Measuring Information Accessibility and Predicting Response-Effects: The Validity of Response-Certainties and Response-Latencies

Volker Stocké[1]

Abstract

Respondents' reports about the frequency of everyday behavior are often found to differ considerably when either low- or high-frequency response scales are used to record the answers. It has been hypothesized that the susceptibility to this type of response effect is determined by the cognitive accessibility of the respective target information in respondents' memories. The *first* aim of the present paper is to test this hypothesis using two alternative, individual level indicators of the cognitive accessibility of information. These measures are the subjects' self-reported response certainty and the time needed to answer the question under consideration. A *second* issue is how response certainties and response latencies should be transformed prior to data analysis in order to maximize their predictive power for response effects. Accordingly, the ability of untransformed measures to predict scale effects is compared with that of logarithmic, square-root and reciprocally transformed versions. The empirical results show that untransformed response certainties and response latencies are equally valid predictors of whether and to what extent subjects' answers are affected by the presentation of response options. A square-root transformation is found to have no effect on both measures, whereas a logarithmic transformation slightly improves the validity of response certainties. In contrast, a reciprocal transformation proves to have a substantially positive effect on both measures and improves their ability to predict the reliability of respondents' survey reports.

[1] Sonderforschungsbereich 504 'Rationality Concepts, Decision Behavior and Economic Modeling', University of Mannheim, L13, 15; D-68131 Mannheim, Germany; vstocke@uni-mannheim.de

1   Introduction

Research has shown that survey respondents' answers are often influenced by seemingly irrelevant differences in the way response options are presented. For instance, reports were found to differ substantially, when numbers 0 to 10 or 1 to 11 were used to mark the response categories on an ordered response scale (Schwarz et al., 1998). One particularly important class of effects are those observed when the frequency of everyday behavior is recorded with response categories, which each represents a distinct frequency range. This is done in order to reduce the cognitive demand necessary to answer behavioral frequency questions, such as for example the number of products purchased during a reference period. However, the way frequency ranges were created was found to affect the frequency reports. Low frequency scale categories at the lower end of the response continuum are more narrow and therefore more numerous at this end of the frequency continuum. In contrast, high frequency scales provide more detailed response categories at the upper end of the continuum. When high- rather than low-frequency response scales were used to collect data, experimental studies reported higher consumer expenditures (Menon et al., 1997, Winter, 2002), more frequent sexual activities (Schwarz and Scheuring, 1988) and more feelings of criminal threat (Gaskell et al., 1994). Furthermore, the type of frequency scales affects subjects' reports about the prevalence of nightmares (Ji et al., 2000) and how often they undertake cultural activities (Bless et al., 1992).

One argument is that an insufficient cognitive accessibility of the requested information is the pivotal determinant for how strongly respondents are affected by differences in the way response scales are presented (Schwarz and Hippler, 1987)[2]. This hypothesis was supported by research in which the susceptibility to scale effects was compared between different subgroups of respondents using varying question topics. In the case of groups of respondents and question topics, in which the respective information was more likely to be accessible to memory, the effects of different types of response scales were attenuated (Schwarz, 1999). However, what has not yet been tested is whether and what measurements of information accessibility at the individual level predict the strength of response effects. Such a test is fruitful, since it provides stronger empirical support for the theoretically assumed role of information accessibility to response effects and about which the accessibility indicator is the best predictor. Furthermore, a valid indicator allows survey researcher to judge the reliability of data provided by different groups of respondents.

---

[2] In the present paper 'cognitive accessibility' refers to the intensity with which information has been encoded in one's memory and therefore the ease with which respondents can retrieve this

Two groups of individual-level measures of information accessibility can be differentiated. *First*, there are meta-cognitive indicators, which are based on respondents' subjective judgments about aspects of the response process or the quality of the resulting answers. Subjects' self-reports about how certain they feel about a particular answer can be regarded as the most often used meta-cognitive indicator (for an overview c.f. Wegener et al., 1995). *Second*, operational indicators employ directly observable and therefore objective characteristics of the response process or the resulting answers. The time respondents need to answer questions is the most prominent indicator in this group of measures (Bassili 1996a, Bassili and Fletcher, 1991, Fazio, 1990). Whether response certainties or response latencies are more valid indicators of information accessibility and a better predictor of response effects has not been compared systematically.

When used as an indicator of information accessibility in applied research, response latencies are often transformed in order to reduce their characteristically skewed distribution. Three transformations are often used for this purpose: the natural logarithm, the square-root and the reciprocal transformation. Whether these transformations have a positive effect on the validity of response latencies and maximizes their predictive power has not been studied systematically. Although the distribution of response certainties is often skewed to a considerable degree as well, there are no studies in which transformed versions of this measure are used to predict information accessibility.

The present paper has three closely related aims: *first*, to analyze whether individual differences in information accessibility can predict how strongly respondents' answers differ when either high- or low frequency scales are used to record behavioral frequencies; *second,* to address the question of whether response certainties or response latencies are the more valid indicator of information accessibility and the better predictor of the analyzed type of response effects; and *third*, to examine the effect of different transformations on the ability of both measures to predict how strongly respondents are influenced by different response scales. The dependent variable is the respondents' reports about the length of their daily TV consumption.

2 Information accessibility and the effect of response scales

A series of split-ballot experiments prove that respondents' reports about their daily TV consumption differ considerably when either high- or low-frequency scales are used to record the answers (Bless et al., 1992; Menon et al., 1995; Rockwood et al., 1997; Schwarz, 1988; Schwarz and Bienias, 1990; Schwarz et

---

information. Conceptually, this refers to the degree of 'chronic' rather than situational information accessibility.

al., 1985; Schwarz and Hippler, 1987; Stocké, 2001; Winter, 2002). In these studies, the response continuum of the number of hours respondents watched TV was categorized as follows: high frequency scales ranged in half hour steps from 'up to 2 hours' to 'more than 4.5 hours', while low frequency scales ranged between 'not at all' to 'more than 2.5 hours'. Responses on these scales were compared by computing the proportion of subjects who reported behavioral frequencies of '2.5 hours of TV or less' and 'more than 2.5 hours of TV'. According to the results of the first experiment, in the case of the high-frequency scale 37.5 percent of subjects, and under the condition of the low-frequency scale 16.2 percent, were classified in the category of 'more than 2.5 hours TV' (Schwarz et al., 1985).

