Asking for Numbers and Quantities: Visual Design Effects in Web Surveys

Marek Fuchs
Social Science Department, University of Kassel, Nora-Platiel-Str. 1, Kassel, Hessen, 34109, Germany

Abstract

When collecting frequencies and quantities in self-administered surveys using open-ended questions researchers face a serious problem: Depending on the concept being measured respondents may not have the exact number readily available to report. Thus they make use of estimation strategies in order to provide a reasonable response. When it comes to the reporting stage of the question answer process, open-ended frequency questions are prone to a specific respondent behavior: Respondents do not report a pure number using digits only, instead they extend their response to alphanumeric information and explicit estimates (e.g., "about 10-20 cigarettes per day") which requires a lot of editing and cleaning in the post-field phase of a survey. This paper aims to provide some insights regarding the optimal visual design of open-ended frequency questions in Web surveys in order to reduce the percentage of alphanumeric elaborations and explicit estimates to frequency questions.

KEY WORDS: Web survey; frequency question, visual design

1. Background and Research Question

In almost every self-administered survey respondents are asked to report frequencies or quantities of behaviors or events in a given time frame or reference period. Respondents answer this type of question by way of two response strategies: They either "recall and count" the instances of the desired behavior or event, or they "estimate" the required number using rate-based computation, pure guessing or by combinations of these heuristics (Sudman et al. 1996). Generally speaking, data based on recall and count is considered to be of superior quality compared to estimated values.

Whether a response is generated by recalling and counting the behaviors or events in question or whether the answer is estimated depends on various characteristics of the concept being asked (Sudman et al. 1996). First, the regularity, similarity and frequency of the occurrence of the behaviors or events determine to what extent respondents recall and count. Also, the length of the reference period has a significant impact on the response quality (Schneider and Sumi 1981). In addition to characteristics of the concept being measured and to the attributes of the reference period, the respondents' motivation affects the extent to which they make use of either estimation strategies or recall and count, as well (Burton and Blair 1991).

In contrast to this exiting body of research we will focus on visual design aspects of the frequency questions. We will assess the effect of formal characteristics of the response field on the response strategies applied by the respondents and on the characteristics of the responses provided.

1.1 Properties of Open-Ended Frequency Questions

Questionnaire designer make use of two different techniques when measuring frequencies or quantities. They either use open-ended questions which provide an open response field, or they rely on closed-ended questions with response options representing frequency ranges. The two question types show differences in the structure of the question answer process. With closed-ended questions respondents rely on the range represented by the response options while generating the answer (Schwarz and Hippler 1987). By contrast, open-ended frequency questions do not support the question answer process in this way, hence they are considered to be more burdensome for the respondents to answer. However, at the same time when using open-ended questions responses are less prone to scale effects (Schwarz and Hippler 1985; Wänke 2002). Also, open-ended questions provide metric data – at least for some variables – while closed-ended question yield ordinal scale level most of the time. Because of these two factors, in many surveys open-ended questions are preferred when collecting frequencies or quantities.

However, the pure numbers collected by open-ended questions suggest a precision and exactness that they often do not provide. With respect to the response strategies discussed above – recalling and counting the instances of the behavior or event vs. estimating – one should be aware of the fact that responses to open-ended frequency questions are also prone to estimation strategies. When answering open-ended frequency questions respondents do not only provided numbers (digits) but at the same time they often add alphanumeric input (e.g. "none", "about", other text entries elaborating the response), other symbols (e.g. "~", "$", "€") or just junk ("3.54654 million doctor visits"). Thus, the responses to

open-ended frequency questions need a lot of cleaning and editing in order to come up with a reasonable response distribution.

Furthermore, many respondents provide ranges ("4-6", "5-10") as well as other indications of explicit estimates ("ca. 10", "~25"). These explicit estimates denote that respondents have problems to come up with an exact number und thus provide a range in which their true response would most likely fall.

