Rating scales: numeric values may change the meaning of scale labels

Schwarz, Norbert; Knäuper, Bärbel; Hippler, Hans-Jürgen; Noelle-Neumann, Elisabeth; Clark, Leslie

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., Knäuper, B., Hippler, H.-J., Noelle-Neumann, E., & Clark, L. (1990). *Rating scales: numeric values may change the meaning of scale labels.* (ZUMA-Arbeitsbericht, 1990/10). Mannheim: Zentrum für Umfragen, Methoden und Analysen -ZUMA-. https://nbn-resolving.org/urn:nbn:de:0168-ssoar-67274

Rating Scales:
Numeric Values May Change the Meaning of Scale Labels

Norbert Schwarz, Bärbel Knäuper, Hans-J. Hippler
Elisabeth Noelle-Neumann, Leslie Clark

ZUMA-Arbeitsbericht Nr. 90/10

Zentrum für Umfragen, Methoden und
Analysen e.V. (ZUMA)
Postfach 12 21 55
6800 Mannheim 1

Seit Juli 1983 sind die ZUMA-Arbeitsberichte in zwei Reihen aufgeteilt:

Die ZUMA-Arbeitsberichte (neue Folge) haben eine hausinterne Begutachtung durchlaufen und werden vom Geschäftsführenden Direktor zusammen mit den übrigen Wissenschaftlichen Leitern herausgegeben. Die Berichte dieser Reihe sind zur allgemeinen Weitergabe nach außen bestimmt.

Die ZUMA-Technischen Berichte dienen zur hausinternen Kommunikation bzw. zur Unterrichtung externer Kooperationspartner. Sie sind nicht zur allgemeinen Weitergabe bestimmt.

The enclosed reprint

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

replaces the previously available **ZUMA-Arbeitsbericht 1990/10** by the same authors.

Public Opinion Quarterly, 1991, 55, 570-582

RATING SCALES
NUMERIC VALUES MAY CHANGE THE MEANING OF SCALE LABELS

NORBERT SCHWARZ
BÄRBEL KNÄUPER
HANS-J. HIPPLER
ELISABETH NOELLE-NEUMANN
LESLIE CLARK

Abstract Three experiments indicate that the numeric values provided as part of a rating scale may influence respondents' interpretation of the endpoint labels. In experiment 1, a representative sample of German adults rated their success in life along an 11-point rating scale, with the endpoints labeled "not at all successful" and "extremely successful." When the numeric values ranged from 0 ("not at all successful") to 10 ("extremely successful"), 34 percent of the respondents endorsed values between 0 and 5. However, only 13 percent endorsed formally equivalent values between −5 and 0, when the scale ranged from −5 ("not at all successful") to +5 ("extremely successful"). Experiment 2 provided an extended conceptual replication of this finding, and experiment 3 demonstrates that recipients of a respondent's report draw different inferences from formally equivalent but numerically different values. In combination, the findings indicate that respondents use the numeric values to disambiguate the meaning of scale labels, resulting in different interpretations and, accordingly, different subjective scale anchors.

NORBERT SCHWARZ is program director at ZUMA and Privatdozent of psychology at the University of Heidelberg. BÄRBEL KNÄUPER is research associate at ZUMA. HANS-J. HIPPLER is project director at ZUMA. ELISABETH NOELLE-NEUMANN is founder and director of the Institute für Demoskopie Allensbach, and professor emeritus of mass communication at the University of Mainz. LESLIE CLARK is assistant professor of psychology at Purdue University. The reported research was supported by grant SWF0044-6 from the Bundesminister für Forschung und Technologie of the Federal Republic of Germany to Norbert Schwarz. We thank George Bishop, Tory Higgins, and Tracy Wellens for stimulating discussions. Correspondence should be addressed to Norbert Schwarz, ZUMA, P.O. Box 12 21 55, D-6800 Mannheim, Germany.

