

How Much Will I Spend? Factors Affecting Consumers' Estimates of Future Expense

Geeta Menon

*Assistant Professor of Marketing
New York University*

Priya Raghurir

*Assistant Professor of Marketing
University of California, Berkeley*

Norbert Schwarz

*Professor of Psychology and Research Scientist
University of Michigan*

Respondents in marketing surveys are often asked to estimate future expenses on products and services to assist marketers in arriving at market-share forecasts. To estimate future expenses for products and services, respondents may use information about past expenses, information about past behavioral frequencies for a related behavior, or both, if they are elicited prior to eliciting estimates of future expenses. To the extent that these sources of information are biased due to the manner in which they are elicited, subsequent estimates of future expense are biased. This article investigates the effect of response alternatives used to elicit behavioral frequencies on subsequent estimates of future expense related to the behavior. Two experiments demonstrate that past and future expense estimates are affected by the range of response alternatives used to elicit a related behavioral frequency prior to eliciting the expense estimate. This effect holds, however, only when the behavior is irregular. Further, the relative use of past expenses versus behavioral frequencies to estimate future expenses is a function of the presence and range of response alternatives used

to elicit the behavioral frequency, the regularity of the behavior, and the order in which these are elicited. The results emphasize that the importance of question sequencing in surveys is a function of the regularity of the target behavior.

What is your estimate for next month's grocery bill? Credit card bill? How much do you think you will spend on telephone calls in the next month? Consumers make estimates like these to not only estimate their disposable income and make decisions about possible future purchases but also to manage their cash flows on a monthly basis. From a decision-making standpoint, therefore, it is important to understand how consumers estimate their future expenses.

The issue of how consumers estimate expenses is also important from a marketing researcher's standpoint. Many marketing surveys determine consumers' estimates of future expenses for different types of products and services. These estimates are used for market-size and brand-share forecasting and also to undertake market segmentation. However, given that questions on future expenses are often preceded by detailed questions on past activity in the related area, it is appropriate to question whether such expense estimates are affected by the type and manner of questions preceding them. For example, many airlines query prior travel patterns before they ask about future travel plans. Financial services companies invariably precede questions relating to the amount you would invest in a variety of financial instruments with questions on past financial behavior. This helps determine market size and plan the marketing mix. Companies in the leisure and vacation industry use questions such as total amount spent per annum on vacations, and number of vacations taken, to identify prospects and design vacation offers for different segments. Media planners ask about readership of magazines before eliciting information related to intended subscriptions and spending. In short, a wide range of surveys, including those by academic researchers (e.g., Granbois & Summers, 1975), collect expenditure-related data. As the accuracy of these reports is important, it is critical to examine conditions under which expense estimates are prone to context effects.

Finally, from a cognitive psychology standpoint, the issue of future expense estimation has some unique aspects. An examination of these aspects may allow generalization of existing theories. Prior literature in this area has primarily studied the effect of question wording and question order on past behavioral frequency questions (e.g., Blair & Burton, 1987; Burton & Blair, 1991; Menon, 1993; Menon Raghbir, & Schwarz, 1995; Schwarz, Hippler, Deutsch, & Strack, 1985; for a recent review see Schwarz & Sudman, 1994), and attitudes (e.g., Menon, Bickart, Sudman, & Blair, 1995; Simmons, Bickart, & Lynch, 1993). Both these categories of judgments are to some extent based on information that exists in memory, albeit subject to errors due to the accessibility of such information. Judgments of the future, on the other hand, involve predictions partially based on a person's assess-

ment of what has occurred in the past. They are, therefore, qualitatively different from “factual” reports of the past or the present. By examining whether contextual factors, which affect behavioral frequency and attitudinal questions, also affect questions relating to estimates of the future, theories of the judgment formulation process in surveys may be extended and generalized.

