency of reported preferences. From an "articulated values" point of view, the context dependency of reported preferences indicates that the measurement methods used are inappropriate for assessing peoples' "true" preferences and need revision until context-independency is reached (if ever). In contrast, the "basic values" point of view suggests that it is the researcher's task to structure the instrument in a way that helps people in constructing meaningful preferences, e.g., by drawing their attention to pertinent information, providing a relevant frame of reference, and so on. Fischhoff and Furby's (1988) extensive discussion of things to do and not do in CV studies illustrates this approach.

In the specific case of CV research, however, the issues seem more complex than in the assessment of other preferences, lets say political preferences in an opinion survey. Survey researchers, for example, want to assess whatever the opinion of people out there is, but are rarely concerned with the rationality of that opinion relative to some normative model. For them, the key criterion to validate reported preferences for a political candidate, for example, is actual voting behavior in an election. If people vote for the candidate for the "right" or "wrong" reasons is of no concern. Not so in the recent debate about CV responses. In this case, observing that a CV survey correctly predicts the outcome of an actual referendum would not satisfy many of the critics. Rather, the compatibility of CV responses with normative economic models assumes crucial importance, shifting the criteria from how people vote on a referendum to whether they vote the way they do for the "right" reason (see the discussions in Hausman, 1993).

Although noneconomists may find it surprising that models that are known to be

wrong as descriptions of real economic behavior can serve as the gold standard for contingent economic behavior, addressing the intricacies of this issue is beyond the scope of my expertise. Two issues, however, are crucial to the current debate and I shall address them by drawing on what we know about how people answer survey questions. These issues are, that CV responses are (a) insensitive to factors that are deemed theoretically relevant, but (b) sensitive to factors that are deemed irrelevant. Taking a mental construal perspective, I argue that proponents as well as opponents of current CV methods need to pay more attention to respondents' construal of the to-be-valued good. Based on related research in survey methodology and attitude measurement, I suggest that many apparently irrelevant features of the questionnaire implicitly provide information that is relevant to the valuation task, whereas much of the explicitly provided information may never find its way into respondents' representations of the good. Based on these considerations, I outline relevant research and discuss methodological implications. Before I turn to these issues, however, it is necessary to review the question answering process in surveys.

Cognition, Communication, and Survey Measurement

That survey data are only as meaningful as the answers that respondents provide has long been recognized by survey researchers. Nevertheless, the cognitive and communicative processes that underlie question answering in the survey context have only recently received theoretical and empirical attention. Drawing on psychological theories of language comprehension, memory, and judgment, psychologists and survey methodologists have begun to formulate explicit models of the question answering process and have tested these models in laboratory experiments and split-ballot surveys. This work links survey methodologists' expertise in the "art of asking questions" to recent developments in cognitive and social

psychology, thus providing a useful theoretical and empirical basis for understanding the processes by which survey respondents arrive at an answer (see the contributions in Hippler, Schwarz, & Sudman, 1987; Schwarz & Sudman, 1992, 1994; Tanur, 1992).

Much of the work on cognitive aspects of survey responding has focused on attitude measurement and it is informative to highlight some of the differences between traditional attitude surveys and CV surveys. From a cognitive perspective, answering an attitude question requires that respondents solve several tasks, on which researchers have reached wide agreement (see, for example, Schwarz & Strack, 1991; Strack & Martin, 1987; Tourangeau & Rasinski, 1988). Although these tasks are plausibly presented in a sequential order, the actual order in which respondents complete the tasks may occasionally differ from the assumed typical sequence.

Question Comprehension

As a first step, respondents have to interpret the question to understand what is meant. If the question is ambiguous, respondents draw on the context of the question to disambiguate its meaning. Relevant contextual information includes the content of preceding questions (see Schuman & Presser, 1981; Schwarz & Strack, 1991, for reviews) as well as apparently formal features of the questionnaire (see Schwarz & Hippler, 1991; Schwarz, 1994, for reviews). Not surprisingly, different substantive interpretations of a question result in different answers. What is more surprising to many researchers is that different substantive interpretations can be triggered by formal features of the questionnaire that are usually not supposed to influence question meaning.

As an example, suppose that respondents are asked to rate how successful they have been in life along a scale from "not at all successful" to "extremely successful." To answer

this question, respondents must determine what the researcher means by "not at all successful". Does this term refer to the absence of outstanding achievements or to the presence of explicit failures? Without additional explanation, respondents may turn to relevant features of question context, making use of what the researcher may consider an irrelevant technical feature of the questionnaire. One such feature are the specific numeric values presented as part of a rating scale. According to textbook knowledge, ratings are affected by the extremity of the verbal endpoint labels and the number of scale points provided, but an 11-point scale is an 11-point scale, no matter whether it presents 11 unnumbered boxes or numeric values that run from 1 to 11, from 0 to 10, or from -5 to +5. Trying to make sense of the specific meaning of the verbal label, however, respondents may turn to such apparently irrelevant factors in interpreting the meaning of the question. Thus, we observed in several studies that "not at all successful" was interpreted as reflecting the absence of outstanding achievements when the rating scale ran from 0 ("not at all successful") to 10 ("very successful"), but was interpreted as reflecting the presence of explicit failure when the rating scale ran from -5 ("not at all successful") to +5 ("very successful"), presenting 0 as the neutral midpoint. As a result of this differential interpretation of the verbal anchor, 34 percent of the respondents endorsed a value below the midpoint of the 0 to 10 scale, whereas only 13 percent endorsed one of the formally equivalent values of the -5 to +5 scale (Schwarz, Knäuper, Hippler, Noelle-Neumann, & Clark, 1991). As this example illustrates, design features that are often considered "nonsubstantive" may profoundly change the meaning of a question. As a result, many influences of apparently "irrelevant" features of questionnaire design can be traced to their impact on respondents understanding of the substantive meaning of the question.

The use of contextual information that is reflected in such findings (see Schwarz,

1994, for a review) is licensed by the tacit rules that govern the conduct of conversation in daily life, as described in Grice's (1975) cooperativeness principle (see Clark & Schober, 1992; Schwarz, 1994; Strack & Schwarz, 1992, for applications to survey research). According to these tacit rules, "communicated information comes with a guarantee of relevance" as Sperber and Wilson (1986, p. vi) noted, and listeners interpret speakers' utterances on the assumption that the speaker is trying to be informative, relevant and clear. If the speaker does not live up to these ideals, listeners may ask for clarification (which they are unlikely to get from the interviewer) or may use related contributions of the speaker to determine the intended meaning. In the survey interview, the contributions of the researcher to the ongoing "conversation" include the content of preceding questions as well as formal features of the questionnaire, rendering them highly relevant as contextual information.

As in other domains of discourse, however, question comprehension requires more than an understanding of the literal meaning of the utterance. Rather, respondents have to identify the pragmatic meaning of the question to determine which information the questioner asks them to provide (Clark & Schober, 1992; Schwarz, 1994). A simple example may again illustrate this point. When asked, "What have you done today?", respondents are likely to understand the literal meaning of the question -- but which information are they supposed to provide? What are the behaviors that the researcher is interested in? Should they report, for example, that they took a shower or not? When the question is asked in an open response format, chances are that hardly any respondent would report taking a shower. But when "taking a shower" is presented as one of the response alternatives in a closed question format, most respondents are likely to endorse it (see Schuman & Presser, 1981; Schwarz & Hippler, 1991).

Findings of this type again reflect that survey respondents rely on the tacit rules that

govern the conduct of conversation in daily life (Grice, 1975). These rules ask speakers to provide information that is "informative" for the recipient, rather than information that the recipient may take for granted anyway. Accordingly, information that seems self-evident, that may be assumed to "go without saying", or that has already been provided in response to a previous question is unlikely to be reported unless the researcher explicitly indicates that it is of interest (e.g., by presenting appropriate response alternatives). This hesitancy to report things that "go without saying" has potentially important implications for questions designed to check respondents' understanding of CV scenarios, as we shall see below.

Recall and Judgment

Once respondents made sense of an attitude question, they presumably recall the relevant opinion from memory. Or at least this is what attitude researchers usually hope for (in line with an articulated values philosophy), reflecting their disdain for what has often been called "door step opinions" in public opinion research. In most cases, however, respondents are unlikely to have a previously formed judgment accessible in memory, in particular a judgment that matches the specific aspects of the issue addressed in the question. Accordingly, respondents need to "compute" a judgment on the spot (as assumed by a basic values philosophy). To do so, they draw on relevant information that is accessible at this point in time to form a mental representation of the target (see Schwarz & Bless, 1992; Schwarz & Strack, 1991; Tourangeau, 1992; Tourangeau & Rasinski, 1987).