These response effects are explained on the basis of two assumptions. *First*, the requested information about the frequency of everyday behavior, in the present case the daily TV consumption, is not available as episodic information in the subjects' Thus respondents cannot simply count instances of this behavior in order to reach an answer. *Second*, respondents are assumed to be cooperative and therefore motivated to answer the question as correctly as possible (Schwarz et al., 1985). Faced with this dilemma, subjects as a *first* step use the respective response scale as a frame of reference to infer the median TV consumption in society, which they assume is represented by the middle response option of the scale. In a *second* step, subjects use this reference point in order to form a judgment about how their own TV consumption compares with the assumed average in society. In a *third* step, respondents select a response option, which is, according to their previous judgments, appropriately located above or below the middle response category. Since the middle response option and therefore the starting point of this response heuristic differs in high- and low-frequency scales, the identical inferential process results in different behavioral reports.

There is empirical support of the assumed informative function of response scales in studies in which respondents estimate higher TV consumption in society, when they have been presented with a high- rather than a low-frequency scale (Schwarz and Hippler, 1987). Furthermore, it has been found that the type of response scale has stronger effects on proxy reports about the TV consumption of friends, compared with responses about their own television viewing behavior (Schwarz and Bienias, 1990). Since subjects' knowledge about TV consumption patterns among their peers is likely to be even more restricted than knowledge of their own habits, this result supports the role of information accessibility as a determinant of scale effects. In other results, there were stronger differences in the response behavior for questions about daily TV consumption, compared to those by students when asked about their average grades from the previous year (Rockwood et al., 1997). Since the former information can be assumed to be less salient and therefore cognitively less accessible, these results were regarded as evidence of the role of information accessibility as well. Furthermore, older respondents' reports about the frequency of their meat consumption were more

strongly affected by the type of frequency scale than by that in the case of younger respondents (Schwarz, 1999). This difference in susceptibility may result from short-term memory deterioration accompanying increasing age and resulting in poorer information accessibility. Interestingly, older respondents proved to be less prone to scale effects when responding to the frequency of health-related symptoms. This reversal in the susceptibility of scale effects can probably be explained by an increased concern with health-related topics, which increases with age and because the respective information is then more intensively encoded in the memory.

3   Response latencies and response certainties as indicators of information accessibility

There is empirical evidence of the validity of response latencies as a measure for information accessibility in research on the determinants of attitude-behavior consistency, the degree of attitude stability and how strong attitude reports are affected by response effects. Here, response time is an indicator of the degree to which an evaluation is associated with an attitude object and the ease with which respondents can therefore answer an attitude question. Thus, response latencies represents an important aspect of attitude strength. Since strong attitudes have been found to be more predictive of behavior and more resistant to change, response latencies can be expected to predict both attitude properties (Krosnick and Petty, 1995). Selecting an attitude answer in surveys can be regarded as a special case of attitude-behavior consistency. Accordingly, the previous arguments about the association between response latencies and attitude properties should apply to the degree of respondents' susceptibility to response effects as well: quick responses to attitude reports should be less affected by irrelevant factors in the response situation.

In a series of empirical studies, the time necessary to answer questions about attitudes towards political candidates predicted the relationship between observed voting behavior and these evaluations: the faster the attitude questions were answered, the more they predicted behavior (Bassili, 1993, Bassili, 1995; Bassili and Bors., 1997; Fazio and Williams, 1986; Fletcher, 2000). In the field of consumer research, quick responses, rather than slow ones about the evaluation of different products, were found to predict the consumers' intention to use these products (Kokkinaki and Lunt, 1997). Quickly expressed attitudes were also found to be more resistant to persuasion: attitudes toward gender-related affirmative action and toward pornography were more resistant to counter arguments, when response latencies were relatively short (Bassili, 1996b).

However, empirical evidence about the ability of response latencies' to predict how strongly respondents' answers are affected by response effects, is

inconsistent. Thus, it has been frequently observed that reports about attitudes toward a general liberalization of abortion were more positive when reported before, rather than after, the question about the legitimacy of abortion in cases of pregnancy from rape. This question of order effect has been found to be less pronounced for respondents with relatively fast response latencies when answering attitude questions (Stocké, 2002a). Furthermore, attitudes towards surveys increasingly predicted the respondents' susceptibility to social desirability bias, when these attitudes were reported fast rather than slow (Stocké, 2002b). In contrast, an extensive study found only negative evidence for the predictive power of response latencies for a range of different response effects: the response speed did not explain the strength of response order effects, influences from middle response options and to what degree responses were affected by acquiescence (Bassili and Krosnick, 2000). Whether response latencies are a valid indicator of the strength of response effects in the field of factual survey questions has not been analyzed.

Results about the validity of self-reported response certainties are mixed as well. Research on the determinants of attitude-behavior consistency has found that attitudes toward political candidates predicted respondents' voting behavior better when these attitudes were characterized as certain (Warland and Sample, 1973). Furthermore, subjects' attitudes toward the liberalization of abortion were substantially more related to their behavioral intentions in this area, when subjects felt increasingly more certain about their attitude reports (Renata, 1999). In contrast, there was no difference in the correlation between self-reported environmental consciousness and behavioral reports according to the certainty of the attitude judgments (Mielke, 1985). Other studies have tested whether the response certainty predicts the respondents' susceptibility to question order effects and only found negative evidence (Krosnick and Schuman, 1988, Schuman et al., 1981). What has been tested as well was whether the effect of preceding questions about liberal or conservative issues on the respondents' support for more welfare spending and improving defendants' rights decreases with greater response certainty (Lavine et al., 1998). Here, the size of this context effect was found to decrease with increasing response certainty. But, this moderator effect did not reach statistical significance. Positive results have been found in a study in which the answers about attitudes toward abortion proved to be less susceptible to question order effects when subjects characterized these evaluations as more certain (Stocké, 2002a). In this study, the predictive power of the certainty measure was directly compared to that of response latencies. In a multivariate analysis, the initially observed relevance of response certainties was strongly reduced and not statistically significant when the moderating role of response latencies was controlled at the same time.

4 Effect of different transformations

Several kinds of transformations are applied in order to reduce the characteristic positive skewness of response latency data and to minimize the effect of outliers when predicting the degree of information accessibility. Most often a natural logarithm, a square root or a reciprocal transformation is used (Bargh and Chartrand, 2000). The available evidence for the validity of raw response latencies and all transformed versions of this measure is positive. The only exceptions are square root transformed response times. However, no comparative research has been done on the relative effect of different transformations on the validity of response latencies.

Bassili (1996b) successfully utilized raw response latencies in order to predict the degree of attitude stability. In research from the area of attitude-behavior consistency, response latencies have been used without any transformation (Bassili and Bors, 1997; Bassili, 1995; Fletcher, 2000). In these studies, raw response latencies proved to be a valid predictor of the degree to which political attitudes explain subjects' behavior toward the respective attitude object. Other research shows the concurrent validity of raw response latencies: self-reports about the importance of different attitude topics are significantly related to the speed with which the attitude questions are answered (Bauman and Dent, 1982). In this study, a log transformation did not affect the predictive power of the latency data.