In addition to those explicit estimation strategies (ranges, "ca.", "~", "about"), survey researchers know another phenomenon that hints towards estimated responses: heaping and bunching (e. g. Jobe et al. 1997). Looking at the response distribution of any give frequency question one can easily detect the relatively high frequencies for the flat figures (e. g. 10, 20, 30, 50 – depending on the range of the frequency distribution) as compared to the uneven numbers in between. This phenomenon is attributed to respondents who were required to estimate the answer because they cannot determine the true number by recall and count. Thus, they rely on an estimation strategy which they express by providing a flat number as an indication of the magnitude of the true response.[1]

In sum, responses to open-ended frequency questions in self-administered surveys are prone to implicit (heaping and bunching) and explicit (range, "about", "ca.", "~") estimates and also to alphanumeric elaborations. While explicit estimates and other alphanumeric input might be advantageous because the fact that the response was obtained using estimation strategies is still visible, they require far more labor on the editing stage of a survey compared to implicit estimates.

1.2 Graphical Characteristics of Response Fields to Open-Ended Questions

In this paper we will especially focus on the impact of the graphical characteristics of the response field on the proportion of implicit estimates, explicit estimates, and alphanumeric elaborations provided by the respondent. Previous research on self administered surveys could show that graphical elements of a survey question may influence the cognitive processing of the question and thus the responses obtained from respondents. Evidence from this line of research in the framework of visual design effects (e.g. Christian/Dillmann 2004; Christian et al. 2007; Couper et al. 2004; Tourangeau et al. 2000) indicates that formal aspects of questions could have an impact on the question answer process that is similar to the effect of the wording of the question, the number and order of response categories and other elements, that are traditionally be considered content bearing.

So far, only a few experiments are reported in the literature, some of which take Sanches (1992) and Smith (1995) as a point of origin. Tourangeau and colleagues (2004) studied the effect of lines, symbols and graphical elements on question understanding as well as on the formatting stage of the question answer process. They found evidence that respondents make inferences about the meaning of survey items based on visual clues such as the spacing of the response options and the graphical grouping of the questions. The authors conclude that based on those visual clues respondents might interpret a survey question in ways that the researcher never intended.

Additional support for the impact of the visual design on the cognitive processing comes from field-experimental studies on graphical and symbolic aspects of self-administered paper-based surveys. Christian and Dillman (2004) found clear evidence supporting the assumption that the size of a response field to an open-ended question affects the amount of information provided by the respondent. "In two of the three questions, the larger answer space significantly produced a greater number of themes and topics mentioned in the answer" (Christian/Dillman 2004: 68). Based on theses findings the authors conclude "that the visual design of questions on self-administered questionnaires can significantly impact the respondent behavior" (Christian/Dillman 2004: 78; similarly Dillman et al. 2005; Smyth et al. 2007).

In this study we focus on the impact of the response field's formal characteristics in a self-administered survey on the degree to which the responses are prone to alphanumeric elaborations, implicit estimates and explicit estimates. This study is part of an ongoing line of research, which is based on the assumption, that formal aspect (=graphical and visual properties) of the response field influence the questions answer process: most dominantly, the reporting stage is affected since respondents react to the formal characteristics of the response field when providing their response. However, other stages of the question answer process might be affected as well.

In previous experiments (Fuchs 2007, see also figure 1 for examples of experimental conditions) we could demonstrate, that the size (length) of a response field to a fre-

[1] Tourangeau and colleagues (2000) call this phenomenon "rounding" which implies that respondents actually know – or could know – the exact response. This true answer would then be rounded for the report. By contrast, we assume that respondents do not know the true answers and have little chance to determine a reasonable frequency using recall and count. Thus, they apply an estimation strategy (Sudman et al. 1996).

quency question has an impact of the proportion of implicit and explicit estimates as well as on the proportion of alphanumeric elaborations (example 1): The longer the field, the more elaboration and explicit estimates were to be observed. Also, a label associated with a response field has an impact on the characteristics of the responses provided (example 2): Generally speaking, labels reduced the proportion of alphanumeric additions and explicit estimates. Finally we could demonstrate that response fields designed as a box with placeholders for each digit yielded fewer explicit estimates and fewer alphanumeric elaborations compared to simple response lines (example 3). These findings proved true especially for those frequency questions that required estimation strategies, because the quantity in question could not be determined by "recall and count".