Public Opinion Quarterly Volume 55:570–582 0033-362X/91/5504-0008$02.50

Rating scales with labeled endpoints are probably the most widely used measurement instrument in social and psychological research. Leaving some concerns about their psychometric properties aside (see Nunnally 1978), the use of these scales does not seem to be very controversial (see Dawes and Smith [1985] for a careful discussion of their general properties and for empirical and psychological justifications for their use). In general, 7-point scales seem to be best in terms of reliability, percentage of undecided respondents, and respondents' ability to discriminate between the scale values (e.g., Cox 1980). Thus, seven plus or minus two is the usual recommendation. Moreover, respondents are able to use rating scales consistently, even in telephone interviews without visual aids (e.g., Hormuth and Brückner 1985). In addition, verbal rating scales, which provide a label for each scale point, have been found to be more reliable than scales that provide labels for the endpoints only (Krosnick and Berent 1990). Finally, researchers have observed that the terms used to label the endpoints, or to designate the separate values of verbal rating scales, may affect the obtained distribution (e.g., Rohrmann 1978; Wegner, Faulbaum, and Maag 1982; Wildt and Mazis 1978). This suggests that respondents pay close attention to the meaning of the labels provided to them, much as one would hope.

Whereas the number of scale points, the inclusion or omission of a neutral point, and the choice of scale labels have received considerable attention in the literature (see Dawes and Smith 1985), the specific numeric values provided have, to our knowledge, not been the topic of theoretical analysis and empirical investigation. Apparently, researchers assume that, for example, a 7-point scale that ranges from "1" to "7" is equivalent to a 7-point scale that ranges from "−3" to "+3," as long as the same endpoint labels are provided. In the present article, we will question this assumption. Drawing on survey data from the Allensbach archive, we will first demonstrate that the specific numeric values provided in a rating scale can have a dramatic impact on the obtained results. We will then discuss different underlying processes and will test their viability in laboratory experiments.

Experiments 1 and 2: The Impact of Numeric Values

SURVEY DATA

In July 1988, the Allensbach Institute conducted a split-ballot experiment as part of a representative survey of the adult population of the

Federal Republic of Germany, using a quota sample of 1,032 respondents, based on an intersection of region, sex, and age. In face-to-face interviews, all respondents were asked to report how successful they have been in life, along an 11-point rating scale with labeled endpoints. The rating scale was presented on a show card, in the form of a ladder, and ranged either from 0 = "not at all successful" to 10 = "extremely successful," or from −5 = "not at all successful" to +5 = "extremely successful." Respondents were randomly assigned to one of the two numeric value conditions, and the question read: "How successful have you been in life, so far? Please use this ladder to tell me. This is how it works: 0 [−5, respectively] means 'not successful at all,' and 10 [+5, respectively] means that you were 'extremely successful.' Which number do you choose?"

As shown in table 1, the numeric values provided on the respective rating scales had a pronounced impact on the obtained reports, χ^2 (10) = 105.1, $p < .001$. For example, whereas 34 percent of the respondents who were given the 0–10 scale endorsed a value between 0 and 5, only 13 percent of the respondents who were given the −5 to +5 scale endorsed one of the presumably equivalent values between −5 and 0 ($z = 8.11$, $p < .0001$, for this contrast; cf. Rosenthal and Rosnow [1985]).

Coding both scales from 0 to 10, this pattern results in mean ratings of $M = 6.4$ on the 0–10, but $M = 7.3$ on the −5 to +5 version of the scale. In addition, an inspection of the distributions along both scales indicates that the responses are dislocated toward the high end of the −5 to +5 scale, as compared to the 0–10 scale. This is also reflected in markedly different standard deviations, SD = 1.03 and .56 for the 0–10 and −5 to +5 scale, respectively.

These findings may reflect either that respondents hesitated to assign themselves a negative score with regard to their success in life or that the numeric values provided on the scale influenced respondents' interpretation of the endpoint labels. Specifically, "not at all successful" may be interpreted as referring to the absence of success if combined with a numeric value of 0, but as referring to the presence of explicit failure if combined with a numeric value of −5. Before we address these possibilities in more detail, however, it is informative to consider the findings of an extended replication.

EXTENDED REPLICATION

To test the reliability of the above finding, we conducted a conceptual replication with a sample of 101 students at a German university. In a self-administered questionnaire, half of the respondents were asked to report, along 11-point rating scales, how successful they have been in

Table 1. The Impact of Numeric Scale Values on Reports along Rating Scales

0–10 Scale			−5 to +5 Scale		
Scale Value	Percentage	Cumulative	Scale Value	Percentage	Cumulative
0	. . .	. . .	−5	1	1
1	. . .	. . .	−4	. . .	1
2	2	2	−3	1	2
3	5	7	−2	1	3
4	7	14	−1	1	4
5	20	34	0	9	13
6	14	48	+1	9	22
7	20	68	+2	23	45
8	20	88	+3	35	80
9	6	94	+4	14	94
10	3	97	+5	4	98
Undecided	3	100	Undecided	2	100
N	480		*N*	552	

SOURCE.—IfD 5007, Juli 1988.