In this article, we investigate a specific context effect that has been shown to bias estimates of behavioral frequencies and subsequent related judgments; the range of response alternatives used to elicit behavioral frequencies (Menon, Raghuram, et al., 1995; Schwarz et al., 1985). We examine whether the range of response alternatives affects subsequently elicited estimates of future expense in domains related to the reported behaviors, even if the latter are elicited in an open-ended format. In keeping with previous research, we predict that response alternatives for a question carry information about the population distribution that is used by the respondent to answer not only that question but may also be used to formulate responses to subsequent, related questions. Further, we predict that the use of this information is an inverse function of the accessibility of alternate sources of information in memory (Feldman & Lynch, 1988; Menon, Raghuram, et al., 1995; Schwarz et al., 1985; Simmons et al., 1993).

In the following sections, we describe two experiments based on our overall framework. We conclude with theoretical and managerial implications of our results.

EXPERIMENT 1

The purpose of Experiment 1 is twofold: (a) to explore the extent to which past expenses will be used as a source of information to estimate future expenses as a function of the presence of response alternatives used to elicit behavioral frequency questions and (b) to examine the effects of the range of response alternatives on estimated future expenses.

Research Hypotheses

The Use of Memory-Based Information in Estimates of Future Expenses

How do respondents estimate future product or service expenses? An extremely simple and intuitive way is to estimate what they spent in the past and extrapolate this to the future. For example, when you are asked to estimate how much you will spend on books in the next year, a simple method would be to estimate how much you spent last year (say \$400), adjust that amount for changed circumstances (say by \$100 for changes in income, more leisure time, etc.), and then come up with an

overall figure (\$500). Because an estimate of past expense may be one of the most diagnostic inputs for judgments of future expense, it follows that past expenses are likely to be directly extrapolated to arrive at estimates of future expense.

H1a: Estimates of past expense will be used by respondents in estimating future expense.

Alternatively, estimates of future expense may reflect individuals' assumptions about the frequency of an underlying behavior. For example, if you estimate that you will purchase 10 books and each of these costs \$50 on average, then your total estimate of book-related expense would be \$500. However, to the extent that the estimate of behavioral frequencies is itself inaccurate (e.g., you didn't actually buy 10 books last year as estimated), this error will follow through to estimates of future expense.

The Use of Context-Based Information in Estimates of Future Expenses

Feldman and Lynch (1988) developed a framework that may be useful in predicting which processes occur when. Specifically, this accessibility–diagnosticity framework predicts that “an earlier response will be used as an input to a subsequent response if the former is accessible and if it is perceived to be more diagnostic than other accessible inputs” (p. 431). Although *accessibility* refers to ease of retrieving an input from memory, *diagnosticity* refers to the sufficiency of the retrieved input to arrive at a solution for the judgment task at hand. Therefore, the likelihood that the response to an earlier question will be used in a response to a later question is a positive function of the accessibility of the earlier question in memory and its diagnosticity for the later question, and a negative function of the accessibility of alternate inputs and their diagnosticity (Simmons et al., 1993). Menon, Raghurir, et al. (1995) adapted this model and tested it in the context of past behavioral frequencies. Figure 1 exemplifies this model.

Figure 1 shows that when a respondent is posed a question in a survey, he or she has two alternate sources of information to draw on in order to come up with a response: information retrieved from memory (i.e., memory-based information) or information inferred from the context of the survey (i.e., context-based information). If information is accessible in memory and also available in the context, the relative diagnosticity of the two information sources would determine their use as inputs for the judgment task at hand. If only one piece of information is available, it will be used as an input for the judgment task, irrespective of its diagnosticity.