Researchers have successfully used the natural log of response latencies in order to predict the increasing cognitive accessibility of subjects' party identification with the length of the ongoing election campaign (Mulligan et al., 2003). Reciprocally transformed response latencies have been found to predict how strongly subjects' attitudes toward consumer products determined the probability that these products are actually chosen (Kokkinaki and Lunt, 1997). Another study using reciprocally transformed response latencies proved that the accessibility of preferences about different TV programs successfully explains the strength of respondents' self-reported preferences (LaBarbera and MacLachlan, 1979). In contrast, square-root transformed response latencies failed to predict how strongly respondents' answers were affected by four different kinds of response effects (Bassili and Krosnick, 2000).

5 Hypothesis

In the following empirical study two hypotheses are tested.

- *Hypothesis 1:* How strong individual respondents' answers about their daily TV consumption are affected by different presentations of the response scale is determined by the cognitive accessibility of the requested information. Thus,

response certainties and response latencies as accessibility measures are expected to moderate the effect of scale types on the response behaviour.

- *Hypothesis 2:* The transformation of response certainties and response latencies reduces the disturbing effect of outliers when respondents' susceptibility to response effects is predicted. Thus, the analysed kinds of transformations are expected to have a positive effect on the predictive power of accessibility measures. However, we do not have a hypothesis about which transformation will be most successful.

6   Empirical study

In order to answer our research questions, we utilized a split ballot experiment in which respondents were randomly assigned to one of two experimental conditions. Under both conditions, subjects were asked to report how many hours they watch television every day. This question was answered either on a high- or a low-frequency response scale. We analyzed whether the typically observed differences in response behavior between the experimental conditions can be predicted with response certainties and the time needed to answer the TV-consumption item. Furthermore, we compared the predictive power of differently transformed versions of both measures.

6.1   Sample

The respondents in this study were a multi-stage, local probability sample of residents in the metropolitan area of Mannheim, Germany (about 300,000 inhabitants). In the first step, households were listed using a random walk procedure. In the second step, respondents were selected among adult residents in the households using the 'last-birthday' method. Altogether 110 interviews were conducted with 53.6 percent female respondents and 46.5 percent male. The mean age of the participants was 46.8 years and they had completed an average of 10.9 years of schooling. The sample consisted of 13.6 percent blue-collar workers, 65.5 percent white collar workers, 9.1 percent self-employed and 11.2 percent subjects not participating in the labor market. The response rate was 34.0 percent.

6.2   Procedure

Data was collected with computer assisted face-to-face interviews in the respondents' homes. Subjects were randomly assigned to one of the two experimental conditions. Depending on these conditions, the interviewers presented show cards with either a low- or a high-frequency response scale. The

experiment was part of a longer interview, which took an average of 58 minutes to be completed. Respondents answered the question about their daily TV consumption and the interviewer recorded the response latency. Immediately thereafter, subjects answered the question about how certain they were about this answer. In the advance letter, the survey topic was described to being about 'habits in everyday life' and 'social problems in society'. Neither the experimental character of the survey nor the fact that response times were measured, were told to the respondents before the end of the interview.

6.3 Operationalization

The high- and low-frequency response scales, the response certainty and response latencies were operationalized as follows:

    - *High and low requency scales:* For both types of frequency scales, the response continuum representing the length of daily TV consumption was categorized into 7 distinct response options. In the case of the low-frequency version, the extreme options were defined by 'no TV consumption' and 'more than 2 hours of TV consumption'. For the high frequency scale these endpoints were 'up to 2 hours' and 'more than 4.5 hours'. Between these extreme response alternatives, both scales ranged in steps of a half hour (Table 1 in the 'descriptive results' section for a presentation of both scales)[3]. Answers on both versions of the response scale can be compared when respondents are classified into groups with less than 2.5 hours and 2.5 or more hours of daily TV consumption.

    - *Raw response certainty*: Directly after the question about the length of daily TV consumption, subjects were asked to report how certain they felt about their answer to this question. Responses were recorded using a seven point likert scale, with endpoints labeled with 'absolutely certain' (scale value 1) and 'not at all certain' (scale value 7)[4].

    - *Raw response latencies:* Response times were recorded together with the responses in a sequence involving four stages during the computer assisted interviews. In the *first* stage, interviewers read the question from the computer screen and switched on the time measurement directly after the question text had been read. *Second*, the time measurement was switched off immediately after respondents answered the question. In the *third* stage of the data collection sequence, the interviewer entered the response into the laptop computer. In the *fourth* stage, the interviewer judged whether the time recorded represented not more and not less than the time which was necessary to answer the question. This

---

    [3] The question wording reads as follows: 'Would you please tell me for a typical weekday, how long do you watch television? Please tell me the number of the appropriate response option from this list'.

    [4] The question reads as follows: 'Would you please indicate how certain you are about your answer about the length of your daily TV consumption'.

was not the case when respondents asked clarifying questions, had to be probed in order to give an appropriate answer or when subjects were distracted by external factors. Under all these conditions the interviewers coded the response latencies to be invalid. This was the case for 21.8 percent of the answers and these cases were excluded from the following analysis. The precision of response latency measurement, based on the technical restrictions of the interview software, is one-hundredth of one second.

   - Transformed response certainties and response latencies: In order to test whether the transformation of response latencies and response certainties affects their validity as an indicator of information accessibility, different versions of both measures have been computed and were included in the analysis. In a first version, the natural logarithm and in a second, the square root function were used to transform the data. In a third version, the reciprocal of the measures was computed. This measure is defined for raw response certainties (RC) as 1/RC and for the raw response latencies (RL) as 1/RL. The reciprocal values were multiplied by –1 and then +1 was added in order to match the direction of all other versions of the accessibility indicator. Accordingly, low values on all eight versions of the indicator represents high accessibility and high values indicate a low cognitive accessibility of the requested information.

6.4   Results

The results of our data analysis are presented in three sections. *First*, the respondents' answers about their daily TV-consumption and the explanatory variables are described. In this section, the associations between the different indicators of information accessibility and therefore their convergent validity is analyzed as well. In the *second* section, it is tested whether raw response certainties and response latencies predict how strongly individual respondents' answers are influenced when either high- or low-frequency scales are used for data collection. In the *third* and final section of the empirical analysis the predictive power of differently transformed variants of both accessibility measures is compared with the validity of their original versions.