NOTE.—Percentages rounded. Data based on a quota sample of 1,032 German adults, randomly assigned to conditions. The question read, "How successful have you been in life, so far?" with scale endpoints labeled "not at all successful" (0 or −5) and "extremely successful" (10 or +5). See Appendix for full wording. χ^2 (10) = 105.1, $p < .0001$.

life and how happy a childhood they had. The other half of the respondents rated the success and childhood happiness of their parents along the same scales, thus extending the present study to proxy reports. The key manipulations consisted in variations of the numeric values and the type of endpoint label provided. Specifically, the scale ranged either from 0 to 10 or from −5 to +5, replicating experiment 1. In addition, the endpoints were either labeled "unsuccessful" and "very successful" (or, "unhappy" and "very happy," respectively) or they were labeled "not so successful" and "very successful" (or, "not so happy" and "very happy," respectively). The latter wordings were introduced to explore the impact of differentially ambiguous scale labels, and the exact wordings of the questions are given in the Appendix. In sum, these manipulations resulted in a 2 (numeric values) × 2 (scale labels) × 2 (self- or proxy reports) factorial between subjects design, which was analyzed by analysis of variance.

The obtained data provide a robust replication of the previously observed impact of numeric values on self- as well as proxy reports of success and childhood happiness. Coding both scales from 0 to 10, respondents reported higher success in life for themselves ($M = 7.38$), as well as for their parents ($M = 8.13$), along the −5 to +5 scale than along the 0–10 scale ($M = 5.96$ and 7.04 for self and parents, respectively), resulting in a pronounced main effect of numeric values ($F[1,93] = 16.21$, $p < .001$). Similarly, respondents reported higher childhood happiness along the −5 to +5 scale than along the 0–10 scale, again both for themselves ($M = 8.08$ and 6.17 for the −5 to +5 and the 0–10 scale, respectively), as well as for their parents ($M = 7.04$ and 5.38, respectively), with $F(1,93) = 5.02$, $p < .03$ for the main effect of numeric values.

The remaining effects that reached significance were of little theoretical interest and reflected that respondents perceived their parents as having been more successful in life, but less happy during their childhood. Neither the self/proxy variable, nor the ambiguity of the scale labels, however, moderated the impact of the numeric values, all Fs < 1 for the interaction terms.

DISCUSSION

As alluded to above, two processes, which are not mutually exclusive, may underlie the consistently observed impact of numeric values. As a first possibility, respondents may hesitate to assign themselves a negative score, reflecting self-presentation concerns. Although we cannot rule out an impact of self-presentation concerns, some aspects of our findings suggest that self-presentation may not be the key factor that drives the observed phenomenon. First, respondents in experi-

ment 2 provided their reports anonymously in a self-administered questionnaire, thus reducing the potential impact of social desirability (cf. Strack et al. 1990). Moreover, self-presentation considerations suggest that the impact of numeric values should be more pronounced for self- than for proxy reports, which was not the case. Finally, why should respondents hesitate to endorse a negative scale value, unless negative scale values communicate a different meaning than positive values?

These considerations suggest that respondents may have used the numeric values provided as part of the rating scale to interpret the meaning of the labels. As Woll et al. (1980, p. 60) note, "Even the most unambiguous words show a range of meaning, or a degree of 'semantic flexibility,' . . . that is constrained by the particular context in which these words occur." (See also Barclay et al. [1974] and Bransford [1979] for a more general discussion.) For example, respondents who are asked to rate their success in life need to determine what the researcher means by "unsuccessful," "not so successful," and the like. Does that term refer to the absence of remarkable successes, or does it refer to the presence of failure? Depending on how respondents interpret the term, respondents who have, for example, neither experienced particular successes nor particular failures but have done "alright" in life may choose very different scale values. Thus, the present findings may reflect that the numeric values changed the meaning of the endpoint labels, resulting in different responses, much as has been observed in studies that explicitly varied the wording of the scale labels (e.g., Rohrmann 1978; Wegner, Faulbaum, and Maag 1982; Wildt and Mazis 1978). According to this account, the finding that the different labels used in experiment 2 did not result in a differential impact of numeric values presumably reflects that the terms "unsuccessful" or "unhappy" are as ambiguous as the terms "not so successful" or "not so happy," which we introduced on the intuitive assumption that they may be more ambiguous than the former.