Returning to the example of estimating the expenses on books for the next year, a respondent may retrieve last year's expense estimate to extrapolate to the future

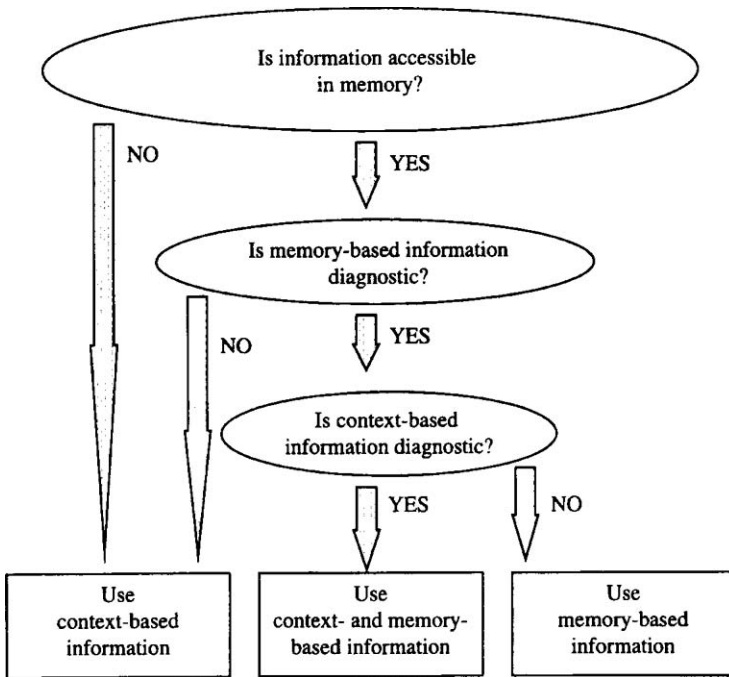


FIGURE 1 Response formulation process in a survey.

or use the estimated number of books to be purchased to construct an expense estimate. What factors might influence the relative use of these alternate sources of information? Because an estimate of past expense may be one of the most diagnostic inputs for judgments of future expense, it follows that past expenses are likely to be directly extrapolated to estimate future expense. However, based on the accessibility–diagnosticity framework, the use of past expenses as a cue would be inversely related to the accessibility of alternative sources of information. Therefore, to the extent that the context makes behavioral frequency estimates more accessible, past expenses would be less likely to be used as a source of information. Further, if behavioral frequencies are elicited using response alternatives, then respondents may not only use these to arrive at their own behavioral frequency estimates (Schwarz et al., 1985) but even when they do not do so, they may still infer the average behavioral frequency estimates from these (Menon, Raghurir, et al., 1995). Therefore, when response alternatives are present, respondents should be less likely to rely on past expense information as compared to when response alternatives have not been provided. Based on this we hypothesize:

H1b: The effect of past expenses on future expenses is less when related frequency judgments have been made using response alternatives versus an open-ended format.

This implies that to the extent future expense estimates are based less on past expenses, they may be based more on contextual cues. One context-based heuristic that has received considerable attention in survey methodology research involves the use of information provided by the range of response alternatives (e.g., Menon, Raghuram, et al., 1995; Schwarz & Hippler, 1987; Schwarz et al., 1985; Schwarz & Scheuring, 1988). Specifically, respondents assume that the researcher constructed a meaningful set of response alternatives that reflects the frequency distribution of the behavior under study. Accordingly, they assume that the extremes of the scale correspond to the extremes of the population distribution and that alternatives in the middle of the scale reflect the average or "typical" frequency of the behavior (Schwarz, 1990; Schwarz & Hippler, 1991). Consequently, respondents use the range of the response alternatives provided to them as a frame of reference in estimating their own behavioral frequency. This results in self-reported higher frequencies along response alternatives that present a high rather than a low range of frequencies. For example, Schwarz et al. (1985, Experiment 1) observed that 37.5% of a German sample reported watching TV for more than 2½ hr per day using response alternatives that ranged from "up to 2½ hr" to "more than 4½ hr," whereas only 16.2% did so when using response alternatives that ranged from "up to ½ hr" to "more than 2½ hr."

As may be expected on theoretical grounds (e.g., Feldman & Lynch, 1988), reliance on the range of response alternatives decreases when relevant behavioral information is easily accessible in memory (see Figure 1). Thus, the impact of response alternatives has been found to be more pronounced for proxy than for self-reports (e.g., Schwarz & Bienias, 1990) and greater as task complexity increases (Bless, Bohner, Hild, & Schwarz, 1992).