Which information is accessible depends on respondents' knowledge and the context of the questionnaire. Some information may always come to mind when a given respondent thinks of a given target, and is hence called <u>chronically accessible</u>. Other information, however, may only come to mind because it has been used just recently, usually to answer

a preceding question. Such temporarily accessible information is the source of most context effects in attitude measurement. The larger the amount of temporarily accessible information relative to chronically accessible information, the larger the size of context effects (see Schwarz & Bless, 1992, for a more detailed discussion). Accordingly, attitude researchers would like respondents to draw solely on chronically accessible information in forming a representation of the target. This follows from their desire to generalize from the opinions reported by their sample to the opinions of a population that has not been exposed to the context of the questionnaire. Accordingly, any influence of the information rendered accessible by the questionnaire is deemed undesirable because such influences would lead to context dependent reports that deviate from the opinions presumably held in the population.

What most researchers hope for in conducting a CV survey, on the other hand, seems quite different. In fully developed CV surveys (see Mitchell & Carson, 1989; Fischhoff & Furby, 1988) respondents are supposed to value a specific good that is described to them in considerable detail in the survey. In most cases, they have no experience with the good. In fact, they are usually asked to value some imagined environmental change for which any previously computed value cannot be available. Accordingly, CV respondents are supposed to rely on the information provided in the questionnaire to arrive at a mental representation of the to-be-valued good. To the extent that they draw on information that is not contained in the scenario, they may construct a good that differs from the good the researcher wants them to value, rendering their reported WTP irrelevant to the goals of the study. Below, I return to these issues in more detail. For the time being, however, I note that the procedures underlying fully developed CV surveys clearly follow a basic values philosophy (despite occasional disclaimers, see Mitchell & Carson, 1989).

Many studies reported by CV critics, on the other hand, follow a methodology that

is more in line with the assumption of an articulated values philosophy. Consistent with what attitude researchers would do, respondents in these studies are provided with descriptions of the good of less than 10 words in length, aptly described as following "the headline method" (Kahneman & Ritov, 1994). The better of these descriptions meet the criteria for good attitude questions and are sufficient to identify the issue, which should allow respondents to retrieve a previously formed preference, assuming such a preference is accessible. From a construal point of view, however, these descriptions provide insufficient specifications of the good, rendering it impossible to tell what exactly respondents are valuing.

Forming an evaluative judgment, however, does not only require a mental representation of the target of the question (the attitude object or good), but also a mental representation of some standard against which the target is evaluated. Much as the representation formed of the target, the representation formed of a standard is context dependent (Schwarz & Bless, 1992) and may include temporarily as well as chronically accessible information. In general, attitude researchers again hope that respondents draw primarily on chronically accessible information in constructing a relevant standard, for the reasons discussed above. In contrast, CV researchers frequently remind respondents of other public goods to provide a framework for the valuation task, again reflecting a basic values approach.

As these comparisons indicate, it is the use of <u>context dependent</u> information, rendered accessible by the questionnaire, that is considered problematic in attitude measurement. In contrast, it is the use of <u>context independent</u> information that is considered problematic in CV surveys, although this issue is rarely addressed in these terms. Given that judgments are likely to be based on <u>both</u> sources of information in either case, attitude researchers and CV researcher are interested in minimizing the impact of different sources

of information.

Formatting

Once respondents have formed a "private" judgment in their mind, they need to communicate it to the researcher. Unless the question uses an open-response format, respondents have to format their judgment to fit the response alternatives provided as part of the question. Attitude researchers as well as CV researchers usually think of response alternatives as simple "measurement devices." As long as the response alternatives allow respondents to express their judgment, they are not considered to influence the obtained results. This assumption, however, is mistaken as the examples reviewed in the context of question comprehension illustrated. The response alternatives constitute an important source of information that respondents draw on in making sense of the question and in constructing mental representations of the target or standard (see Schwarz & Hippler, 1991). Below, I address the implications of this perspective for different forms of WTP questions.

Reporting the Answer

Finally, respondents may wish to edit their response before they report it, due to influences of social desirability and situational adequacy (see DeMaio, 1984). In this regard, attitudes as well as CV researchers attempt to safeguard against influences of social desirability and survey methodologists developed various procedures that reduce these influences (see Sudman & Bradburn, 1983).

Modes of Data Collection

In summary, interpreting the question, generating an opinion, formatting the response,

and editing the answer are the main psychological components of a process that starts with respondents' exposure to a survey question and ends with their overt report. Several aspects of this process are influenced by the mode of data collection used, resulting in differences between face-to-face or telephone interviews and self-administered questionnaires (see Schwarz, Strack, Hippler, & Bishop, 1991). In the final section of this paper, I address some of the relevant mode differences and their implications for the use of mail surveys in contingent valuation research.

Mental Representations and Scope Effects: What are Respondents Valuing?

"According to consumer theory, for any good, more of that good is preferred to less. Therefore, a scenario offering a higher level of services should yield a higher value" (Desvouges, Hudson, & Ruby, 1994, p. 8), at least if the good has positive marginal utility. Much of the recent debate has focused on the extent to which the responses obtained in CV surveys meet this assumption of economic theorizing. Several studies demonstrated that reported WTP can be insensitive to variations in the scope of the described good, whereas others observed satisfactory sensitivity (see Carson & Mitchell, 1993, in press, for reviews). The crucial question is what demonstrations of insensitivity to scope actually demonstrate: that a given specific survey instrument is insensitive or that CV methodology in general is insensitive to scope? Whereas the former conclusion suggests that the specific instrument needs improvement, the latter conclusion suggests that respondents' valuations may be based on processes that are incompatible with the economic framework that provides the theoretical justification for CV methods (see Carson & Mitchell, 1993 and Kahneman & Knetsch, 1992, for opposing perspectives).

In addressing this question, we face the usual problems involved in interpreting null results, in particular in the presence of other studies that obtained the predicted differences. Psychologists' usual response to null results is to demonstrate (a) that the appropriate experimental conditions have been realized, based on proper manipulation checks, and/or (b) that some other dependent variable shows a pattern that plausibly supports that the experimental manipulation was successful. In the absence of such evidence, psychologists usually refrain from drawing substantive conclusions from null results. Note, however, that making a case for the interpretability of null results does not require more than demonstrating the successful realization of the experimental conditions and appropriate statistical power. As Fischhoff (this conference) noted, some CV researchers advocate standards that raise the cost of methodological studies to a level that renders basic research unfeasible. Whereas these criteria may be appropriate for damage assessments, where the absolute value of WTP is of key interest, many of them (including representativeness, sample size, and related variables) are overly restrictive for methodological studies interested in relative differences between experimental conditions.

What, however, does it take to demonstrate that differences in scope have been successfully realized in a scenario? At the minimum, it requires a manipulation check that indicates that respondents do indeed expect that (a) different quantities of (b) the same quality of the described good will (c) be delivered with equal likelihood, if proper payment is provided. This requires that respondents understand the differences in quantity conveyed in the scenario; that their mental construal of the good is not more or less inclusive than intended in the scenario; that the scope does not affect the perceived likelihood of delivery or the perceived quality of what is delivered; and so on (see Fischhoff & Furby, 1988, for an extensive discussion of relevant criteria). Unfortunately, it is unclear to what extent such

prerequisites were met in many of the better known demonstrations of embedding effects -- as well as in the studies that demonstrated sensitivity to differences in scope.

Do Respondents Understand the Description of Scope?

Obviously, differences in scope described in the scenario will only affect the obtained responses if they find their way into respondents' own mental representations of the good. This requires that respondents notice the scope described in the scenario and that this description makes sense to them. In general, respondents assume that the space and time allocated to an issue in the questionnaire reflects its relative importance (Fischhoff & Furby, 1988; Schwarz, 1994). Accordingly, scope information needs to be elaborated on, rather than mentioned in passing, and respondents' attention needs to be drawn to it.

To illustrate this rather trivial point, consider two studies reported by Desvouges et al. (1993) and Schkade and Payne (1993), based on the same scenarios. Using a self-administered questionnaire in a mall intercept survey, Desvouges et al. asked respondents to value a program that would prevent the deaths of either 2,000, 20,000, or 200,000 migratory waterfowl from oil ponds in the Rocky Mountain flyway. Respondents were further informed that 2,000 birds represent "much less" than 1% of the population, 20,000 birds "less" than 1%, and 200,000 birds "about 2%". The results showed that the described scope had no impact on reported WTP, which seems surprising if one draws on differences in the absolute number of birds involved, but not if one draws on the minor differences in the percentage of birds involved. Interestingly, however, the difference between 2,000 and 200,000 birds resulted in differential median WTP values of $25 and $50, respectively (or $25 and $60, $p = .11$, when zero answers were excluded) when Schkade and Payne used exactly the same scenario, but asked respondents to think aloud while answering the

questions. Although these differences failed to reach significance with the small sample size used, the pattern of WTP suggests that the think-aloud task fostered attention to detail, resulting in the elaboration of the scope information provided.