6.4.1 Descriptive results

In Table 1 the respondents' answers about their daily TV consumption are presented for the conditions of high- and low-frequency scales. As in other experiments, responses were strongly affected by the type of response scale. Whereas under the condition of a low frequency scale, 80.4 percent of respondents reported watching TV between 0 and 2.5 hours a day. This proportion is only 57.4 percent for the high frequency scale.

Table 1: Reports about daily TV consumption for high- and low-frequency scales.

Low-frequency response scale	N	%	High-frequency response scale	N	%
(1) not at all	1	1.8			
(2) up to half a hour	4	7.1			
(3) between 0.5 and 1 hour	7	12.5			
(4) between 1 and 1.5 hours	17	30.4			
(5) between 1.5 and 2 hours	10	17.9	(1) up to 2 hours	24	44.4
(6) between 2 and 2.5 hours	6	10.7	(2) between 2 and 2.5 hours	7	13.0
(7) more than 2.5 hours	11	19.6	(3) between 2.5 and 3 hours	13	24.1
			(4) between 3 and 3.5 hours	4	7.4
			(5) between 3.5 and 4 hours	5	9.3
			(6) between 4 and 4.5 hours	0	0.0
			(7) more than 4.5 hours	1	1.9

Summarized response behavior

	N	%		N	%
0 to 2.5 hours	45	80.4	0 to 2.5 hours	31	57.4
More than 2.5 hours	11	19.6	More than 2.5 hours	23	42.6
Total	56	100	Total	54	100

According to their self-reports, respondents are very confident about the time they spend watching TV every day. On the response scale between 1 (absolutely certain) and 7 (not at all certain) the sample mean is 1.9 and therefore close to the certainty endpoint of the scale (Table 2). In contrast, the average raw response latency is 6.6 seconds and indicates a relatively long response process: answering the target question seems to be objectively a difficult task.[5] The conclusions about the degree of information accessibility therefore differ whether the subject self-reports or the response latencies as an objective indicator are taken into account.

Raw response certainties and response latencies deviate to a different, but in both cases significant, degree from a normal distribution. The Kolmogorov-Smirnov test statistics are 0.30 ($p < 0.01$) for the certainty measure and 0.12 ($p < 0.01$) for response latencies. According to this result, the distribution of certainties is skewed more positively than the response latencies. Furthermore, our results show that the transformation of data is not very effective in solving this problem. In the case of response certainties, none of the transformations reduces the deviation from a normal distribution. The Kolmogorov-Smirnov test statistics are 0.34 ($p < 0.01$) for log certainties, 0.33 ($p < 0.01$) for the square root transformed version and 0.36 ($p < 0.01$) for the reciprocal transformation. In the case of

---

[5] On average it took respondents 4.3 seconds to answer each of the other 126 questions in the interview. The response time necessary to report the length of their daily TV consumption is therefore 2.3 seconds longer compared with the other questions in the questionnaire.

response latencies, a square root transformation reduces to a substantial degree the skew of the distribution. Here, the Kolmogorov-Smirnov test statistic is 0.06 (p > 0.1) and indicates a non-significant deviation from the normal distribution. However, all other transformations of response time proved to be ineffective in this respect. The Kolmogorov-Smirnov test is 0.14 (p < 0.01) for the log transformation and 0.26 (p < 0.01) for the reciprocal transformation. In summary, with the exception of square root transformations in the case of response latencies, different types of transformation proved of little value in improving the shape of data distributions.

Table 2: Descriptive statistics for the different indicators of information availability.

	Min/Max	Mean	Median	Standard deviation
Raw response uncertainty [a]	1/7	1.86	1.00	1.30
Logarithmic response uncertainty	0.00/1.95	0.45	0.00	0.56
Square root response uncertainty	1.00/2.65	1.30	1.00	0.41
Reciprocal response uncertainty	0.00/0.86	0.27	0.00	0.32
Raw response latency [b]	27.00/3263.00	662.27	566.00	529.10
Logarithmic response latency	3.30/8.09	6.15	6.34	0.93
Square root response latency	5.20/57.12	23.82	23.79	9.80
Reciprocal response latency	0.96/1.0	0.99	1.00	0.005

[a] Scale value 1= 'absolutely certain', 7= 'not at all certain'.

[b] The unit of measurement is one-hundredth of a second; Sample size: N=110.

The structure of correlations between different accessibility indicators is as expected in the case of their convergent validity: the association between all versions has a positive sign (Table 3). In particular, relations between differently transformed versions of the same type of accessibility indicator are strong and in each case statistically significant. Correlations between response certainties range between 0.88 and 0.99 and those between response latencies between 0.51 and 0.96. The positive associations between response certainties and response latencies indicates that the more time subjects need to answer the TV consumption question, the more uncertain they feel about their reports. However, the strength of these relationships is weak and in many cases statistically insignificant. The inter-indicator correlation ranges between 0.05 and 0.23. Accordingly, response certainties and response latencies are related, but by no means identical measures.

Table 3: Correlations between different indicators of information accessibility (Pearson's correlation coefficients).

	RC	LOG-RC	SR-RC	REC-RC	RL	LOG-RL	SR-RL	REC-RL
Raw response certainty (RC)	1.00	--	--	--	--	--	--	--
Logarithmic RC (LOG-RC)	0.96***	1.00	--	--	--	--	--	--
Square root RC (SR-RC)	0.99***	0.99***	1.00	--	--	--	--	--
Reciprocal RC (REC-RC)	0.88***	0.98***	0.94***	1.00	--	--	--	--
Raw response latency (RL)	0.15	0.20**	0.18*	0.23**	1.00	--	--	--
Logarithmic RL (LOG-RL)	0.11	0.15	0.13	0.17*	0.85***	1.00	--	--
Square root RL (SR-RL)	0.14	0.19**	0.17*	0.21**	0.96***	0.96***	1.00	--
Reciprocal RL (REC-RL)	0.05	0.07	0.06	0.07	0.51***	0.84***	0.67***	1.00

Sample size: N = 110; Significance: *** $p \leq 0.01$; ** $p \leq 0.05$; * $p \leq 0.1$

6.4.2 Predictive power of raw response certainties and response latencies for respondents' susceptibility to response effects

In the following section, we consider whether and to what extent untransformed versions of response certainties and response latencies predict the degree to which respondents' answers are affected by the type of response scale. As presented in Table 1, the dependent variable in the following analysis is the dichotomous and across scale versions comparable response variable. Logistic regression analysis is utilized to analyze the data. According to our hypothesis, the effect of response scales is expected to increase when the requested information becomes less accessible. An interaction parameter between the type of response scale on the one hand and response certainties as well as response latencies on the other should prove to be a significant predictor of response behavior.