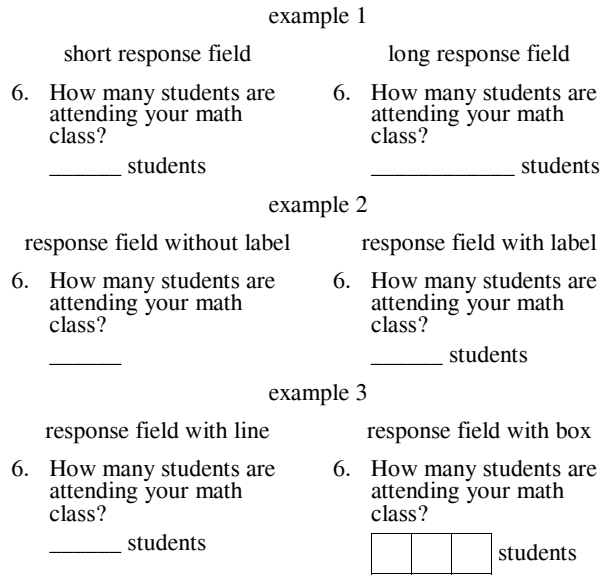

Figure 1: Examples for experimentally varied properties of response fields associated to frequency questions

So far, these findings have been proven for self-administered paper-based questionnaires. In contrary to these findings, for Web surveys the effects of the response fields' size and the associated label are by far less pronounced or not existent at all. We have attributed this inconsistency to the respondent's perception of Web interfaces. In our view, respondents assume that the computer would accept digits only (digits are the "language" of computers), thus, they restrict their responses to an open frequency question in Web survey predominantly to digits. As a consequence, the proportion of explicit estimates and alphanumeric elaborations is rather small in Web surveys. In addition, the percentage of alphanumeric elaborations, explicit estimates, and implicit estimates is not affected by the design properties that were tested so far (size and label), because the perception of the Web interface as a computer program overrides the impact of other formal characteristics of the response field (Fuchs 2007).

In concordance with this assumption, Smyth and colleagues (2007) have demonstrated that formal characteristics of the response field associated with an open text question in a Web survey have a relatively small effect on the responses obtained compared to a paper and pencil surveys. In their study, the effect size in Web surveys is by far smaller and less consistent compared to paper and pencil surveys which shades light on the question to what extent Web surveys and paper and pencil surveys are comparable in their underlying visual design principles.

1.3 Hypothesis

For Web surveys we predict a modification of the graphical elements' impact on the response behavior. In a Web survey as compared to a paper-based self-administered survey the question answer process differs to a certain degree; the questionnaire is answered rather segmented and each question is treated separated from previous questions (Fuchs 2002; 2003a) which in part is due to the interactive character of most current Web surveys. For the research reported in this paper it is important to recognize that responding to a Web survey requires a computer-mediated interaction. The respondent is well aware that he or she interacts with a computer interface and the interface reminds the respondent frequently throughout the survey of its computer mediated character.

While a paper form allows all kinds of deviations when formatting the response, most users are aware that computer interfaces a less flexible and less accepting. While they can take notes and elaborate their responses in a paper-based self-administered questionnaire, most respondents have experienced difficulty doing similar things on a Web page. Thus, they adapt their response behavior to the specifics of the Web survey. The underlying assumption is that Web surveys per se yield far less input beyond pure digits. Thus, in a Web survey respondents are less prone to alphanumeric input and explicit estimates compared to respondents in a similar paper-based self-administered survey.

In particular, a special property of the computer interface is anticipated by the respondents: the type of variables or the format of the desired data respectively. Respondents are assumed to be aware of the fact that a computer interface does not accept characters when it asks for an input that usually requires digits. Thus they adapt their response behavior accordingly. As a consequence, the proportion of alphanumeric input and explicit estimates is reduced compared to a similar frequency question embedded into a paper-based survey.

In this study we will experiment with a formal characteristic of a response field associated with a frequency question in Web survey that strengthens its perception of the computer-interface. In order to do so, we will convey the expected format of the response using a pre-defined value displayed in the response field printed in shaded gray (e. g. "0,000.00"). If the respondent clicks into the response field in order to provide his or her answer, the pre-set value would disappear and give room for the respondent's answer (see figure 2). This will convey the format of the desired response and remind the respondents of properties of the expected answer which in turn will decrease the proportion of alphanumeric input and explicit estimates. By contrast, the percentage of implicit estimates will be larger when a pre-set value is displayed in a Web survey because the pre-set value will affect the reporting stage of the question answer process only, but not prior stages of the cognitive processing, like retrieval of relevant information (including the decision on "recall and count" vs. "estimation"). Since the underlying response strategy (recall and count vs. estimation) will not be affected by the pre-set value, respondents who have estimated their response will denote the character of the response as an estimate using heaping and bunching instead of explicit estimates. Since responses which are based on recall and count are not prone to implicit or explicit estimates on the reporting stage, this effect will be most visible for those questions that require estimation strategies in order to find an answer. Question that can be answered applying "recall and count" by most respondents are not affected by the visual design effects studied in this experiment.