As this example illustrates, scope information requires appropriate encoding, which needs to be ensured by procedures that draw respondents' attention to scope information and encourage its active elaboration. This may be accomplished by questions that stimulate respondents to draw on the scope information during the information acquisition phase and to incorporate scope information in their mental representations of the good. That this can be accomplished even for information that is notoriously difficult to convey, such as information about low-level risks, has been successfully demonstrated in several studies (e.g., Mitchell & Carson, 1986). Simply observing that scope information does not affect reported WTP, however, is insufficient to challenge CV methodology at a basic level as the observed insensitivity may always reflect an unsuccessful manipulation that should be addressed in a revision of the instrument.

Inclusive Representations

A different problem arises when differences in scope are introduced by respondents' implicit assumptions about the intervention that is supposed to provide the relevant good. This problem increases with respondents' knowledge about the content domain and decreases, at least to some extent, with the explicitness of the scenario. As an illustration, consider the construals of highly sophisticated respondents in a study reported by Schulze, McClelland, and Lazo (1994). According to economic models, the jointly assessed value of preserving a species of blue winged <u>and</u> a species of green winged butterflies in the Amazon rain forest should be equivalent to the sum of the values assigned to preserving only the blue winged

butterflies or only the green winged butterflies, when each species is valued by itself (assuming that the two are not substitutable). This prediction presupposes, however, that respondents do indeed treat the preservation of both species as unrelated. Yet, sophisticated respondents may be unlikely to do so, as focus group results reported by Schulze et al. (1994) illustrate. Drawing on their knowledge about the interdependence of environmental processes, sophisticated respondents may reason,

> "Butterfly species in the Amazon are becoming extinct because of a loss of habitat. The only way to save one species is to save all of them by saving the forest as well" (Schulze et al. 1994, p. 16).

As a result, respondents' representation of the to-be-valued good includes much more than the butterfly species that is to be saved, resulting, in the present case, in a valuation of the complete forest along with an unknown number of other species. Not surprisingly, asking these respondents to value the preservation of several species in the same habitat does not result in higher WTP, reflecting that the valuation of any species includes the value of all other species given the respondents' mental model of the content domain.

To which extent such joint valuations underlie non-additive valuations is an empirical issue. The embedding effect that results from joint valuations, however, does not violate normative assumptions. Rather, it reflects that respondents in the narrow scope condition are valuing a more inclusive good than the researcher intended them to begin with, thus rendering explicit extensions in scope uninfluential. Accordingly, it is crucial to understand what knowledge respondents draw on in forming a mental representation of the good. Because this knowledge is domain dependent, its exploration presents a formidable task.

Of course, better CV surveys usually try to assess what respondents have been valuing after WTP is assessed. Based on what we learned from cognitive research into question

comprehension in surveys, however, such questions are not guaranteed to uncover diverging construals of the to-be-valued good. In line with conversational norms, respondents may not elaborate on information they consider self-evident, assuming that some things go without saying (Clark, 1985). To return to Schulze et al.'s (1994) example, a respondent who assumes that butterflies can only be saved by saving their habitat, may also assume that the researcher, who is presumably an expert on the issue, shares this obvious insight. When asked, she may hence respond that she valued the butterfly species, taking it for granted that this implies whatever it obviously takes to reach the goal. One way to address this issue is to explore respondents' reasoning processes in think-aloud studies during the pretesting phase and to develop specific queries for inclusion in the main survey. Much as in the case of the "What-have-you-done-today?" example discussed above, such specific probes would be likely to uncover assumptions that respondents take for granted, although these assumptions would not be reported in response to a more general question.

Scope and Certainty of Provision

Similar concerns apply to assessments of the scenario's plausibility and the perceived likelihood that the promised good is actually provided. As Fischhoff and Furby (1988, p. 162) noted, "Agreeing to a proposed transaction involves an exchange of promises: to deliver a good and to make a payment. As a result, two defining aspects of any agreement are the probabilities of each promise being kept. The less likely a good is to be received, the less an offer to provide it should be valued."

In general, CV researchers want respondents to assume that provision of the good is certain. To the extent that respondents' perceived likelihood of provision is lower, they undervalue the good relative to the researcher's intended interpretation of the scenario.

Although the perceived certainty of provision is affected by many aspects of the scenario (see Mitchell & Carson, 1989), a particularly crucial one is the described scope: The larger the scope, the less likely it may seem that the promised good is actually delivered. Tempting as an investment plan that promises an annual return of 200% may be, we are likely to put more money into a more conservative plan, reflecting that exaggerated promises raise doubts -- in particular if the relevant mechanism of provision does not strike one as plausible. Similar concerns may apply to a plan that promises, for example, to "Rehabilitate all recently released criminals," as opposed to "Rehabilitate recently released young offenders" (see Kahneman & Knetsch, 1992, p. 65, for relevant examples). To the extent that the certainty of provision seems lower in the former case than in the latter, we shouldn't expect that people pay more for the former plan.

In considering respondents' perception of the certainty of provision it is helpful to distinguish between the provision of the promised intervention and the provision of the good that the intervention is supposed to produce or protect. As the rehabilitation example illustrates, both are likely to be affected by scope. On the one hand, the implementation of a rehabilitation program directed at all released criminals may seem less realistic than a program directed at the smaller subset of young offenders. On the other hand, respondents may assume that "young offenders" are more likely to be successfully rehabilitated than seasoned "criminals", as may already be indicated by the respective terms. If so, they may assume that the smaller program delivers more for the money than the larger program. Moreover, one may wonder if respondents spontaneously include young offenders in their mental representation of criminals, reiterating the concerns about the mental representation of scope addressed above.

Again, respondents' general acceptance of the scenario and their perception of

provision is assessed in the better studies, usually after respondents reported their WTP. In many cases, these acceptance checks ask respondents to report what they thought the plan would deliver. This question, however, may be difficult to answer as doubts about the plan may suggest it won't deliver what it promises, without specifying what the scope of the actual delivery might be. Hence, the easiest answer is to reiterate the promises detailed in the scenario, possibly resulting in an overestimate of scenario acceptance. To assess the impact of scope on the perceived likelihood and quality of delivery, simple ratings may provide a better way to gauge the plan's perceived effectiveness.

Conclusions

As the above examples illustrate, there is a need for basic research into the mental representations that respondents form of the to-be-valued good. This research needs to take respondents' real world knowledge into account, exploring how the application of this knowledge transforms the information provided in typical scenarios. The bottom line is simple: Respondents do not value the good as described, but the good as represented in their own mental construal of the scenario. Such construal processes have been explored in attitude measurement (e.g., Schwarz & Bless, 1992), person perception (e.g., Griffin & Ross, 1991), and behavioral decision research (e.g., Payne, Bettman, & Johnson, 1992), but they have found insufficient attention in recent CV discussions. Understanding these construal processes may raise as well as solve problems that are crucial to CV research. On the one hand, such research may uncover that the apparent insensitivity of CV responses to factors that should matter may reflect that these factors are not well represented in respondents' mental construal of the to-be-valued good. On the other hand, a mental construal approach to valuation emphasizes that the exact nature of the valued good cannot

be inferred from the scenario alone, rendering it difficult to determine what people value.

How Much Does it Take?

WTP and Implicit Cost Information

Given that respondents have no experience with the good, they need to determine their WTP under conditions of considerable uncertainty. According to CV theorizing, respondents are supposed to imagine the described good and to determine what this good would be worth to them. A key variable that people are likely to consider in determining what they are willing to pay for a good is the cost involved in producing the good. Suppose, for example, that your car breaks down and you are asked what you'd be willing to pay to get it restored to its previous condition. Chances are that you'd like to know if all it takes is to fix a loose wire, as opposed to, lets say, replacing the complete engine. And although the outcome is the same -- your car is running again -- your reaction to a $800 bill is likely to be very different in these two cases. As this example illustrates, our willingness to pay is not only determined by the outcome, but also by our assessment of the fair cost of producing that outcome.

Nevertheless, information about the amount of money it takes to produce a good is usually not explicitly provided in CV scenarios, for theoretical as well as practical reasons. In the case of natural resource damage assessments, the scenario is only supposed to serve as a vehicle that provides a plausible framework for eliciting the valuation of a good (e.g., a clean beach) that the plan is supposed to protect. In terms of the above car example, it is the running car that is to be valued, not the repair effort. Much as in the car example, however, respondents are likely to consider "how much it takes" to produce the outcome, in addition to the outcome itself.