In the *first* stage, the results of our analysis, as shown in Table 4, prove that using either a high- or a low-frequency scale has a statistically significant effect on the length of the reported daily TV consumption (c.f. model 1). Furthermore, the respondents' education is a significant and their age a marginally significant predictor of their reported TV consumption. Less educated respondents reported watching more TV and younger subjects tended to report less TV consumption[6].

---

[6] Additionally, the effect of subjects' sex, socio-economic status, income, religious denomination and marital status on response behavior has been tested (results not reported). None

Table 4: Effect of response scales on the probability of reporting high vs. low TV consumption and the moderating role of raw response certainties and response latencies for the strength of this response effect (logistic regression results)7.

|	Model 1 B (Wald-Statistic)	Model 2 Response Certainty B (Wald-Statistic)	Model 3 Response Latency B (Wald-Statistic)
(1) EDUCATION (years)	-0.35 (5.96)**	-0.26 (2.91)*	-0.46 (7.57)***
(2) AGE (years)	0.03 (2.91)*	0.03 (4.35)**	0.02 (1.66)
(3) RESPONSE SCALE (low-freq) a)	-1.33 (7.80)***	-1.55 (9.11)***	-1.67 (10.14)***
(4) RESPONSE ACCESSIBILITY (scale values)	--	1.14 (4.93)**	-1.38 (6.84)***
(5) RESPONSE SCALE ACCESSIBILITY	--	-1.20 (4.25)**	-1.25 (4.28)**
Constant	2.36 (1.59)	1.25 (0.39)	4.05 (3.51)*
Pseudo r² reduction when removing 'RESPONSE SCALE • ACCESSIBILITY'	--	0.046	0.046
Total pseudo r²	0.154	0.204	0.221
N	110	110	110

a) Reference category: 'high-frequency'; Significance: *** $p \leq 0.01$; ** $p \leq 0.05$; * $p \leq 0.1$

Regression model 2, presented in Table 4, tests whether subjects' susceptibility to effects of different response scales differs according to their self-reported response certainty. This is confirmed, since the interaction parameter between the type of response scale and the certainty measure is found to be a statistically significant predictor for the behavioral reports about the intensity of TV consumption. In model 3, the relevance of response latencies in this respect is analyzed. Here, response times prove to be a valid predictor of the susceptibility to response effects as well: the response latency times response scale interaction explains the response behavior in a significant way. According to the reduction of predictive power for response behavior, when the interaction between the scale type and response certainties as well as response latencies is removed from the regression equation, both accessibility indicators have exactly the same predictive power for scale effects. In both cases, the explained deviance, as measured with pseudo r², is reduced by 4.6 percent points.

---

of these factors proved to be related to the response behavior and are therefore not included in this and the following analysis.

7 The estimation of multiplicative parameters in regression analysis is likely to cause high multicollinearity. Without addressing this problem, the tolerance of interaction parameters in our analysis is in some cases as low as 0.03. In order to reduce multicollinearity, it is recommended to include the respective variables in z-standardized form into regression models (Cronbach, 1987). Although this affects regression parameters on the lower level of hierarchical models, this treatment leaves interaction parameters on the highest level unaffected (Aiken and West, 1991: 28ff.). After the metric measures of information accessibility in our analysis were standardized, the tolerance of all parameters in all regression models is found to be 0.60 or higher.

Figure 1 illustrates the significant interaction effect between the type of response scale and subjects' response certainty. According to the results, those respondents who feel confident about the reported length of their TV consumption are found to be practically unaffected by the type of response scale used to record their answers. In this group 29.6 percent of subjects reported watching more than 2.5 hours of TV a day with the low-frequency scale and 26.5 percent did so when a high frequency scale was used. In contrast, answers of subjects with low response certainty differ strongly depending on the type of response scale. Here, with the low-frequency scale, 10.3 percent reported watching more than 2.5 hours TV per day, whereas this proportion is 70.0 percent with the high-frequency scale. Accordingly, self-reported response certainty is a strong predictor for subjects' susceptibility to the analyzed type of response effects.
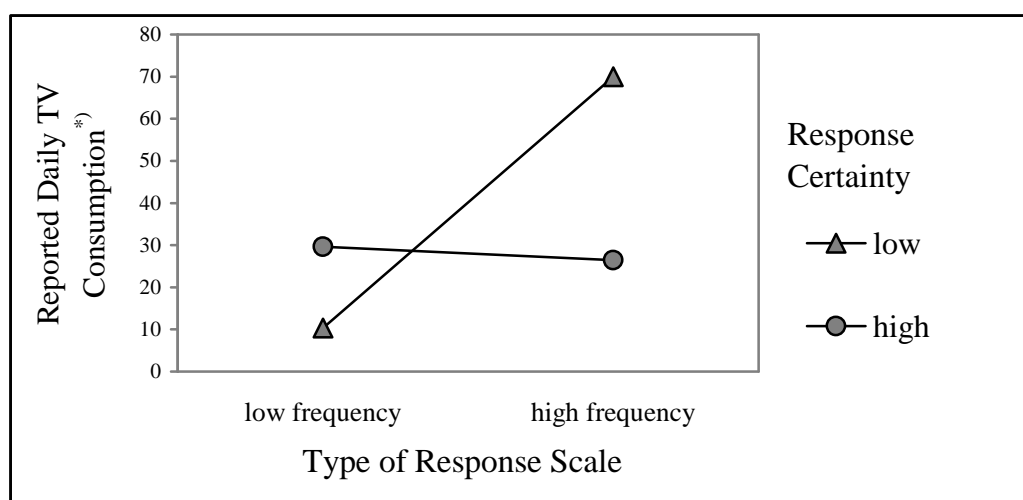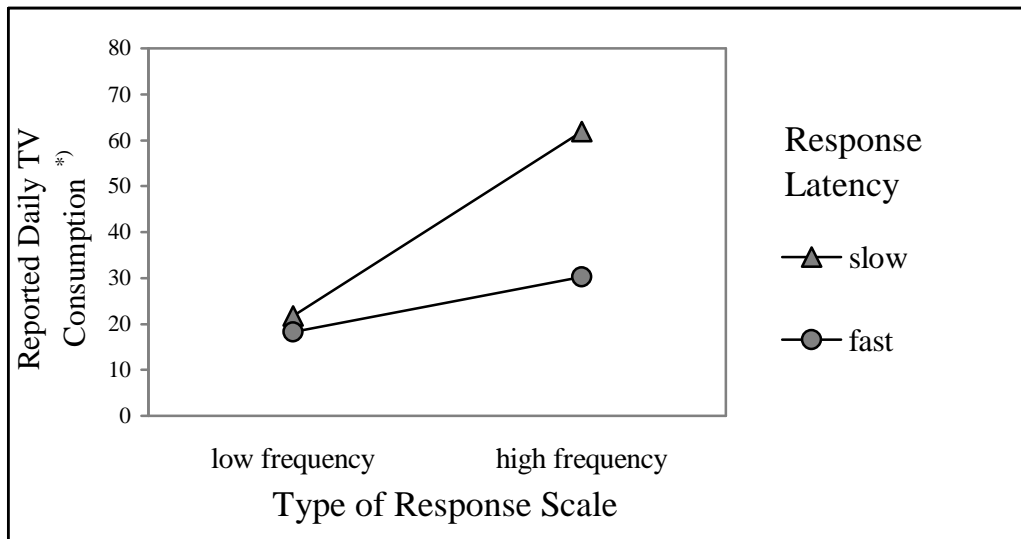