Note that this interpretation does not require self-presentation considerations as a necessary condition to account for the observed findings. Nevertheless, it implies that the impact of numeric values may increase with increasing self-presentation concerns, because these concerns are known to increase respondents' reluctance to endorse values with unfavorable implications (Schlenker 1980). Accordingly, semantic interpretation and self-presentation processes are not mutually exclusive. Rather, the operation of the latter requires the operation of the former in the first place.

Experiment 3 was designed to provide a more direct test of the shift in meaning hypothesis offered here, under conditions that render self-presentation concerns very unlikely. To accomplish this, respondents were given transcripts of another person's reports along rating

scales with different numeric values and were asked to draw inferences about the target person. If the numeric values change the meaning of the scale labels, this should be reflected in different inferences about the target.

Experiment 3: Inferences Based on Numeric Values

Twenty-two students at a German university participated in a study that was purportedly concerned with the accuracy of the inferences that people can draw about others on the basis of minimal information. All participants received a short description of two target persons and a transcript of these persons' responses to a survey question (see Appendix for question wordings). The first target person reported his health satisfaction along an 11-point rating scale, ranging from "dissatisfied" (0 or −5) to "very satisfied" (10 or +5). Depending on the numeric values provided on the scale, this person had allegedly checked a value of −4, or a numerically equivalent value of 1. Similarly, the second target person had allegedly rated his success on academic exams, again along 11-point rating scales, with the endpoints labeled "not so successful" (0 or −5) and "successful" (10 or +5), and had checked a minus three or a two. Subjects were randomly assigned to one of the two numeric values conditions constituted by the scale along which the target persons had allegedly given their reports.

As dependent variables, subjects were asked, in an open response format, to estimate how frequently the first target person had to see a doctor during the last month and how often the second target person had to repeat an exam due to failure. If the numeric values affect subjects' interpretation of the endpoint labels of the rating scales along which the target persons gave their reports, one should expect that they draw more extreme inferences if the numeric values range from −5 to +5 rather than from 0 to 10.

The findings support this assumption. Specifically, subjects estimated that the first target person had to see a doctor twice as often when he checked a minus four on the −5 to +5 scale ($M = 2.2$), than when he checked one on the 0–10 scale ($M = 1.0$), $F(1,20) = 4.86$, $p < .04$. Similarly, they assumed that the second target person had failed on twice as many exams ($M = 1.73$) when he checked a minus three on the −5 to +5 scale, than when he checked a 2 on the 0–10 scale ($M = .91$), $F(1,20) = 5.63$, $p < .03$.

In combination, these findings indicate that respondents drew more extreme inferences from reports given along a −5 to +5 scale than from formally identical reports given along a 0–10 scale. This is consistent

with the assumption that the numeric values presented on the rating scales affected respondents' interpretation of the meaning of the report.

Conclusions

We conclude from the reported findings that respondents may use the numeric values provided on a rating scale to disambiguate the meaning of scale labels. If the numeric values range from 0 to 10, as was the case in our studies, their very structure seems to suggest that the researcher is interested in the absence or presence of the attribute to which the scale pertains, that is, success or happiness in the above examples. If the numeric values range from −5 to +5, including a zero at the midpoint, their structure seems to suggest that the absence of the attribute corresponds to zero, whereas the negative values refer to the presence of its opposite, that is, failure or unhappiness in the above examples. In more general terms, scales that provide a continuum from negative to positive values may indicate that the researcher has a bipolar conceptualization of the respective dimension in mind, whereas scales that present only positive values may indicate a unipolar conceptualization. If so, the choice of numeric values may either facilitate or dilute the polarity implications of the endpoint labels that are provided to respondents. Accordingly, researchers may be well advised to match the numeric values that they provide to respondents with the intended conceptualization of the underlying dimension as uni- or bipolar.