As Schkade and Payne (1993, p. 283) observed on the basis of think-aloud protocols, "Perhaps the most common strategy in our sample involved first acknowledging that something should be done and then trying to figure out how much an appropriate amount would be." To determine the "appropriate amount", respondents are likely to consider cost information, resulting in increasing WTP with increasing cost estimates. To arrive at a cost estimate, respondents have two relevant sources of information, namely the complexity of the scenario and the $-value of the WTP question. Thus, scenarios that describe more extended plans, involving, for example, ten rather than one escort ships, are likely to convey higher cost and may hence be likely to produce higher WTP -- even under conditions where the same alleged outcome is described. A more directly relevant source of information, however, is provided by the $-value presented in a referendum question. In general, respondents assume that the numeric values provided as part of a question reflect the researcher's expert knowledge about the issue under study (see Schwarz & Hippler, 1991, for a review of relevant research). From this perspective, the $-value of the WTP question conveys the experts' cost estimate for the described plan.

To illustrate this point, I recently asked 20 students at the University of Michigan if they'd vote YES or NO on a plan to spend either $5 or $50 of their tuition to cover the start-up cost of an extensive recycling program on campus. Subsequently, the students were asked to estimate what the actual start-up cost would be. Not surprisingly, they estimated the start-up cost to be about eight times higher when the plan involved a contribution of $50 rather than $5.

That the $-value offered as part of a WTP question provides implicit cost information has methodological as well as theoretical implications. Whereas the methodological implications are likely to be appreciated by CV researchers, the theoretical implications are

more problematic for CV research. On the methodological side, the present perspective suggests that any discrepancies between open-ended WTP questions and referendum questions are a function of the specific $-value presented in the referendum question. If the suggested $-value is higher than the modal value offered in an open-ended format, referendum questions will result in higher WTP estimates; if it is lower, referendum questions will result in lower WTP estimates. Hence, general discussions of whether referendum questions over- or underestimate WTP (e.g., Desvouges, Hudson, & Ruby, 1994; Green, 1992) are likely to be futile because the specific outcome depends on the specific value presented as the experts' estimate of "what it takes."

On the theoretical side, this perspective implies that the $-value offered in a referendum question is best conceptualized as part of the scenario, informing respondents about the cost of the plan. Whereas this cost information is not explicitly provided in the scenario, it is offered by the elicitation question and most likely used by respondents in their mental construal. From this perspective, respondents who are provided with different $-values as part of the WTP question provide valuations of somewhat different goods rather than different valuations of the same good. As a methodological implication, it follows that discrepancies between WTP questions that use an open-ended versus closed response format (e.g., Holmes & Kramer, 1993; Cameron & Huppert, 1991) do not indicate a faulty measurement process that reflects a lack of convergent validity in assessing the value of the same good (in contrast to the conclusion reached by Desvouges et al., 1994). Rather, it indicates that valuing somewhat different goods results in different valuations -- further emphasizing the need to understand what it is that respondents actually value.

Note, however, that the perspective offered here has a troublesome implication: To the extent that respondents draw on cost information in determining their WTP, the cost

information provided by the researcher predetermines respondents' valuation of the good, essentially anticipating what the study is supposed to reveal. As suggested by Schkade and Payne's (1993) protocols, respondents may first determine if something needs to be done and may then try "to figure out how much an appropriate amount would be" (p. 283). If they interpret the referendum question as reflecting the experts' estimate of the contribution that it takes to meet the cost, this latter task is relatively easy, reducing the decision to whether one is willing, and able, to contribute one's share towards the provision of a collective good.

That respondents take such a cost sharing perspective is suggested by Schkade and Payne's (1993) protocols as well as by data reported by Green, Kahneman, and Kunreuther (1994). Schkade and Payne (1993, p. 283), for example, observed that "41% of the sample mentioned the idea that, if everyone did their part, each household would not have to give all that much. In fact, respondents who used this reasoning did give substantially lower WTPs." This theme is reiterated in Green et al.'s (1994) findings. Respondents in their study were asked to report their WTP for a program to teach English to immigrants to the United States. Some respondents were informed that 20 million other households would be asked the same question, whereas other respondents did not receive this information. Those whose attention was drawn to the large number of other households asked, were more likely to offer some payment, but the payment they did offer was lower than the payment offered by those who did not receive this information. Apparently, respondents who were aware that 20 million households would be asked to make a contribution inferred that each household's contribution could be relatively small. This finding suggests that respondents were willing to pay "what it takes", at least within limits, but certainly not more than it takes.

Although these findings are nicely compatible with a contribution model of CV measurement, as Green et al. (1994) suggest, they are not necessarily incompatible with a

purchase model. To return to the car example, suppose that your repair bill includes $100 to cover the mechanic's trip to your car. Would your response to this charge differ if the mechanic also repaired three additional cars that broke down within a few yards, charging each one $100 for the trip? It seems that considerations of cost sharing may be as important in a purchase frame as in a contribution frame.

Conclusions

In summary, the conjectures offered here bear on the concern that "CV responses are sensitive to methodological factors that, from the standpoint of economic theory, are irrelevant" (Schkade & Payne, 1993, p. 273), such as the specific format of the elicitation question used. From a psychological point of view, the key assumption of this argument is mistaken: Far from reflecting formal, methodological features without substantive relevance, different forms of the elicitation question provide different information that respondents actively use in solving a complex task. Much as many formal aspects of attitude questions, such as the numeric values of a rating scale (Schwarz, Knäuper, et al., 1991) have been found to change the substantive meaning of the question, so do different elicitation formats provide more than exchangeable "measurement devices." Whereas future systematic evidence bearing on these conjectures may put the concern to rest that CV responses are sensitive to "irrelevant" factors, these conjectures raise a potentially more troublesome concern by suggesting that the referendum questions recommended by the NOAA panel predetermine respondents' valuations by providing implicit cost information.

Mode of Data Collection

Survey responses are to some extent affected by the mode of data collection used. This

reflects, in part, that face-to-face interviews, telephone interviews and self-administered questionnaires pose somewhat different cognitive tasks (see Schwarz, Strack, Hippler, & Bishop, 1991, for a detailed discussion) and that these modes of data collection provide differential opportunities for respondent self-selection. Below I address two key differences between (face-to-face or telephone) interviews and mail questionnaires that seem particularly relevant to CV surveys, namely respondent self-selection and the degree to which the researcher can control respondents' exposure to relevant information. In combination, these two factors argue against the use of mail surveys in CV research.

Exposure to Information

One of the key differences between self-administered questionnaires and face-to-face or telephone interviews is that self-administered questionnaires provide no control over the extent to which respondents understand the material presented to them and the order in which questions are read (see Schwarz, Strack, et al., 1991). As several studies indicated, an unknown proportion of respondents is likely to read parts or all of the questionnaire before filling in an answer, resulting in influences of subsequent questions on answers given to preceding questions. In addition, relevant issues may be discussed within the household, introducing other influences that are difficult to identify. For CV surveys, this lack of control has several potentially problematic implications.

First, it remains unclear to what extent respondents have carefully read the often lengthy and cumbersome scenarios. Hence, an unknown proportion of respondents may answer the questions on the basis of a rather superficial understanding of the scenario. The previously discussed observation that Schkade and Payne (1993) obtained differences in WTP as a function of scope when they requested respondents to think aloud while answering the

questionnaire, whereas Desvouges et al. (1993) did not when they administered the questionnaire without a think-aloud requirement illustrates this point. On the other hand, respondents who are sufficiently interested in the topic have more time to think about the scenario under self-administered conditions than they would under interview conditions. At the extreme, they may return to the questionnaire at a later time, discuss it with others, gather additional information, and so on. Such an increased elaboration of the scenario is only likely for highly motivated respondents and its specific impact on WTP is likely to depend on how plausible the scenario seems, once one scrutinizes its details.

Second, to the extent that respondents read ahead, their interpretation of the scenario may be affected by their knowledge of the questions that follow it, including the $-value mentioned in the WTP question or the manipulation checks that ask them if they believed the information that was presented in the scenario. At present, it is unclear what the exact implications of knowing these questions are, but it seems likely that knowing the $-value affects the encoding of the scenario and that reading the manipulation checks may raise doubts about the reliability of the information offered.

Third, the possibility to discuss the issue with other household members or to gather more information may introduce information that is not contained in the scenario, further affecting respondents' construal of the to-be-valued good.

As these possibilities illustrate, mail survey conditions render it even more difficult to determine what respondents are actually valuing than interview surveys, reflecting that the researcher has less control over respondents' exposure to relevant information. Note, however, that some of these potential problems are of less concern when self-administered questionnaires are used under controlled laboratory conditions, where respondents may be discouraged from reading ahead and have to complete the questionnaire without interruption.