1.4 Methods

400 students of the University of Kassel (Germany) participated in a field experimental comparison of two versions of an open-ended frequency question. The study was embedded in a survey on the consumption of alcohol, nicotine, and other drugs among university students (see Fuchs/Funke 2007 for details). The respondents belong to a student online access panel at the university maintained by the author. All participants were randomly assigned to either of two versions of a questionnaire that contained identical substantive questions. As part of the study the students were supposed to report the amount of money that they have spent on alcoholic beverages within the past four weeks. The question was presumed to require an "estimation" strategy in order to produce a meaningful answer.

Each questionnaire contained either of two experimentally modified question versions: In version 1 (see figure 1, right) respondents answered the question using standard Web survey methodology. Next to the question wording an empty open input field of appropriate size was displayed. In version 2 (figure 1, left) respondents answered the same question, however, the response field was not empty – like in the traditional Web survey approach. Instead, we have pre-filled the response field with a shaded grey "0,000.00" (which would disappear when respondents clicked with the mouse into the field in order to enter their response).

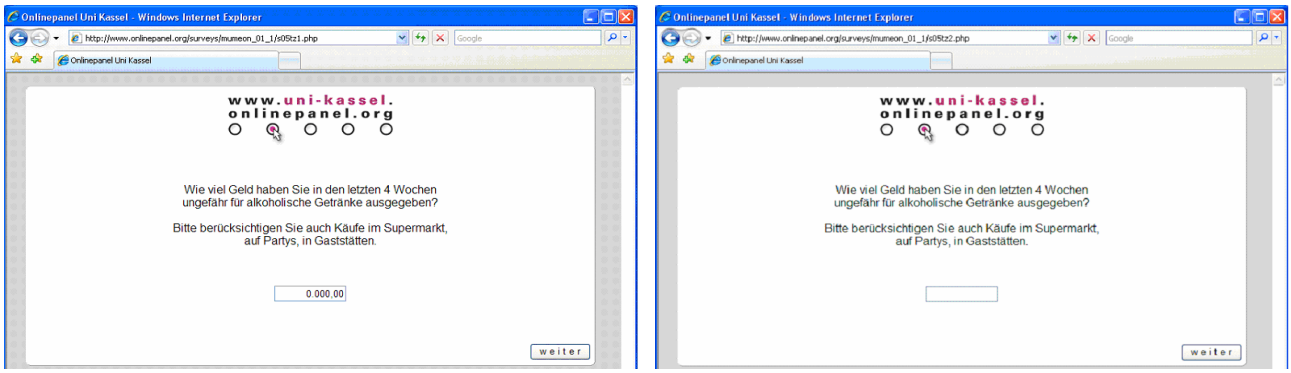

Figure 1: Questionnaire version with (left) and without (right) a pre-defined default value (Fuchs/Funke 2007)

Note: Questions asked in an interactive Web survey (one question at a time). Question wording: "How much money have you spent on alcoholic beverages within the past 4 weeks? Please consider also purchases in grocery stores, at parties, and in bars."

We assumed that this default value would work as a signal indicating the properties of the expected answer. Thus, we reminded respondents to report a pure number without any elaboration or other alphanumeric supplemental information. We assumed that this pre-set default value would increase the respondent's perception of the Web survey as a computer interface, which in turn would affect their response behavior. The Web interface was designed

in a way so that it would accept every kind of input whatsoever. All responses obtained from the respondents were stored in a database and manually coded thereafter: For each response we determined whether or not it contained an alphanumeric elaboration, an explicit estimate, or it represented an implicit estimate (the following answers were coded as an implicit estimate: 5, 10, 15, 20, 30, 50, 100, 200, 300).

2. Results

Results are displayed in table 1. Evidence indicates a clear effect of the experimental condition in the expected direction. Respondents who answered the question with the default value present in the response field are by far more prone to reporting a pure number (digits only). By contrast, in the control group with no default value being visible in the response field, the proportion of respondents providing any alphanumeric input (also in addition to a number) is significantly larger ($p < .001$).
The proportion of explicit estimates is not affected by the presence or absence of the pre-defined value in the response field. However, as has been shown with in the context of previous experiments (Fuchs 2007), the overall proportion of explicit estimates is rather small in Web surveys compared to a paper-based self-administered survey. Also, the proportion of implicit estimates is only faintly affected by the presence of the pre-set value.