Whereas this point may seem obvious in the studies reported above—where the combination of an apparently unidirectional verbal scale ("not at all" to "extremely successful") with a bipolar numeric scale may strike some readers as awkward—the strong effects obtained here suggest that the impact of numeric values should be even more pronounced when the polarity of the verbal labels themselves is more ambiguous. For example, political issue questions in the National Election Studies are intended to introduce two opposing positions, reflecting a liberal and a conservative viewpoint. Thus, respondents may be asked, "Some people believe that we should spend much less for defense. Others feel that spending should be greatly increased. Where would you place yourself on this scale? Greatly decrease defense spending (1); greatly increase defense spending (7)." The current findings suggest that using values from −3 to +3, rather than from 1 to 7, may help to emphasize the intended liberal-conservative bipolarity of these options. Similarly, our reasoning suggests that scales that

follow the format of semantic differentials (Osgood 1952) by providing polar opposites should make use of numeric values that range from negative to positive, as is frequently but not always the case. In contrast, scales that are intended to assess the intensity of a single attribute should follow a zero-to-positive-values format to emphasize that the question pertains to the absence or presence of this specific attribute, rather than the presence of its opposite.

Note, however, that the use of different scale formats may affect the obtained item variance. The conditions under which this is the case are not well understood, however, and further research is needed to explore this issue. Suppose that most people are more likely to experience success rather than failure in life. If so, using a -5 to $+5$ format restricts the meaningful response alternatives for most respondents to the positive half of the scale, resulting in reduced item variance relative to a 0–10 format. Suppose, on the other hand, that many people would experience more failure than success in life. If so, a 0–10 format would provide fewer meaningful response alternatives for these respondents than a -5 to $+5$ format, resulting in reduced item variance in the former case. This suggests that the choice of a scale format should be based on researchers' knowledge about the relevant distribution to avoid undesirable restrictions in item variance.

Finally, the present reasoning bears on the comparability of data obtained under different administration modes (see Schwarz et al. [1991] for a general conceptualization of mode effects). The use of numeric values is most prevalent in face-to-face interviews and self-administered surveys, whereas the increasing use of an unfolding format in telephone interviews avoids the presentation of numeric values. Based on the present findings, we hypothesize that data obtained in an unfolding format are more compatible with verbal rating scales that present labels for each scale point, or with scales that present unnumbered boxes or similar devices, than with scales that use numeric values. These considerations echo the general insight that the absolute values, or marginals, obtained in response to any survey question are difficult to interpret (see Schuman 1986). To what extent changes in scale format affect the obtained relationships or changes over time, on the other hand, needs to be explored in future research.

From a theoretical perspective, the present findings support the general conclusions that we have drawn from related research into the impact of response alternatives on frequency reports of mundane behaviors and related judgments (see Schwarz 1990; Schwarz and Hippler 1987). Far from being "neutral measurement devices," the response alternatives that are provided to respondents do constitute a source of information that respondents actively use in determining their task and in constructing a reasonable answer. While survey meth-

odologists have traditionally focused on the information that is provided by the wording of the question, we need to pay equal attention to the information that is conveyed by apparently formal features of the questionnaire. Respondents apply many of the rules that govern the conduct of conversation in everyday life (cf. Clark 1985; Grice 1975) to the survey interview. In doing so, they assume that every contribution to the ongoing conversation is relevant and meaningful. In the survey interview, the researcher's contributions include the response alternatives, the numeric values of rating scales, and the ordering of questions, as well as other features of questionnaire design. (See Schwarz and Hippler [in press], Schwarz and Strack [1991], and Strack and Schwarz [in press] for reviews and theoretical analyses.) The analysis of the informational functions of apparently formal features of questionnaire design is therefore a key task in the collaboration of cognitive psychologists and survey researchers.

Appendix
English Translations of Question Wordings

EXPERIMENT 1

See text for question wording.

EXPERIMENT 2

Self-reports.

"What is your opinion: How happy was your own childhood?"—"unhappy" (0 or −5) to "happy" (10 or +5); or "not so happy" (0 or −5) to "very happy" (10 or +5), respectively.

"How successful have you been in life?"—"unsuccessful" (0 or −5) to "very successful" (10 or +5); or "not so successful" (0 or −5) to "very successful" (10 or +5).