Menon, Raghuram, et al. (1995) proposed that the effect of response alternatives on behavioral frequency reports would also be greater for irregular behaviors than regular, frequently performed behaviors. They derived this prediction based on survey methods research that had shown that respondents estimate frequencies of a commonly performed behavior rather than recalling and counting individual instances (e.g., Blair & Burton, 1987, Menon, 1993). One such estimation strategy is to use a summary representation of frequency per unit of time—or a rate-of-occurrence (Blair & Burton, 1987; Burton & Blair, 1991). Rates-of-occurrence are formulated after extensive experience with a behavior and have been shown to be more accurate than the use of episodic recall for frequent behaviors (Menon, 1993, 1997). For example, using the verbal protocol technique to ascertain the manner in which respondents formulated frequencies, Menon (1993) demonstrated that when

rates-of-occurrence were used by respondents, reported frequency was more strongly correlated to an independent measure of actual frequency recorded in a diary. Moreover, Menon also observed that rates-of-occurrence were unlikely to be available for irregular behaviors that were engaged in without a fixed periodicity. Thus, if it is difficult to retrieve individual instances of frequent behaviors and rate-of-occurrence information is not available in memory for irregular ones, respondents may rely on the information provided by the context of the questionnaire. In sum, the range of response alternatives presents a contextual source of information that can be used by respondents to arrive at behavioral frequency judgments, but this is moderated by the regularity of the behavior. When respondents did not have access to memory-based rate-of-occurrence information (for irregular behaviors), the effect of the range of response alternatives was greater.

Using the accessibility–diagnosticity framework, we now argue that the use of the information contained in response alternatives is inversely related to the accessibility of alternative sources of information—in this case, memory-based behavioral frequency estimates. We propose that respondents base estimates of future expense on estimates of past behavioral frequencies. However, these frequencies are a function of the response alternatives presented by the researcher if the behavior is irregular, but not if the behavior is regular. If respondents consider behavioral frequency information in making future expense estimates, such that the higher the behavioral frequency, the greater the estimate of expense, these estimates should mirror the pattern of respondents' estimates of behavioral frequencies. Accordingly, future expense estimates should be affected by the response alternatives provided if the behavior is irregular, but not if the behavior is regular, or

- H2: The effect of response alternatives used to elicit past behavioral frequencies on estimates of future expenses on categories related to the target behavior is moderated by the regularity of the behavior, such that
- a. Estimates are higher when response alternatives use high rather than low frequencies for an irregular behavior.
 - b. The range of response alternatives does not affect estimates of future expense for a regular behavior.

Method

Participants

Ninety undergraduate students enrolled in an introductory marketing course at New York University participated for partial course credit.

Design

A 2×3 (Regularity of the Behavior: Regular and Irregular \times Response Alternatives: Low vs. High Frequency vs. None) mixed design was used. Regularity of the behavior was manipulated within-subjects, and response alternatives were manipulated between-subjects. The order of administration of the within-subjects factor was fully counterbalanced. As order of administration did not affect the results, this factor was dropped from the analyses.

Menon, Raghubir, et al. (1995) established regularity as a variable moderating the effects of response alternatives over 9 different behaviors in their Experiment 1. In Experiments 2 and 3 they used one behavior each to represent a regular behavior and an irregular behavior. In this article, we used the two behaviors studied by Menon, Raghubir, et al. (1995): The regular behavior was “washing hair” and the irregular behavior was “making unplanned stops to talk with friends.”

We also used the two sets of response alternatives used by Menon, Raghubir, et al. (1995). The low frequency set had the following alternatives: up to 1, 2–3, 4–5, 6–7, and more than 7 times. The high-frequency set comprised the following alternatives: up to 7, 8–9, 10–11, 12–13, and more than 13 times. Note that the first four categories of the low-frequency response alternatives are accommodated by the first category of the high-frequency response alternatives. Hence, reports along these two sets of response alternatives can be compared by determining the percentage of respondents who report frequencies of “less” (or “more”) than 7 times. We also included a no-response alternative condition in which frequency estimates were elicited in an open-ended fashion.