Accordingly, self-administered questionnaires are a feasible instrument for basic research under controlled administration conditions, although their use under the uncontrollable conditions of mail surveys cannot be recommended.

Respondent Self-Selection

Adding to the above problems, mail surveys are particularly prone to respondent self-selection, as Schuman (1994) noted. This follows from the simple fact that mail survey respondents can screen the questionnaire before they decide to participate, whereas face-to-face and telephone surveys do not provide this opportunity. As a result, nonresponse in interview surveys reflects the impact of variables that are unrelated to the specific content of the survey, whereas nonresponse in mail surveys is, to an unknown degree, affected by content. Hence, mail surveys are likely to overrepresent individuals who find the topic interesting and important.

Although this issue becomes less problematic as the response rate of mail surveys increases, self-selection problems remain even at high response rates. As an illustration, suppose that a mail and a telephone survey both obtain an 80% response rate. In the telephone survey, the 20% nonresponse includes participants who refuse because they are called at a bad time, are on vacation, or whatever. However, it does not include respondents who thought about the issue and decided it is not worth their time. In contrast, mail respondents can work on the questionnaire at a time of their choice, thus potentially reducing nonresponse due to timing problems. On the other hand, however, they have the possibility to screen the questionnaire and are more likely to participate in the survey if they find the issue of interest. As a result, an identical nonresponse rate of 20% under both modes is likely to be unrelated to the topic under interview conditions, but not under mail conditions.

Hence, similar response rates under different modes do not necessarily indicate comparable samples.

Moreover, the variable that drives self-selection under mail conditions is respondents' interest in the topic, which may be only weakly related to sociodemographic variables. Accordingly, topic driven self-selection may be present even if the completed sample seems representative with regard to sociodemographic variables like age, sex, income, and so on. To assess problems of topic related self-selection, one needs to assess respondents' interest in the topic. The current considerations predict that mail surveys include a higher percentage of respondents who report high interest than face-to-face or telephone surveys, even if the samples are comparable on sociodemographic grounds. If this differential self-selection results in an over- or underestimation of WTP, however, is difficult to determine because interest in the topic may reflect favorable as well as unfavorable attitudes towards the issue.

In combination with the lack of control over respondents' understanding of the provided information, the problem of topic related self-selection renders mail surveys unfeasible for CV research.

References

Cameron, T., & Huppert, D. (1991). Referendum contingent valuation estimates: Sensitivity to the assignment of offered values. Journal of the American Statistical Association, 86, 910-918.

Carson, R. T., & Mitchell, R.C. (1993). Sequencing and nesting in contingent valuation surveys. Unpublished manuscript.

Carson, R. T., & Mitchell, R. C. (in press). The issue of scope in contingent valuation studies. American Journal of Agricultural Economics.

Carson, R. T., Mitchell, R. C., Hanemann, W. M., Kopp, R. J., Presser, S., & Ruud, P. A. (1992). A contingent valuation study of lost passive use values resulting from the Exxon Valdez oil spill. San Diego: NRDA.

Clark, H.H. (1985). Language use and language users. In G. Lindzey, & E. Aronson (Eds.), Handbook of social psychology (Vol. 2, pp. 179 - 232). New York: Random House.

Clark, H. H., & Clark, E. V. (1977). Psychology and language. New York: Harcourt, Brace, Jovanovich.

Clark, H. H., & Schober, M. F. (1992). Asking questions and influencing answers. In J. M. Tanur (Ed.), Questions about questions (pp. 15-48). New York: Russel Sage.

DeMaio, T. J. (1984). Social desirability and survey measurement: A review. In C. F. Turner & E. Martin (Eds.), Surveying subjective phenomena (Vol. 2, pp. 257-281). New York: Russell Sage.

Desvouges, W. H., Hudson, S. P., & Ruby, M. C. (1994, May). Evaluating CV performance: Separating the light from the heat. EPA/DOE Workshop, Herndon, VA.

Desvouges, W. H., Johnson, R., Dunford, R. W., Boyle, K. J., Hudson, S. P., & Wilson, K. N. (1993). Measuring natural resource damages with contingent valuation. In J.

A. Hausman (Ed.), Contingent valuation: A critical assessment. Amsterdam: Elsevier.

Diamond, P. A., Hausman, J. A., Leonard, G.K., & Denning, M. A. (1993). Does contingent valuation measure preferences? Experimental evidence. In J. A. Hausman (Ed.), Contingent valuation: A critical assessment. Amsterdam: Elsevier.

Fischhoff, B. (1991). Value elicitation: Is there anything there? American Psychologist, 46, 835-847.

Fischhoff, B. & Furby, L. (1988). Measuring values: A conceptual framework for interpreting transactions with special reference to contingent valuation of visibility. Journal of Risk and Uncertainty, 1, 147-184.

Green, D. P. (1992). The price elasticity of mass preferences. American Political Science Review, 86, 128-148.

Green, D. P., Kahneman, D., & Kunreuther, H. (1994). How the scope and method of public funding affect willingness to pay for public goods. Public Opinion Quarterly, 58, 49-67.

Grice, H. P. (1975). Logic and conversation. In P. Cole, & J.L. Morgan (Eds.), Syntax and semantics, 3: Speech acts. pp. 41 -58. New York: Academic Press.

Griffin, D. W., & Ross, L. (1991). Subjective construal, social inference, and human misunderstanding. In M. Zanna (Ed.), Advances in experimental social psychology (vol. 24, pp. 319-360).

Hausman, J.A. (Ed.) (1993). Contingent valuation. A critical assessment. Amsterdam: Elsevier.

Hippler, H.J., Schwarz, N., & Sudman, S. (Eds.). (1987). Social information processing and survey methodology. New York, NY: Springer Verlag.

Holmes, T. P., & Kramer, R. A. (1993). An independent sample test of yea-saying in

dichotomous choice contingent valuation. Unpublished manuscript. Research Triangle Park: Forestry Sciences Lab.

Kahneman, D., & Knetsch, J. A. (1992). Valuing public goods: The purchase of moral satisfaction. Journal of Environmental Economics and Management, 22, 57-70.

Kahneman, D., & Ritov, I. (1994). Determinants and stated willingness to pay for public goods: A study in the headline method. Journal of Risk and Uncertainty, in press.

Kopp, R. J., & Pease, K. A. (1994). Contingent valuation: Economics, law and politics. Manuscript. Washington, DC: Resources for the Future.

Mitchell, R. C., & Carson, R. T. (1986). Valuing drinking water risk reduction using the contingent valuation method. Report to the U. S. Environmental Protection Agency, Washington, D.C.

Mitchell, R. C., & Carson, R. T. (1989). Using surveys to value public goods: The contingent valuation method. Washington, D.C.: Resources for the Future.

Payne, J. W., Bettman, J. R., & Johnson, E.J. (1992). Behavioral decision research: A constructive processing perspective. Annual Review of Psychology, 43, 87-131.

Schkade, D. A., & Payne, J. W. (1993). Where do the numbers come from? How people respond to contingent valuation questions. In J. A. Hausman (Ed.), Contingent valuation. A critical assessment (pp. 271-303). Amsterdam: Elsevier.

Schulze, W., McClelland, G., & Lazo, J. (1994, May). Methodological issues in using contingent valuation to measure non-use values. EPA/DOE Workshop, Herndon, VA.

Schwarz, N. (1994). Judgment in a social context: Biases, shortcomings, and the logic of conversation. In M. Zanna (Ed.), Advances in experimental social psychology (Vol. 26). San Diego, CA: Academic Press.

Schwarz, N., & Bless, H. (1992). Constructing reality and its alternatives: Assimilation and

contrast effects in social judgment. In L.L. Martin & A. Tesser (Eds.), The construction of social judgment (pp. 217-245). Hillsdale, NJ: Erlbaum.

Schwarz, N., & Hippler, H.J. (1987). What response scales may tell your respondents: Informative functions of response alternatives. In H.J. Hippler, N. Schwarz, & S. Sudman (Eds.), Social information processing and survey methodology (pp. 163-178). New York, NY: Springer Verlag.

Schwarz, N., & Hippler, H.J. (1991). Response alternatives: The impact of their choice and ordering. In P. Biemer, R. Groves, N. Mathiowetz, & S. Sudman (Eds.), Measurement error in surveys (pp. 41-56). Chichester, UK: Wiley.

Schwarz, N., Knäuper, B., Hippler, H. J., Noelle-Neumann, E., & Clark, F. (1991). Rating scales: Numeric values may change the meaning of scale labels. Public Opinion Quarterly, 55, 570-582.