Table 1: Percentage of response characteristics of the answer obtained for the conditions with vs. without pre-set default value in the response field

	With default value	Without default value	Total
Pure numbers	34	17	26
Implicit estimates	58	60	59
Explicit estimates	2	1	2
Alphanumeric input	6	22	13
Missing data	0	0	0
Total	100	100	100 ***

Note: N = 399; *** $p < .001$. "How much money have you spent on alcoholic beverages within the past 4 weeks? Please consider also purchases in grocery stores, at parties, and in bars."

Even though the pre-set default value does not change the percentage of explicit and implicit estimates dramatically, it increases the proportion of pure number responses and at the same time it reduces the proportion of responses containing any alphanumeric characters or elaborations. Thus, the pre-set default value used in this Web survey seems to convey an instruction similar to a response box with placeholders for each digit in a paper-based self-administered survey (see figure 1, example 3, and also Fuchs 2007 for details). It signals that nothing but a pure digit response is expected.

3. Discussion

In Web surveys the proportion of alphanumeric input and explicit estimates is by far smaller compared to paper-based self administered surveys. In our view, this is due to the fact that a Web survey is predominantly been perceived by the respondents as a computer interface that will not accept any other input than just pure digits. The perception of the Web survey as a computer interface overrides the possible effects of those visual design elements that were identified as important factors in paper-based self-administered surveys (the length of input field, a label associated with the response field, and the design as a box with placeholders for each digit, see figure 1 and Fuchs 2007 for details). However, we assumed that any hint to the computer-mediated character of the survey would decrease alphanumeric elaborations and explicit estimates while implicit estimates would be increased respectively.

Our findings are partially in accordance with this assumption: in this experiment the proportion of alphanumeric input and (less pronounced) of explicit estimates could be reduced by providing a pre-set default value in the response field that reminds the respondent of the type of the desired response (digits only). We assume that the pre-set value strengthens computer mediated character of the survey and reminds respondents of the desired response format. As a consequence, responses beyond pure digits are less prevalent.

Interestingly, the proportion of implicit estimates (heaping and bunching) is not affected by the presence or absence of a pre-set value. Even though, the proportion of pure digit responses rises, the relative proportion of implicit estimates among those pure digit answers is reduced (overall, the proportion of implicit estimates among all responses remains stable). In our view this finding indicates that the pre-set value does not only change the way the respondents report their answers that are based on an estimation strategy. At the same time the pre-set value seems to increase the proportion of responses that generated based on "recall and count". This gives room for the assumption that the design of the response field changes not only the reporting of the response but improves data quality as well. In turn, we have to reconsider our previous assumption that a pre-set value would affect the reporting stage of the question answer process, only.

The findings reported here (as compared to the results of similar experiments with paper and pencil surveys) suggest that paper-based self-administered surveys and Web

surveys differ in terms of their visual design language. They still belong to the same group of survey modes (self-administration) however, respondents react to varying sets of formal properties and the response behavior might differ. In a paper-based self-administered survey respondents assume a highly resilient questionnaire which allows notes, alphanumeric extension of the response and the like. Thus, when they feel the need to answer the question with elaborations or extensions they do not hesitate to do so. By contrast, in a Web survey the respondents' perception of the questionnaire is dominated by its character as a computer-mediated interaction. Thus, respondents do not choose ways to convey additional information that they would have chosen in a paper-based environment.

These results remind us of the importance of the notion that "little things matter" (Smith 1995). In addition, we have to be aware that those little things matter to various degrees depending on the type of survey mode been used. Even though, a Web survey might provide a lot of empty space on each questionnaire page (especially when using an interactive design), respondents cannot annotate, provide elaborations and the like. Since respondents are used to these limitations of Web pages and Web forms, they adapt their response behavior in Web surveys to the available means of data input.

Since mixed mode studies including a Web component are becoming more and more popular in order to overcome decreasing response rates, it is important to note that identically designed questions on paper and on the Web do not necessarily induce identical cognitive processing on the respondent's side.