*) Percentage of respondents who reported watching more than 2.5 hours of TV every day

Figure 1: Predictive power of subjects' response certainties for effects of different types of response scales on their reports about the duration of daily TV consumption[8].

Figure 2 presents the significant interaction effect between the type of response scale and the time necessary to answer the question about the extent of TV consumption. The results closely correspond to those found for the moderating role of response certainties. Here, subjects' behavioral reports are only slightly affected by the presentation of response options when these responses were made fast. Under these conditions, 21.7 percent of subjects reported watching more than 2.5 hours of TV every day with the low-frequency scale and 30.3 percent, when the high-frequency scale was used. Subjects' answers with longer response times were found to be much more strongly affected. Here, the low-frequency scale leads to 18.2 and the high-frequency scale to 61.9 percent in reports of heavy TV

---

[8] In this figure, the sample has been split into two groups with different response certainty. The groups with low certainty represents 44.5 and with high certainty 55.5 percent of the respondents.

consumption. In summary, response latencies as well as response certainties can be regarded as valid indicators of the cognitive accessibility of information and good predictors for the analyzed type of response effects.

*) Percentage of respondents who reported watching more than 2.5 hours of TV every day

Figure 2: Predictive power of subjects' response latencies for effects of different types of response scales on their reports about the duration of daily TV consumption[9].

6.4.3 Effect of transformation on the predictive power for response effects

In the following part of the analysis, we test whether the transformation of response certainties and response latencies affects the ability to predict how strongly subjects are influenced by the way response options are presented. This is done with a total of six multivariate logistic regression analyses, in which each combination of accessibility indicator and transformation type is tested with respect to its predictive power. For each of these analyses, it is computed how much explained variance in the response behavior can be added, when the respective 'scales type' times 'accessibility indicator' interaction is introduced into the regression model. Figure 3 presents these increases in pseudo r^2 for different transformations of response certainties and response latencies.

Taking raw response latencies as a starting point, logarithmic and square root transformations are found to be mostly irrelevant for how effectively this measure predicts subjects' susceptibility to response effects. When raw response latencies are used to predict the effect of different response scales, this increases the pseudo

---

[9] The sample has been split into equally-sized subgroups with fast and slow response latencies. The group with fast response times amount to 50.9 and the group with slow response latencies 49.1 percent of the respondents.

r^2 of the regression model by 4.6 percentage points (see Table 4) and this figure is 4.7 in the case of log as well as square root transformed response latencies. In contrast, the reciprocal transformation clearly has a positive effect on the predictive power of this accessibility measure: with a reciprocal transformation, the inclusion of response latencies and their predictive power for how strongly subjects are affected by scale types increases the explained variance by 5.5 percent. Compared with raw response latencies, the reciprocally transformed version of this accessibility measure improves the predictive power of the analyzed type of response effect by 0.9 percent points in pseudo r^2.
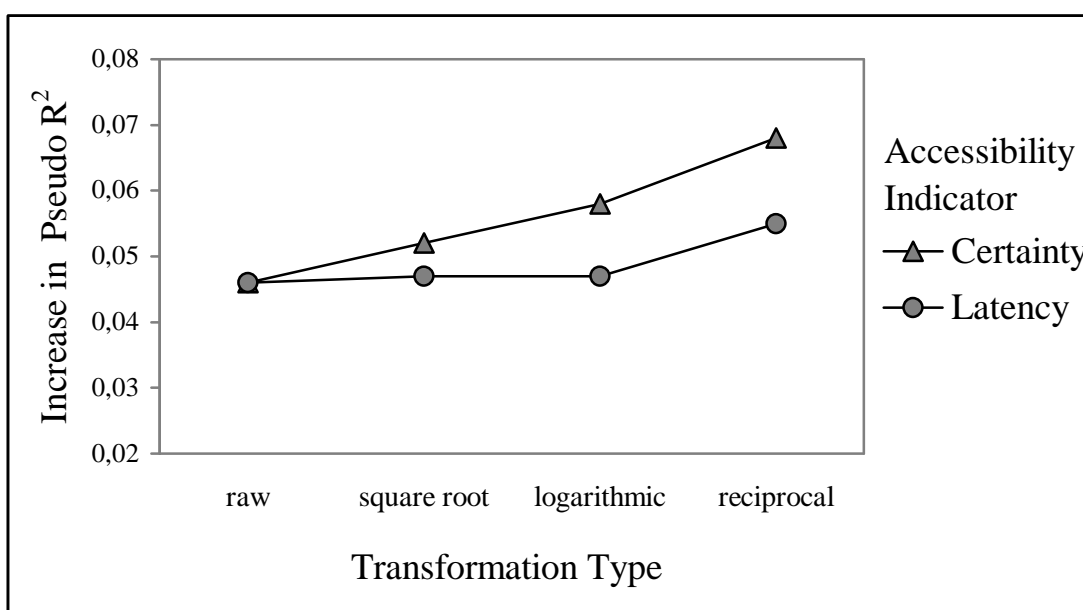

Figure 3: Effect of transformation of response certainties and response latencies on their predictive power for subjects' susceptibility to different types of response scales when reporting their daily TV consumption.

In the case of response certainties, different transformations have a much stronger effect on the predictive power of this accessibility measure. When raw response certainties are included into the regression model in order to predict how strongly individual respondents are influenced by different response scales, the explained variance increases by 4.6 percentage points in terms of pseudo r^2. A log transformation increases this value to 5.2 and a square root transformation to 5.8 percentage points. But, the strongest improvement of predictive power is found in the case of a reciprocal transformation. Using reciprocal transformed response certainties to predict the susceptibility to the analyzed type of response effects increases the explained variance of the regression model by 6.8 percentage points. Compared with raw response certainties, this is an improvement in the ability to

predict how strongly subjects' answers are affected by the way response options are presented by 2.2 percentage points.