Schwarz, N., & Strack, F. (1991). Context effects in attitude surveys: Applying cognitive theory to social research. In W. Stroebe & M. Hewstone (Eds.), European Review of Social Psychology (Vol. 2, pp. 31-50). Chichester: Wiley.

Schwarz, N., Strack, F., Hippler, H.J., & Bishop, G. (1991). The impact of administration mode on response effects in survey measurement. Applied Cognitive Psychology, 5, 193-212.

Schwarz, N. & Sudman, S. (Eds.) (1992). Context effects in social and psychological research. New York, NY: Springer Verlag.

Schwarz, N., & Sudman, S. (Eds.) (1994). Autobiographical memory and the validity of retrospective reports. New York, NY: Springer Verlag.

Sperber, D., & Wilson, D. (1986). Relevance: Communication and cognition. Cambridge, MA: Harvard University Press.

Strack, F., & Martin, L. (1987). Thinking, judging, and communicating: A process account of context effects in attitude surveys. In H.J. Hippler, N. Schwarz, & S. Sudman (Eds.), Social information processing and survey methodology (pp. 123 - 148). New York: Springer Verlag.

Strack, F., & Schwarz, N. (1992). Communicative influences in standardized question situations: The case of implicit collaboration. In K. Fiedler & G. Semin (Eds.), Language, interaction and social cognition (pp. 173-193). Beverly Hills: Sage.

Strack, F., Schwarz, N., & Wänke, M. (1991). Semantic and pragmatic aspects of context effects in social and psychological research. Social Cognition, 9, 111-125.

Sudman, S., & Bradburn, N. M. (1983). Asking questions. San Francisco: Jossey-Bass.

Tanur, J. M. (Ed.) (1992). Questions about questions. New York: Russel Sage.

Thaler, R. H. (1992). The winner's curse: Paradoxes and anomalies in economic life. New York: Free Press.

Tourangeau, R. (1992). Attitudes as memory structures: Belief sampling and context effects. In N. Schwarz & S. Sudman (Eds.), Context effects in social and psychological research. New York: Springer Verlag.

Tourangeau, R., & Rasinski, K.A. (1988). Cognitive processes underlying context effects in attitude measurement. Psychological Bulletin, 103, 299 - 314.

ZUMA-Arbeitsberichte

80/15 Gerhard Arminger, Willibald Nagl, Karl F. Schuessler
Methoden der Analyse zeitbezogener Daten. Vortragsskripten der ZUMA Arbeitstagung vom 25.9. bis 5. 10. 79

81/07 Erika Brückner, Hans-Peter Kirschner, Rolf Porst, Peter Prüfer, Peter Schmidt
Methodenbericht zum "AllBUS 1980"

81/19 Manfred Küchler, Thomas P. Wilson, Don H. Zimmerman
Integration von qualitativen und quantitativen Forschungsansätzen

82/03 Gerhard Arminger, Horst Busse, Manfred Küchler
Verallgemeinerte Lineare Modelle in der empirischen Sozialforschung

82/08 Glenn R. Carroll
Dynamic analysis of discrete dependent variables: A didactic essay

82/09 Manfred Küchler
Zur Messung der Stabilität von Wählerpotentialen

82/10 Manfred Küchler
Zur Konstanz der Recallfrage

82/12 Rolf Porst
"ALLBUS 1982" - Systematische Variablenübersicht und erste Ansätze zu einer Kritik des Fragenprogramms

82/13 Peter Ph. Mohler
SAR - Simple AND Retrieval mit dem Siemens-EDT-Textmanipulationsprogramm

82/14 Cornelia Krauth
Vergleichsstudien zum "ALLBUS 1980"

82/21 Werner Hagstotz, Hans-Peter Kirschner, Rolf Porst, Peter Prüfer
Methodenbericht zum "ALLBUS 1982"

83/09 Bernd Wegener
Two approaches to the analysis of judgments of prestige: Interindividual differences and the general scale

83/11 Rolf Porst
Synopse der ALLBUS-Variablen. Die Systematik des ALLBUS-Fragenprogramms und ihre inhaltliche Ausgestaltung im ALLBUS 1980 und ALLBUS 1982

84/01 Manfred Küchler, Peter Ph. Mohler
Qualshop (ZUMA-Arbeitstagung zum "Datenmanagement bei qualitativen Erhebungsverfahren") - Sammlung von Arbeitspapieren und -berichten, Teil I + II

84/02 Bernd Wegener
Gibt es Sozialprestige? Konstruktion und Validität der Magnitude-Prestige-Skala

84/03 Peter Prüfer, Margrit Rexroth
Erfahrungen mit einer Technik zur Bewertung von Interviewerverhalten

84/04 Frank Faulbaum
Ergebnisse der Methodenstudie zur internationalen Vergleichbarkeit von Einstellungsskalen in der Allgemeinen Bevölkerungsumfrage der Sozialwissenschaften (ALLBUS) 1982

84/05 Jürgen Hoffmeyer-Zlotnik
Wohnquartiersbeschreibung. Ein Instrument zur Bestimmung des sozialen Status von Zielhaushalten

84/07 Gabriele Hippler, Hans-Jürgen Hippler
Reducing Refusal Rates in the Case of Threatening Questions: The "Door-in-the-Face" Technique

85/01 Hartmut Esser
Befragtenverhalten als "rationales Handeln" - Zur Erklärung von Antwortverzerrungen in Interviews

85/03 Rolf Porst, Peter Prüfer, Michael Wiedenbeck, Klaus Zeifang
Methodenbericht zum "ALLBUS 1984"

86/01 Dagmar Krebs
Zur Konstruktion von Einstellungsskalen im interkulturellen Vergleich

86/02 Hartmut Esser
Können Befragte lügen? Zum Konzept des "wahren Wertes" im Rahmen der handlungstheoretischen Erklärung von Situationseinflüssen bei der Befragung

86/03 Bernd Wegener
Prestige and Status as Function of Unit Size

86/04 Frank Faulbaum
Very Soft Modeling: The Logical Specification and Analysis of Comlex Process Explanations with Arbitrary Degrees of Underidentification and Variables of Arbitrary Aggregation and Measurement Levels

86/05 Peter Prüfer, Margrit Rexroth (Übersetzung Dorothy Duncan)
On the Use of the Interaction Coding Technique

86/06 Hans-Peter Kirschner
Zur Kessler-Greenberg-Zerlegung der Varianz der Meßdifferenz zwischen zwei Meßzeitpunkten einer Panel-Befragung

86/07 Georg Erdmann
Ansätze zur Abbildung sozialer Systeme mittels nicht-linearer dynamischer Modelle

86/09 Heiner Ritter
Einige Ergebnisse von Vergleichstests zwischen den PC- und Mainframe-Versionen von SPSS und SAS

86/11 Günter Rothe
Bootstrap in generalisierten linearen Modellen

87/01 Klaus Zeifang
Die Test-Retest-Studie zum ALLBUS 1984 - Tabellenband

87/02 Klaus Zeifang
Die Test-Retest-Studie zum ALLBUS 1984 - Abschlußbericht

87/04 Barbara Erbslöh, Michael Wiedenbeck
Methodenbericht zum "ALLBUS 1986"

87/05 Norbert Schwarz, Julia Bienias
What Mediates the Impact of Response Alternatives on Behavioral Reports?

87/06 Norbert Schwarz, Fritz Strack, Gesine Müller, Brigitte Chassein
The Range of Response Alternatives may determine the Meaning of the Question: Further Evidence on Informative Functions of Response Alternatives

87/07 Fritz Strack, Leonard L. Martin, Norbert Schwarz
The Context Paradox in Attitude Surveys: Assimilation or Contrast?