Proxy reports.

"What is your opinion: How happy was the childhood of your parents?"—"unhappy" (0 or −5) to "happy" (10 or +5); or "not so happy" (0 or −5) to "very happy" (10 or +5), respectively.

"How successful have your parents been in life?"—"unsuccessful" (0 or −5) to "very successful" (10 or +5); or "not so successful" (0 or −5) to "very successful" (10 or +5).

EXPERIMENT 3

Scenario 1. Peter K., an MBA student, was asked how successful he was on his exams. He gave the following response: "How successful have you

been on your exams?"—"not so successful" (0 or −5) to "very successful" (10 or +5). Depending on type of scale, Peter K. allegedly checked "2" or "−3."

Scenario 2. Manfred D. was asked how satisfied he is with his health. He gave the following response: "How satisfied are you with your health?"—"not so satisfied" (0 or −5) to "very satisfied" (10 or +5). Depending on type of scale, Manfred D. allegedly checked "1" or "−4."

German Question Wording

EXPERIMENT 1

"Wie erfolgreich waren Sie bisher in Ihrem Leben? Sagen Sie es bitte nach dieser Leiter hier. Es geht so: Null (−5) bedeutet überhaupt nicht erfolgreich und 10 (+5) bedeutet, Sie waren bisher außerordentlich erfolgreich. Welche Zahl nehmen Sie?"

EXPERIMENT 2

Self-reports.

"Was meinen Sie: Wie glücklich war Ihre eigene Kindheit?"—"unglücklich" (0 or −5) to "glücklich" (10 or +5) "nicht so glücklich" (0 or −5) to "glücklich" (10 or +5).

"Wie erfolgreich waren Sie im Leben?"—"erfolglos" (0 or −5) to "sehr erfolgreich" (10 or +5) "nicht so erfolgreich" (0 or −5) to "sehr erfolgreich" (10 or +5).

Proxy reports.

"Was meinen Sie: Wie glücklich war die Kindheit Ihrer Eltern?"—"unglücklich" (0 or −5) to "glücklich" (10 or +5) "nicht so glücklich" (0 or −5) to "glücklich" (10 or +5).

"Wie erfolgreich waren Ihre Eltern im Leben?"—"erfolglos" (0 or −5) to "sehr erfolgreich" (10 or +5) "nicht so erfolgreich" (0 or −5) to "sehr erfolgreich" (10 or +5).

EXPERIMENT 3

Scenario 1. Der BWL-Student Peter K. wurde gefragt, wie erfolgreich er im BWL-Vordiplom war. Er machte die folgende Angabe: "Wie erfolgreich waren Sie in Ihrem Vordiplom?"—"nicht so erfolgreich" (0 or −5) to "sehr erfolgreich" (10 or +5).

Scenario 2. Manfred D. wurde gefragt, wie zufrieden er mit seiner Gesundheit ist. Er machte die folgende Angabe: "Wie zufrieden sind Sie mit Ihrer Gesundheit?"—"nicht so zufrieden" (0 or −5) to "sehr zufrieden" (10 or +5).

References

Barclay, J. R., J. D. Bransford, J. J. Franks, N. S. McCarrell, and K. E. Nitsch. 1974. "Comprehension and Semantic Flexibility." *Journal of Verbal Learning and Verbal Behavior* 13:471–81.

Bransford, J. D. 1979. *Human Cognition: Learning, Understanding, and Remembering*. Belmont, CA: Wadsworth.

Clark, H. H. 1985. "Language Use and Language Users." In *Handbook of Social Psychology*, ed. G. Lindzey and E. Aronson, 2:179–232. New York: Random House.

Cox, E. P. 1980. "The Optimal Number of Response Alternatives for a Scale: A Review." *Journal of Marketing Research* 17:407–22.

Dawes, R. M., and T. Smith. 1985. "Attitude and Opinion Measurement." In *Handbook of Social Psychology*, ed. G. Lindzey and E. Aronson, 2:509–66. New York: Random House.

Grice, H. P. 1975. "Logic and Conversation." In *Syntax and Semantics: Speech Acts*, ed. P. Cole and J. L. Morgan, 3:41–58. New York: Academic Press.