7   Summary and discussion

In the first step, our study replicates the well-documented effects of differently categorized response scales on the respondents' answers about the length of their daily TV consumption. According to the results, 19.6 percent of the subjects reported watching more than 2.5 hours of TV every day when a low-frequency response scale is used to record the answers. This proportion is 42.6 percent when a high-frequency scale is used. This difference in response behavior of 23 percentage points is nearly identical to the 21.3 percentage points observed in the first study with the same experimental design (Schwarz et al., 1985). The way response options are presented therefore has a robust effect on the subjects' behavioral reports.

The *first* stage of our study focused on whether individual differences in the cognitive accessibility of the requested information can be used to judge how strongly respondents' answers are affected by the presentation of response options. In this part of the analysis, the predictive power of response certainties and response latencies is compared. According to our results, both measurements are significant predictors of how strongly respondents are affected by the presentation of the response scales: the more certain subjects were about the correctness of their behavioral reports and the less time they needed to answer the frequency question, the weaker are the effects of the scale type on the response behavior. Furthermore, the predictive power of both accessibility measures is found to be equally strong. Our results provide additional and more direct support for the assumed role of information accessibility as a pivotal determinant of response effects (Schwarz and Hippler, 1987).

The *second* stage of our study tested whether a logarithmic, square root and reciprocal transformation of response certainties and response latencies affected their validity as measurements of the cognitive accessibility of information. Such transformations are frequently used in order to correct the characteristically positively skewed distribution of latency data and to attain a better approximation to a normal distribution. Since self-reported response certainties are often also skewed, the same argument applies in the case of this measure as well. However, according to our results only the square root transformation of response latencies has the assumed positive effect on data distribution. All other transformations, in particular in the case of response certainties, are either inconsequential or have a negative effect on the skewed nature of the data distributions. Nevertheless, a comparison of the two raw versions, with the in-total six transformed types of accessibility measures partly reveals clear differences in their predictive validity.

In the case of response latencies, a log and square root transformation has no impact on the ability of this measure to predict response effects: the explanatory power of these versions is practically identical with that of raw response latencies. However, a reciprocal transformation clearly has a positive effect and increases the explained variance, compared to raw response latencies by 0.9 percentage points and therefore the predictive power for response effects by 20 percent. In the case of response certainties, all transformations improve the validity of the raw measure. Here, the logarithmic and square root transformations increase the predictive power by 0.6 and 1.2 percentage points. This is an improvement in the ability to predict the susceptibility to scale effects of between 13 and 26 percent of the predictive power of raw response certainties. However, the reciprocal transformed certainty measure is by far the best measure: the predictive power of raw certainties is improved by 2.2 percentage points and therefore the ability to predict subjects' susceptibility to scale effects by 48 percent.

In summary, both, response certainties and response latencies, are found to be valid indicators for individual differences in the availability of information and, with respect to their raw versions, equally good predictors for response effects. Furthermore, applying a reciprocal transformation increases the predictive power of both indicators. However, since this transformation has a more positive effect on response certainties, this indicator slightly outperforms the response latencies' ability to predict response effects. This disadvantage of response latencies is probably offset by the more cost-efficient way of collecting this data. In principle, response times are a byproduct of computer assisted survey interviews and can easily be recorded. In contrast, the extra questions necessary for collecting response certainties doubles the interview time and therefore the financial cost, as well as placing an extra burden on respondents, because they have to answer additional questions.

According to our results, response certainties *and* response latencies can be used in order to predict how strongly respondents are susceptible to how response scales are presented. Survey researchers can thus utilize these measures in order to judge the reliability and quality of survey data obtained from different groups of respondents. However, further research is needed on whether these accessibility indicators are valid predictors for other kinds of response effects.

Acknowledgements

Sonderforschungsbereich 504 'Rationality Concepts, Decision Behavior and Economic Modeling' at the University of Mannheim.

References

[1]   Aiken, L.S. and West, S.G. (1991): *Multiple Regression: Testing and Interpreting Interactions*. Newbury Park: Sage.

[2]   Bargh, J.A. and Chartrand, T.L. (2000): The mind in the middle. A practical guide to priming and automaticity research. In H.T. Reis and C.M. Judd (Eds.): *Handbook of Research Methods in Social and Personality Psychology*. Cambridge, UK: Cambridge University Press.

[3]   Bassili, J.N. (1993): Response latency versus certainty as indexes of the strength of voting intentions in a CATI survey. *Public Opinion Quarterly*, **57,** 54-61.

[4]   Bassili, J.N. (1995): Response latency and the accessibility of voting intentions: What contributes to accessibility and how it affects vote choice. *Personality and Social Psychology Bulletin*, **21,** 686-695.

[5]   Bassili, J.N. (1996a): The how and why of response latency measurement in telephone surveys. In N. Schwarz and S. Sudman (Eds.): *Answering Questions. Methodology for Determining Cognitive and Communicative Processes in Survey Research.* San Francisco: Jossey-Bass.

[6]   Bassili, J.N. (1996b): Meta-judgemental versus operative indexes of psychological attributes: The case of measures of attitude strength. *Journal of Personality and Social Psychology*, **71,** 637-653.

[7]   Bassili, J.N. and Bors, D.A. (1997): Using response-latency to increase lead time in election forecasting. *Canadian Journal of Behavioural Science*, **29,** 231-238.

[8]   Bassili, J.N. and Fletcher, J.F. (1991): Response-time measurement in survey research. A method for CATI and a New Look at nonattitudes. *Public Opinion Quarterly*, **55,** 331-346.

[9]   Bassili, J.N. and Krosnick, J.A. (2000): Do strength-related attitude properties determine susceptibility to response effects? New evidence from response latency, attitude extremity, and aggregate indices. *Political Psychology*, **21,** 107-132.

[10] Bauman, K.E. and Dent, C.W. (1982): Influence of an objective measure on self-reports of behavior. *Journal of Applied Psychology*, **67,** 623-628.

[11] Bless, H., Bohner, G., Hild, T., and Schwarz, N. (1992): Asking difficult questions: Task complexity increases the impact of response alternatives. *European Journal of Social Psychology* , **22,** 309-312.

[12] Cronbach, L.J. (1987): Statistical tests for moderator variables: Flaws in analyses recently proposed. *Psychological Bulletin*, **102,** 414-417.