87/08 Gudmund R. Iversen
Introduction to Contextual Analysis

87/09 Seymour Sudman, Norbert Schwarz
Contributions of Cognitive Psychology to Data Collection in Marketing Research

87/10 Norbert Schwarz, Fritz Strack, Denis Hilton, Gabi Naderer
Base-Rates, Representativeness, and the Logic of Conversation

87/11 George F. Bishop, Hans-Jürgen Hippler, Norbert Schwarz, Fritz Strack
A Comparison of Response Effects in Self-Administered and Telephone Surveys

87/12 Norbert Schwarz
Stimmung als Information. Zum Einfluß von Stimmungen und Emotionen auf evaluative Urteile

88/01 Antje Nebel, Fritz Strack, Norbert Schwarz
Tests als Treatment: Wie die psychologische Messung ihren Gegenstand verändert

88/02 Gerd Bohner, Herbert Bless, Norbert Schwarz, Fritz Strack
What Triggers Causal Attributions? The Impact of Valence and Subjective Probability

88/03 Norbert Schwarz, Fritz Strack
The Survey Interview and the Logic of Conversation: Implications for Questionnaire Construction

88/04 Hans-Jürgen Hippler, Norbert Schwarz
"No Opinion"-Filters: A Cognitive Perspective

88/05 Norbert Schwarz, Fritz Strack
Evaluating One's Life: A Judgment of Subjective Well-Being

88/06 Norbert Schwarz, Herbert Bless, Gerd Bohner, Uwe Harlacher, Margit Kellenbenz
Response Scales as Frames of Reference: The Impact of Frequency Range on Diagnostic Judgments

88/07 Michael Braun
Allbus-Bibliographie (7. Fassung, Stand: 30.6.88)

88/08 Günter Rothe
Ein Ansatz zur Konstruktion inferenzstatistisch verwertbarer Indices

88/09 Ute Hauck, Reiner Trometer
Methodenbericht International Social Survey Program - ISSP 1987

88/10 Norbert Schwarz
Assessing frequency reports of mundane behaviors: Contributions of cognitive psychology to questionnaire construction

88/11 Norbert Schwarz, B. Scheuring (sub.)
Judgments of relationship satisfaction: Inter- and intraindividual comparison strategies as a function of questionnaire structure

88/12 Rolf Porst, Michael Schneid
Ausfälle und Verweigerungen bei Panelbefragungen - Ein Beispiel -

88/13 Cornelia Züll
SPSS-X. Anmerkungen zur Siemens BS2000 Version

88/14 Michael Schneid
Datenerhebung am PC - Vergleich der Interviewprogramme "internerv+" und "THIS"

88/15 Norbert Schwarz, Bettina Scheuring
Die Vergleichsrichtung bestimmt das Ergebnis von Vergleichsprozessen: Ist - Idealdiskrepanzen in der Partnerwahrnehmung

88/16 Norbert Schwarz, Bettina Scheuring
Die Vergleichsrichtung bestimmt das Ergebnis von Vergleichsprozessen: Ist-Idealdiskrepanzen in der Beziehungsbeurteilung

89/01 Norbert Schwarz, George F. Bishop, Hans-J. Hippler, Fritz Strack
Psychological Sources of Response Effects in Self-Administered And Telephone Surveys

89/02 Michael Braun, Reiner Trometer, Michael Wiedenbeck
Methodenbericht. Allgemeine Bevölkerungsumfrage der Sozialwissenschaften - ALLBUS 1988

89/03 Norbert Schwarz
Feelings as Information: Informational and Motivational Functions of Affective States

89/04 Günter Rothe
Jackknife and Bootstrap: Resampling-Verfahren zur Genauigkeitsschätzung von Parameterschätzungen

89/05 Herbert Bless, Gerd Bohner, Norbert Schwarz und Fritz Strack
Happy and Mindless? Moods and the Processing of Persuasive Communications

89/06 Gerd Bohner, Norbert Schwarz und Stefan E. Hormuth
Die Stimmungs-Skala: Eine deutsche Version des "Mood Survey" von Underwood und Froming

89/07 Ulrich Mueller
Evolutionary Fundamentals of Social Inequality, Dominance and Cooperation

89/08 Robert Huckfeldt
Noncompliance and the Limits of Coercion: The Problematic Enforcement of Unpopular Laws

89/09 Peter Ph. Mohler, Katja Frehsen und Ute Hauck
CUI - Computerunterstützte Inhaltsanalyse. Grundzüge und Auswahlbibliographie zu neueren Anwendungen

89/10 Cornelia Züll, Peter Ph. Mohler
Der General Inquirer III - Ein Dinosaurier für die historische Forschung

89/11 Fritz Strack, Norbert Schwarz, Brigitte Chassein, Dieter Kern, Dirk Wagner
The Salience of Comparison Standards and the Activation of Social Norms: Consequences for Judgments of Happiness and their Communication

89/12 Jutta Kreiselmaier, Rolf Porst
Methodische Probleme bei der Durchführung telefonischer Befragungen: Stichprobenziehung und Ermittlung von Zielpersonen, Ausschöpfung und Nonresponse, Qualität der Daten

89/13 Rainer Mathes
Modulsystem und Netzwerktechnik. Neuere inhaltsanalytische Verfahren zur Analyse von Kommunikationsinhalten

89/14 Jutta Kreiselmeier, Peter Prüfer, Margrit Rexroth
Der Interviewer im Pretest. Evaluation der Interviewerleistung und Entwurf eines neuen Pretestkonzepts. April 1989

89/15 Henrik Tham
Crime as a Social Indicator

89/16 Ulrich Mueller
Expanding the Theoretical and Methodological Framework of Social Dilemma Research

89/17 Hans-J. Hippler, Norbert Schwarz, Elisabeth Noelle-Neumann
Response Order Effects in Dichotomous Questions: The Impact of Administration Mode

89/18 Norbert Schwarz, Hans-J. Hippler, Elisabeth Noelle-Neumann, Thomas Münkel
Response Order Effects in Long Lists: Primacy, Recency, and Asymmetric Contrast Effects

89/19 Wolfgang Meyer
Umweltberichterstattung in der Bundesrepublik Deutschland

89/20 Michael Braun, Reiner Trometer
ALLBUS Bibliographie (8. Fassung, Stand: 30.6.1989)

89/21 Günter Rothe
Gewichtungen zur Anpassung an Statusvariablen. Eine Untersuchung am ALLBUS 1986

89/22 Norbert Schwarz, Thomas Münkel, Hans-J. Hippler
What determines a "Perspective"? Contrast Effects as a Function of the Dimension Tapped by Preceding Questions

89/23 Norbert Schwarz, Andreas Bayer
Variationen der Fragenreihenfolge als Instrument der Kausalitätsprüfung: Eine Untersuchung zur Neutralisationstheorie devianten Verhaltens

90/01 Norbert Schwarz, Andreas Bayer
Assimilation and Contrast Effects in Part-Whole Question Sequences: A Conversational Logic Analysis

90/02 Norbert Schwarz, Fritz Strack, Hans-J. Hippler, George Bishop
The Impact of Administration Mode on Response Effects in Survey Measurement

90/03 Norbert Schwarz, Herbert Bless, Gerd Bohner
Mood and Persuasion: Affective States Influence the Processing of Persuasive Communications

90/04 Michael Braun, Reiner Trometer
ALLBUS-Bibliographie 90

90/05 Norbert Schwarz, Fritz Strack
Context Effects in Attitude Surveys: Applying Cognitive Theory to Social Research

90/06 Norbert Schwarz, Herbert Bless, Fritz Strack, Gisela Klumpp, Annette Simons
Ease of Retrieval as Information: Another Look at the Availability Heuristic

90/07 Norbert Schwarz, Fritz Strack, Hans-J. Hippler
Kognitionspsychologie und Umfrageforschung: Themen und Befunde eines interdisziplinären Forschungsgebietes

90/08 Norbert Schwarz, Hans-J. Hippler
Response Alternatives: The Impact of their Choice and Presentation Order

90/09 Achim Koch
Externe Vergleichsdaten zum ALLBUS 1984, 1986, 1988

90/10 Norbert Schwarz, Bärbel Knäuper, Hans-J. Hippler, Elisabeth Noelle-Neumann, Leslie Clark
Rating Scales: Numeric Values May Change the Meaning of Scale Labels

91/01 Denis J. Hilton
Conversational Inference and Rational Judgment

91/02 Denis J. Hilton
A Conversational Model of Causal Explanation

91/03 Mood Effects on Interpersonal Preferences: Evidence for Motivated Processing Strategies

91/04 Joseph P. Forgas
Affective Influences on Interpersonal Perception

91/05 Norbert Schwarz, Herbert Bless
Constructing Reality and Its Alternatives: An Inclusion/Exclusion Model of Assimilation and Contrast Effects in Social Judgment

91/06 Herbert Bless, Roland F. Fellhauer, Gerd Bohner, Norbert Schwarz
Need for Cognition: Eine Skala zur Erfassung von Engagement und Freude bei Denkaufgaben

91/07 Norbert Schwarz, Bärbel Knäuper, E. Tory Higgins
Der Einfluß von Rangordnungsaufgaben auf nachfolgende Denkprozesse: Zur Aktivierung prozeduraler Sets

91/08 Bettina Scheuring, Norbert Schwarz
Selbstberichtete Verhaltens- und Symptomhäufigkeiten: Was Befragte aus Antwortvorgaben des Fragebogens lernen

91/09 Norbert Schwarz, Herbert Bless
Scandals and the Public's Trust in Politicians: Assimilation and Contrast Effects

91/10 Rolf Porst
Ausfälle und Verweigerungen bei einer telefonischen Befragung

91/11 Uwe Blien, Heike Wirth, Michael Müller
Identification risk for microdata stemming from official statistics

