

## Short Note

### **Asking Difficult Questions: Task Complexity Increases the Impact of Response Alternatives**

HERBERT BLESS, GERD BOHNER and  
TRAUDEL HILD  
*Universität Mannheim,  
Mannheim, FRG*

*and*

NORBERT SCHWARZ  
*Zentrum für Umfragen und Analysen, ZUMA,  
Mannheim, FRG*

#### *Abstract*

*In providing behavioural frequency reports, respondents use the range of the response alternatives as a frame of reference, resulting in higher estimates on scales that offer high rather than low values. The present study demonstrates that the size of this effect increases with increasing question difficulty.*

#### **INTRODUCTION**

In psychological testing, laboratory experiments, and survey research, respondents are often asked to report the frequency with which they engage in a certain behaviour by checking the appropriate alternative from a set of response categories provided to them. Recent research indicated that respondents simplify their task by using the range of the response alternatives as a frame of reference in computing a frequency estimate (see Schwarz (1990) and Schwarz and Hippler (in press) for reviews). The use of this estimation procedure is based on the assumption that the scale reflects the researcher's knowledge about the frequency distribution, and that values in the middle range of the scale reflect the 'average' or 'usual' frequencies, whereas the extreme values of the scale correspond to the extremes of the distribution. As a result, respondents report higher behavioural frequencies when the response scale

Mailing addressee for correspondence: Herbert Bless, FP Subjektive Erfahrung, Universität Mannheim, Steubenstr. 46, D-6800 Mannheim, Germany.

offers high rather than low frequency response alternatives (e.g. Schwarz, Hippler, Deutsch and Strack, 1985; Schwarz and Scheuring, 1988).

Theoretically, the use of estimation procedures should be more likely, the less that respondents can rely on a 'recall and count' strategy (Bradburn, Rips and Shevell, 1987). Accordingly, the impact of response alternatives has been found to be more pronounced for frequent and mundane behaviours, which are not well represented in memory (Schwarz, 1990), than for rare and important events, and for reports about others' rather than one's own behaviour (Schwarz and Bienias, 1990). Whereas this research indicated that reliance on estimation strategies increases as it becomes more difficult to derive a memory-based answer, the present research extends this work by exploring a different aspect of task difficulty. Specifically, we asked respondents to provide either reports of the *absolute* or the *relative* frequency with which they engage in five different behaviours. For example, a respondent would either be asked, how many hours he or she spends watching TV (absolute frequency) or what percentage of his or her leisure time is spent watching TV (relative frequency). Whereas the memorability of watching TV is presumably the same under both conditions, reports of relative frequencies pose a considerably more complex task by requiring an estimate of one's total leisure time, an estimate of one's TV consumption, and the computation of the respective proportion. We assume that respondents will simplify this task by relying on the range of the response alternatives in computing their estimate. If so, the impact of response alternatives should be more pronounced for relative rather than absolute frequency reports of the same behaviour. This finding would extend previous research by indicating that increasing task difficulty elicits an increased reliance on estimation strategies, independent of the memorability of the respective behaviour.

## METHOD

One hundred and thirty-three subjects were randomly assigned to a 2 (low versus high frequency scale)  $\times$  2 (absolute versus relative frequency report) factorial design. Subjects were provided with a questionnaire including five critical and two filler items. The critical items referred to different activities (watching TV, buying educational materials or clothes, cultural activities, drinking coffee). Subjects were asked to report either the absolute or the relative frequency with which they engaged in these behaviours, on a response scale that provided either high or low frequency response alternatives. Pretesting had demonstrated the relative frequency questions to be significantly more difficult to answer than the absolute frequency questions. For TV consumption, these questions would read, for example:

Absolute frequency:

'How many hours do you watch TV on a typical weekday?

- |                 |                  |
|-----------------|------------------|
| low frequency   | high frequency   |
| — up to ½ hour  |                  |
| — ½ to 1 hour   |                  |
| — 1 to 1½ hours |                  |
| — 1½ to 2 hours |                  |
| — 2 to 2½ hours | — up to 2½ hours |
-

- |   |   |
|---|---|
| <input type="checkbox"/> more than 2½ hours | <input type="checkbox"/> 2½ to 3 hours      |
|   | <input type="checkbox"/> 3 to 3½ hours      |
|   | <input type="checkbox"/> 3½ to 4 hours      |
|   | <input type="checkbox"/> 4 to 4½ hours      |
|   | <input type="checkbox"/> more than 4½ hours |

Relative frequency:

‘How much of your leisure time do you spend watching TV on a typical weekday?’