[13] Fazio, R.H. (1990): A practical guide to the use of response latency in social psychological research. In C. Hendrick and M.S. Clark (Eds.): *Review of Personality and Social Psychology: Vol. 11. Research Methods in Personality and Social Psychology.* Newbury Park, CA: Sage.

[14] Fazio, R.H. and Williams, C.J. (1986): Attitude accessibility as a moderator of the attitude-perception and attitude-behavior relations: An investigation of the 1984 presidential election. *Journal of Personality and Social Psychology*, **51,** 505-514.

[15] Fletcher, J.F. (2000): Two timing: Politics and response latencies in a bilingual survey. *Political Psychology*, **21,** 27-55.

[16] Gaskell, G.D., O' Muircheartaigh, C.A., and Wright, D.B. (1994): Survey questions about the frequency of vaguely defined events. *Public Opinion Quarterly*, **58,** 241-254.

[17] Ji, L.J., Schwarz, N., and Nisbett, R.E. (2000): Culture, autobiographical memory, and behavioral frequency reports: Measurement issues in cross-cultural studies. *Personality and Social Psychology Bulletin*, **26,** 585-593.

[18] Kokkinaki, F. and Lunt, P. (1997): The relationship between involvement, attitude accessibility and attitude-behaviour consistency. *British Journal of Social Psychology*, **36,** 497-509.

[19] Krosnick, J.A. and Petty, R.E. (1995): Attitude strength: An overview. In R.E. Petty and J.A. Krosnick (Eds.): *Attitude Strength. Antecedents and Consequences.* Mahwah, N.J.: Erlbaum.

[20] Krosnick, J.A. and Schuman, H. (1988): Attitude intensity, importance, and certainty and susceptibility to response effects. *Journal of Personality and Social Psychology*, **54,** 940-952.

[21] LaBarbera, P.A. and MacLachlan, J.M. (1979): Response latency in telephone interviews. *Journal of Advertising Research*, **3,** 49-55.

[22] Lavine, H., Huff, J.W., Wagner, S.H., and Sweeney, D. (1998): The moderating influence of attitude strength on susceptibility to context effects in attitude surveys. *Journal of Personality and Social Psychology*, **75,** 359-373.

[23] Menon, G., Raghubir, P., and Schwarz, N. (1995): Behavioral frequency judgements: An accessibility-diagnosticity framework. *Journal of Consumer Research*, **22,** 212-228.

[24] Menon, G., Raghubir, P., and Schwarz, N. (1997): How much will I spend? Factors affecting consumer's estimates of future expense. *Journal of Consumer Psychology*, **6,** 141-164.

[25] Mielke, R. (1985): Eine Untersuchung zum Umweltschutz-Verhalten (Wegwerf-Verhalten): Einstellung, Einstellungs- Verfügbarkeit und Soziale

Normen als Verhaltensprädiktoren. *Zeitschrift für Sozialpsychologie*, **16,** 196-205.

[26] Mulligan, K., Grant, J.T., Mockabee, S.T., and Monson, J.Q. (2003): Response latency methodology for survey research: Measurement and modeling strategies. *Political Analysis*, **11,** 289-301.

[27] Renata, F. (1999): Attitude strength and the attitude-behavior domain: Magnitude and independence of moderating effects of different strength indices. *Journal of Social Behavior & Personality*, **14,** 177-197.

[28] Rockwood, T.H., Sangster, R.L., and Dillman, D.A. (1997): The effect of response categories on questionnaire answers. *Sociological Methods & Research*, **26,** 118-140.

[29] Schuman, H., Presser, S., and Ludwig, J. (1981): Context effects on survey responses to questions about abortion. *Public Opinion Quarterly*, **45,** 216-223.

[30] Schwarz, N. (1988): Was Befragte aus Antwortvorgaben lernen: Zur informativen Funktion von Antwortvorgaben bei Verhaltensfragen. *Planung und Analyse*, **15,** 103-107.

[31] Schwarz, N. (1999): Frequency reports of physical symptoms and health behaviors: How the questionnaire determines the results. In D.C. Park, R.W. Morell, and K. Shifren (Eds.): *Processing in Medical Information in Aging Patients: Cognitive and Human Factors Perspectives* . Mahwah, N.J.: Erlbaum.

[32] Schwarz, N. and Bienias, J. (1990): What mediates the impact of response alternatives on frequency reports of mundane behaviors. *Applied Cognitive Psychology*, **4,** 61-72.

[33] Schwarz, N., Grayson, C.E., and Knäuper, B. (1998): Formal features of rating scales and the interpretation of question meaning. *International Journal of Public Opinion Research*, **10,** 177-183.

[34] Schwarz, N. and Hippler, H.J. (1987): What response scales may tell your respondents: Informative functions of response alternatives. In H. J. Hippler, N. Schwarz, and S. Sudman (Eds.): *Social Information Processing and Survey Methodology.* New York: Springer.

[35] Schwarz, N., Hippler, H.J., Deutsch, B., and Strack, F. (1985): Response scales: Effects of category range on reported behavior and comparative judgments. *Public Opinion Quarterly*, **49,** 388-395.

[36] Schwarz, N. and Scheuring, B. (1988): Judgements of relationship satisfaction: Inter- and intraindividual comparisons as a function of questionnaire structure. *European Journal of Social Psychology*, **18,** 485-496.

[37] Stocké, V. (2001): The influence of frequency scales on the response behavior. A theoretical model and its empirical examination. Working-Paper No. 01-24: Sonderforschungsbereich 504, University of Mannheim.

[38] Stocké, V. (2002a): Die Vorhersage von Fragenreihenfolgeeffekten durch Antwortlatenzen: Eine Validierungsstudie. *ZUMA-Nachrichten*, **50,** 26-53.

[39] Stocké, V. (2002b): Soziale Erwünschtheit bei der Erfassung von Einstellungen gegenüber Ausländern. Theoretische Prognosen und deren empirische Überprüfung. Arbeitspapier No. 02-09: Sonderforschungsbereich 504.

[40] Warland, R.H. and Sample, J. (1973): Response certainty as a moderator variable in attitude measurement. *Rural Sociology*, **38,** 174-186.

[41] Wegener, D.T., Downing, J., Krosnick, J.A., and Petty, R.E. (1995): Measures and manipulations of strength-related properties of attitudes: Current practice and future directions. In R.E. Petty and J.A. Krosnick (Eds.): *Attitude Strength. Antecedents and Consequences.* Mahwah, N.J.: Erlbaum.

[42] Winter, J. (2002): Bracketing effects in categorized survey questions and the measurement of economic quantities. Working-Paper No. 02-35: Sonderforschungsbereich 504, University of Mannheim.