Hormuth, S. E., and E. Brückner. 1985. "Telefoninterviews in Sozialforschung und Sozialpsychologie: Ausgewählte Probleme der Stichprobengewinnung, Kontaktierung und Versuchsplanung." *Kölner Zeitschrift für Soziologie und Sozialpsychologie* 37:526–45.

Krosnick, J. A., and K. M. Berent. 1990. "The Impact of Verbal Labeling of Response Alternatives and Branching on Attitude Measurement Reliability in Surveys." Paper presented at the annual meeting of the American Association for Public Opinion Research, Lancaster, PA.

Nunnally, J. C. 1978. *Psychometric Theory*. New York: McGraw-Hill.

Osgood, C. 1952. "The Nature and Measurement of Meaning." *Psychological Bulletin* 49:197–237.

Rohrmann, B. 1978. "Empirische Studien zur Entwicklung von Antwortskalen für die sozialwissenschaftliche Forschung." *Zeitschrift für Sozialpsychologie* 9:222–45.

Rosenthal, R., and R. L. Rosnow. 1985. *Contrast Analysis*. Cambridge: Cambridge University Press.

Schlenker, B. R. 1980. *Impression Management*. Belmont, CA: Brooks/Cole.

Schuman, H. 1986. "Ordinary Questions, Survey Questions, and Policy Questions." *Public Opinion Quarterly* 50:431–42.

Schwarz, N. 1990. "Assessing Frequency Reports of Mundane Behaviors: Contributions of Cognitive Psychology to Questionnaire Construction." In *Research Methods in Personality and Social Psychology*, ed. C. Hendrick and M. S. Clark, pp. 98–119. Review of Personality and Social Psychology, vol. 11. Beverly Hills, CA: Sage.

Schwarz, N., and H.-J. Hippler. 1987. "What Response Scales May Tell Your Respondents: Informative Functions of Response Alternatives." In *Social Information Processing and Survey Methodology*, ed. H.-J. Hippler, N. Schwarz, and S. Sudman, pp. 163–78. New York: Springer Verlag.

———. In press. "Response Alternatives: The Impact of Their Choice and Ordering." In *Measurement Error in Surveys*, ed. P. Biemer et al. Chichester: Wiley.

Schwarz, N., and F. Strack. 1991. "Context Effects in Attitude Surveys: Applying Cognitive Theory to Social Research." In *European Review of Social Psychology*, ed. M. Hewstone and W. Stroebe, 2:31–50. Chichester: Wiley.

Schwarz, N., F. Strack, H.-J. Hippler, and G. Bishop. 1991. "Psychological Sources of Response Effects in Surveys: The Impact of Administration Mode." In "Special Issue: Cognitive Aspects of Survey Methodology," ed. J. Jobe and E. Loftus. *Applied Cognitive Psychology* 5:193–212.

Strack, F., and N. Schwarz. In press. "Communicative Influences in Standardized Question Situations: The Case of Implicit Collaboration." In *Language and Social Cognition*, ed. K. Fiedler and G. Semin. Beverly Hills, CA: Sage.

Strack, F., N. Schwarz, B. Chassein, D. Kern, and D. Wagner. 1990. "The Salience of Comparison Standards and the Activation of Social Norms: Consequences for Judgments of Happiness and Their Communication." *British Journal of Social Psychology* 29:303–14.

Wegner, B., F. Faulbaum, and G. Maag. 1982. "Die Wirkung von Antwortvorgaben bei Kategorialskalen.," *ZUMA-Nachrichten* 10:3–20.

Wildt, A. R., and M. B. Mazis. 1978. "Determinants of Scale Responses: Label verus Position." *Journal of Marketing Research* 15:261–67.

Woll, S. B., D. G. Weeks, C. L. Fraps, J. Pendergrass, and M. A. Vanderplas. 1980. "Role of Sentence Context in the Encoding of Trait Descriptors." *Journal of Personality and Social Psychology* 39:59–68.

ZUMA-Arbeitsberichte

80/15 Gerhard Arminger, Willibald Nagl, Karl F. Schuessler
Methoden der Analyse zeitbezogener Daten. Vortragsskripten der ZUMA-Arbeitstagung vom 25.09. - 05.10.79

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 Complex 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 strcuture

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 "interv^{+}" und "THIS"

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

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 Kreiselmaier, 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, Fritz Strack, Hans-Peter Mai
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.