91/12 Petra Beckmann
Methodological Report ISSP 1989

91/13 Martina Wasmer, Achim Koch, Michael Wiedenbeck
Methodenbericht zur "Allgemeinen Bevölkerungsumfrage der Sozialwissenschaften" (Allbus) 1990

91/14 Uwe Blien, Oded Löwenbein
Einkommensanalysen auf der Grundlage amtlicher Daten und Umfragedaten: Ergebnisse zur betrieblichen Seniorität und Arbeitslosigkeit

91/15 Petra Beckmann, Peter Mohler, Rolf Uher
ISSP Basic Information on the ISSP Data Collection 1985 -1994

91/16 Norbert Schwarz
In welcher Reihenfolge fragen? Kontexteffekte in standardisierten Befragungen

91/17 Ellen D. Riggle, Victor C. Ottati, Robert S. Wyer, Jr., James Kuklinski, Norbert Schwarz
Bases of Political Judgments: The Role of Stereotypic and Non-stereotypic Information

91/18 Dagmar Krebs
Was ist sozial erwünscht? Der Grad sozialer Erwünschtheit von Einstellungsitems

91/19 Michael Braun, Reiner Trometer
ALLBUS-Bibliographie

91/20 Michael Schneid
Einsatz computergestützter Befragungssysteme in der Bundesrepublik Deutschland

91/21 Rolf Porst, Michael Schneid
Software-Anforderungen an computergestützte Befragungssysteme

91/22 Ulrich Mueller
The Reproductive Success of the Elites in Germany, Great Britain, Japan and the USA during the 19th and 20th Century

92/01 P.H. Hartmann, B. Schimpl-Neimanns
Zur Repräsentativität sozio-demographischer Merkmale des ALLBUS - multivariate Analysen zum Mittelschichtbias der Umfrageforschung

92/02 Gerd Bohner, Kimberly Crow, Hans-Peter Erb, Norbert Schwarz
Affect and Persuasion: Mood Effects on the Processing of Message Content and Context Cues and on Subsequent Behavior

92/03 Herbert Bless, Gerd Bohner, Traudel Hild, Norbert Schwarz
Asking Difficult Questions: Task Complexity Increases the Impact of Response Alternatives

92/04 Wolfgang Bandilla, Siegfried Gabler, Michael Wiedenbeck
Methodenbericht zum DFG-Projekt Allbus Baseline-Studie

92/05 Frank Faulbaum
Von der Variablenanalyse zur Evaluation von Handlungs- und Prozeßzusammenhängen

92/06 Ingwer Borg
Überlegungen und Untersuchungen zur Messung der subjektiven Unsicherheit der Arbeitsstelle

92/07 Ingwer Borg, Michael Braun
Arbeitsethik und Arbeitsinvolvement als Moderatoren der psychologischen Auswirkungen von Arbeitsunsicherheit

92/08 Eleanor Singer, Hans-Jürgen Hippler, Norbert Schwarz
Confidentiality Assurances in Surveys: Reassurance or Threat?

92/09 Herbert Bless, Diane M. Mackie, Norbert Schwarz
Mood Effects on Attitude Judgments: The Independent Effects of Mood before and after Message Elaboration

92/10 Ulrich Mueller, Carola Schmid
Ehehäufigkeit und Fruchtbarkeit weiblicher Mitglieder der deutschen Elite

92/11 Herbert Bless, Fritz Strack, Norbert Schwarz
The Informative Functions of Research Procedures: Bias and the Logic of Conversation

92/12 Norbert Schwarz, Herbert Bless, Michaela Wänke
Subjective Assessment and Evaluation of Change: Lessons from Social Cognition Research

92/13 Norbert Schwarz, Hans-J. Hippler
Buffer Items: When Do They Buffer and When Don't They?

92/14 Hans-J. Hippler, Norbert Schwarz
The Impact of Administration Modes on Response Effects in Surveys

92/15 Michaela Wänke, Norbert Schwarz
Comparative Judgments: How the Direction of Comparison Determines the Answer

92/16 Michael Braun, Reiner Trometer
ALLBUS-Bibliographie (11. Fassung, Stand: 30.6.1992)

92/17 Anke Nau, Rolf Porst
Projektbericht ZUMA-Methodenpanel Teil 1: Konzeption, Vorbereitung und Durchführung

92/18 Michael Schneid
Handbuch ZUMA-Feldsteuerungsprogramm

92/19 Paul Lüttinger, Bernhard Schimpl-Neimanns
Amtliche Bildungsstatistik und empirische Sozialforschung

92/20 Rolf Porst, Michael Schneid
Fragebogenschreiben mit Microsoft WORD 5.0: Druckformate, Textbausteine und Makros zur effizienten Gestaltung von Fragebogen

92/21 Michael Braun, Jacqueline Scott, Duane F. Alwin
Economic Necessity or Self-Actualization? Attitudes toward Women's Labor-force Participation in the East and West

92/22 Duane F. Alwin, Michael Braun, Jacqueline Scott
The Separation of Work and the Family: Attitudes Towards Women's Labour-Force Participation in Germany, Great Britain, and the United States

92/23 Michaela Wänke, Norbert Schwarz, Herbert Bless
The Availability Heuristic Revisited: Experienced Ease of Retrieval in Mundane Frequency Estimates

93/01 Michael Braun, Carmen Eilinghoff, Siegfried Gabler, Michael Wiedenbeck
Methodenbericht zur Allgemeinen Bevölkerungsumfrage der Sozialwissenschaften (Allbus) 1992

93/02 Michael Häder, Sabine Novossadeck
Einstellungsdynamik in Ostdeutschland. Auswertung von Wiederholungsuntersuchungen zur individuellen Bewältigung des sozialen Wandels

93/03 Norbert Schwarz
Judgment in a Social Context: Biases, Shortcomings, and the Logic of Conversation

93/04 Norbert Schwarz, Michaela Wänke, Herbert Bless
Subjective Assessments and Evaluations of Change: Some Lessons from Social Cognition Research

93/05 Norbert Schwarz
Context Effects in Attitude Measurement

93/06 Reiner Trometer
Die Operationalisierung des Klassenschemas nach Goldthorpe im ALLBUS

93/07 Carola Schmid
Der Zugang zu den Daten der Demographie

93/08 Michael Schneid
Fragebogenschreiben mit Word für Windows

93/09 Achim Koch
Die Nutzung demographischer Informationen in den Veröffentlichungen mit Allbus-Daten

93/10 Helmut Schröder
Über den Zusammenhang zwischen Aktivitäten und Zufriedenheit: "Eine kommunale Seniorenbefragung"

93/11 Michael Braun, Reiner Trometer
ALLBUS-Bibliographie, 12. Fassung, Stand: 30.9.93

93/12

93/13 Steven E. Finkel, Peter R. Schrott
Campaign Effects on Voter Choice in the German Election of 1990

93/14 Jürgen Hoffmeyer-Zlotnik, Dagmar Krebs
Subjektive Statuszuweisung; Objektive Schichtmessung

93/15 Dagmar Krebs
Richtungseffekte von Itemformulierungen

93/16 Dagmar Krebs
Social Desirability: The collective conscience? Judging the degree of social desirability in attitude items

93/17 Bernhard Krüger, Heiner Ritter, Cornelia Züll
SPSS Einsatz auf unterschiedlichen Plattformen in einem Netzwerk: Daten und Ergebnisaustausch

94/01 Jürgen H.P. Hoffmeyer-Zlotnik, Michael Wiedenbeck
Überlegungen zu Sampling, Qualitätsprüfung und Auswertung von Daten aus Teilpopulationen

94/02 Michael Häder, Sabine Häder
Die Grundlagen der Delphi-Methode: - Ein Literaturbericht -

94/03 Sabine Häder
Auswahlverfahren bei Telefonumfragen

94/04 Peter Prüfer, Margrit Rexroth
Ein Verfahren zur Erfassung von Erhebungsproblemen bei Interviews der Hauptstudie

94/05 Michael Häder, Sabine Häder
Ergebnisse einer Experimentellen-Studie zur Delphi-Methode

94/06 Bernhard Schimpl-Neimanns, Heike Wirth
Bestandsaufnahme und Nutzungsmöglichkeiten amtlicher Mikrodaten der DDR für Sekundäranalysen zur Bildungs- und Einkommensungleichheit

94/07 Norbert Schwarz, Hans-J. Hippler
Subsequent Questions May Influence Answers to Preceding Questions in Mail Surveys

94/08 Norbert Schwarz, Hans-J. Hippler
The Numeric Values of Rating Scales: A Comparison of their Impact in Mail Surveys and Telephone Interviews

94/09 Norbert Schwarz
Cognition, Communication, and Survey Measurement:
Some Implications for Contingent Valuation Surveys