- | low frequency                                  | high frequency                                 |
|--|--|
| <input type="checkbox"/> up to 5 per cent      |  |
| <input type="checkbox"/> 5 to 10 per cent      |  |
| <input type="checkbox"/> 10 to 15 per cent     |  |
| <input type="checkbox"/> 15 to 20 per cent     |  |
| <input type="checkbox"/> 20 to 25 per cent     | <input type="checkbox"/> up to 25 per cent     |
| <hr/>  |  |
| <input type="checkbox"/> more than 25 per cent | <input type="checkbox"/> 25 to 35 per cent     |
|  | <input type="checkbox"/> 35 to 45 per cent     |
|  | <input type="checkbox"/> 45 to 55 per cent     |
|  | <input type="checkbox"/> 55 to 65 per cent     |
|  | <input type="checkbox"/> more than 65 per cent |

Reports provided on the different scales are comparable by computing the proportion of subjects who report behavioural frequencies above or below the cut off point<sup>1</sup> of the respective scale (indicated by a line in the above example). Following suggestions by Rosenthal and Rosnow (1985) for the analysis of proportions, respondents above the cut-off point were assigned a value of 1, and respondents below the cut-off point a value of 0, and these data were analysed by a 2 (scale) × 2 (type of judgment) MANOVA.<sup>2</sup>

## RESULTS AND DISCUSSION

Consistent with previous results, subjects reported higher frequencies when provided with high rather than low frequency response alternatives,  $F(5,125) = 17.01$ ,  $p < 0.0005$ , as shown in Table 1.

Although the impact of response alternatives was reliable for both types of question, the effect was more pronounced for reports of relative,  $F(5,125) = 20.96$ ,  $p < 0.0005$  for the simple main effect, rather than absolute frequencies,  $F(5,125) = 2.07$ ,  $p < 0.074$  for the simple main effect. This pattern of findings is reflected in the predicted interaction of scale range and type of judgment,  $F(5,125) = 5.58$ ,  $p < 0.005$ .

These findings indicate that respondents' reliance on estimation strategies increases with increasing task difficulty. Whereas task difficulty was a function of the availability of relevant episodic memories in previous research (e.g. Schwarz and Bienias, 1990), the present findings extend this research by suggesting that any other variable that

<sup>1</sup> The cut-off points represent the modal response for the respective behaviour obtained in pretests using an open answer format.

<sup>2</sup> All  $F$ -ratios are based on Wilks's lambda.

Table 1. Percentages of subjects above the cut-off point of the respective scale

Scale range difference	Type of question					
	Absolute frequency		Difference	Relative frequency		
	High	Low		High	Low	
Topic						
watching TV	10	09	01	12	00	12
buying educational material	81	62	19	59	06	53
cultural activities	39	15	24	65	09	55
drinking coffee	32	29	03	85	29	56
buying clothes	97	76	21	88	24	64

Percentages of subjects above the cut-off point of the respective scale. The higher the difference in percentages the higher the frequency reports on the high relative to the low frequency scale.

increases task difficulty may also increase respondents' reliance on the range of response alternatives presented to them. The more demanding the computation of a frequency report is, the more likely respondents are to use the range of the response alternatives as a frame of reference, resulting in reports that are largely a function of the response alternatives offered to them. Accordingly, researchers would be well advised to assess frequency reports in an open response format (*cf.* Schwarz, 1990), and the more so, the more complex the task is.

### ACKNOWLEDGEMENTS

The reported research was supported by grants from the Deutsche Forschungsgemeinschaft (Str. 264/2 to F. Strack and N. Schwarz, and 278/5 to N. Schwarz, H. Bless and G. Bohner), and the Bundesminister für Forschung und Technologie of the Federal Republic of Germany (SWF0044 to N. Schwarz).

### REFERENCES

- Bradburn, N., Rips, L. J. and Shevell, S. K. (1987). 'Answering autobiographical questions: The impact of memory and inference on surveys', *Science*, **236**: 157-161.
- Rosenthal, R. and Rosnow, R. L. (1985). *Contrast Analysis*, Cambridge University Press, Cambridge.
- Schwarz, N. (1990). 'Assessing frequency reports of mundane behaviors: contributions of cognitive psychology to questionnaire construction'. In: Hendrick, C. and Clark, M. S. (Eds) *Research Methods in Personality and Social Psychology Review of Personality and Social Psychology*, Vol. 11, Sage, Beverley Hills, CA, pp. 98-119.
- Schwarz, N. and Bienias, J. (1990). 'What mediates the impact of response alternatives on frequency reports of mundane behaviours?' *Applied Cognitive Psychology*, **4**: 61-72.
- Schwarz, N. and Hippler, H. J. (In press). 'Response alternatives: The impact of their choice and ordering'. In: Biemer, P. *et al.* (Eds) *Measurement Error in Surveys*, Wiley, Chichester, pp. 00-00.
- Schwarz, N., Hippler, H. J., Deutsch, B. and Strack, F. (1985). 'Response categories: effects on behavioral reports and comparative judgments', *Public Opinion Quarterly*, **49**: 388-395.
- Schwarz, N. and Scheuring, B. (1988). 'Judgments of relationship satisfaction: inter- and intraindividual comparisons as a function of questionnaire structure', *European Journal of Social Psychology*, **18**: 485-496.