References

Burton, S., and Blair, E. (1991), "Task Conditions, Response Formulation Processes, and Response Accuracy for Behavioral Frequency Questions in Surveys," *Public Opinion Quarterly*, **55**, 50-79.

Christian, L. M., and Dillman, D. A. (2004), "The Influence of Graphical and Symbolic Language Manipulations on Responses to Self-Administered Questions," *Public Opinion Quarterly*, **68**, 57-80.

Christian, L. M., Dillman, D. A., and Smyth, J. D. (2007), "Helping the Respondents Get It Right the First Time: The Influence of Words, Symbols, and Graphics in Web Surveys," *Public Opinion Quarterly*, **71**, 113-125.

Couper, M. P., Tourangeau, R., and Kenyon, K. (2004), "Picture This! Exploring Visual Design Effects in Web Surveys," *Public Opinion Quarterly*, **68**, 255-266.

Dillman, D. A., Gertseva, A., and Mahon-Haft, T. (2005), "Achiving Usability in Establishment Surveys through the Application of Visual Design Principles," *Journal of Official Statistics*, **21**, 183-214.

Fuchs, M. (2002), "Comparing Internet and Paper and Pencil Surveys. Question Order Effects and Response Order Effects as Indicators for Differences of Data Collection Techniques," *Egyptian Journal of Public Opinion Research*, **3**, 21-43.

Fuchs, M. (2003), "Kognitive Prozesse Und Antwortverhalten in Einer Internet-Befragung," *Österreichische Zeitschrift für Soziologie*, **28**, 19-45.

Fuchs, M. (2007), "Asking for numbers and quantities Visual design effects in paper&pencil and Web surveys," (under review).

Fuchs, M., and Funke, F. (2007), "Video Web Survey. Results of an Experimental Comparison with a Text-Based Web Survey," in *Survey and Statistical Computing Iv. The Impact of Technology on the Survey Process*, ed. M. Trotman, Berkeley: Association for Survey Computing.

Jobe, J. B., Watt, W., Tourangeau, R., K., B. A., and Rasinski, K. A. (1997), "Effects of Interview Mode on Sensitive Questions in a Fertility Survey," in *Survey Measurement and Process Quality*, eds. L. Lyberg, P. Biemer, M. Collins, E. deLeeuw, C. Dippo, N. Schwarz and D. Trewin, New York: Wiley, pp. 311-329.

Sanchez, M. E. (1992), "Effects of Questionnaire Design on the Quality of Survey Data," *Public Opinion Quarterly*, **56**, 206-217.

Schneider, A. L., and Sumi, D. (1981), "Patterns of Forgetting and Telescoping: An Analysis of Leaa Survey Victimization Data," *Criminology*, **19**, 400-410.

Schwarz, N., and Hippler, H.-J. (1987), "What Response Scales May Tell Your Respondents: Information Functions of Response Alternatives," in *Social Information Processing and Survey Methodology*, eds. N. Schwarz, H.-J. Hippler and S. Sudman, New York: Springer, pp. 163-178.

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.

Smith, T. (1995), "Little Things Matter: A Sample of How Differences in Questionnaire Format Can Affect Survey Responses," in *annual conference of the American Association for Public Opinion Research*, American Statistical Association, pp. 1046-1051.

Smyth, J. D., Dillman, D. A., Christian, L. M., and Stern, M. J. (2006), "Effects of Using Visual Design Principles to Group Response Options in Web Surveys," *International Journal of Internet Science*, **1**, 6-16.

Smyth, J. D., Dillman, D. A., and Christian, L. M. (2007), "Improving Response Quality in List-Style Open-Ended Questions in Web and Telephone Surveys," *Annual conference of the American Association for Public Opinion Research.*

Sudman, S., Bradburn, N., and Schwarz, N. (1996), *Thinking About Answers. The Application of Cognitive Processes to Survey Methodology*, San Francisco: Jossey-Bass.

Tourangeau, R., Rips, L., and Rasinski, K. (2000), *The Psychology of Survey Response*, Cambridge, UK: Cambridge University Press.

Tourangeau, R., Couper, M. P., and Conrad, F. G. (2004), "Spacing, Position, and Order: Interpretive Heuristics for Visual Features of Survey Questions," *Public Opinion Quarterly*, **68**, 368-393.

Wänke, M. (2002), "Conversational Norms and the Interpretation of Vague Quantifiers," *Applied Cognitive Psychology*, **16**, 301-307.