Choice of Expense Categories

The expense categories were selected based on two pretests.

Pretest 1. In this pretest ($n = 45$), we generated three expense categories that participants associated with each of the two behaviors of interest. With regard to the regular behavior, “washing hair,” these were (a) shampoo; (b) conditioner; and (c) hair spray, gel, etc.; whereas “making unplanned stops” was associated with socializing activities. Participants perceived that a person who makes a lot of unplanned stops must be a sociable person and therefore must engage in more social activities. Consequently, the three expense categories related to “making unplanned stops” were (a) going out to eat with friends, (b) going to bars, and (c) going to the movies.

Pretest 2. Note that our hypotheses regarding expenses relate to the regularity of the behaviors they are associated with rather than any inherent difference in the regularity of the expenses themselves. Therefore, it is important to match the

expenditure categories on (a) their regularity and (b) their frequency of occurrence. In the second pretest, therefore, a separate group of participants ($n = 41$) was asked to rate the regularity of the three expenses relating to each of the two behaviors and the frequency with which they made purchases in each of these categories. Of the three expenditures associated with each behavior, we determined that expenses on shampoo (expenditure associated with the regular behavior) and going to the movies (expenditure associated with the irregular behavior) were matched on regularity; means on a 7-point scale ranged from *irregular* to *regular*: 4.20 versus 4.17, $t(40) = .07$, $p = .946$; and on frequency; mean purchase frequency to an open-ended question about the last 3 months: 2.80 versus 3.37, $t(40) = 1.12$, $p = .269$. Note that these expense categories also have a stable expenditure-per-occurrence; that is, they do not vary much across purchase occasions. Based on these pretests, we selected expenses on shampoo as the expense category associated with the regular behavior and expenses on movies as the category associated with the irregular behavior.

Procedure

In the main experiment, past expenses pertaining to nine different expense categories were elicited through open-ended questions, two of which were “shampoo” and “movies.”¹ The order in which the expense categories were presented was randomized. To reduce demand effects, we then included a filler task before the next question. The filler task took approximately 5 min, after which participants estimated their own frequency of the target behavior using low- or high-frequency response alternatives or in an open-ended manner, depending on the condition they were assigned to. The question used was, “How often, in an average week, do you ...?” Next, they estimated their own future expenses for the next 3 months on the three expense categories related to the target behavior (as per Pretest 1) in an open-ended format: “Please estimate the total dollar amount you plan to spend on the following in the next 3 months.” This sequence was repeated for the second behavior, and the order in which the behaviors were presented was fully counter-balanced.

¹Note that out of the seven other categories, four were expenses associated with “conditioners,” “hair spray, gel, etc.,” “going out to eat with friends,” and “going to bars”; that is, the expense categories generated in Pretest 1. We conducted all the tests of hypotheses reported in this article using an index of the three expense categories for the regular and irregular behaviors. These analyses lead to similar results to those reported using only one expense category. In the interest of internal validity, only those expenses for “shampoo” and “going to the movies” are reported.

Results

Manipulation Checks

The effects of the range of response alternatives on behavioral frequency judgments. The predictions of the proposed model presuppose that the range of response alternatives affects behavioral frequency estimates for irregular behaviors, but not for regular behaviors. Therefore, to replicate the findings of Menon, Raghbir, et al. (1995), we conducted a loglinear analysis using respondents' frequency reports as a binary dependent variable: 0 (*up to 7 times*) and 1 (*more than 7 times*), with the range of response alternatives (low vs. high) and regularity of the behavior (regular vs. irregular) as independent variables. This analysis revealed the expected significant interaction between regularity and response alternatives, $\chi^2(4) = 26.04, p < .001$. The form of this interaction indicates that response alternatives did not affect the frequency judgments of the regular behavior; 21.4% versus 23.3%, $\chi^2(1) = .03, p = .86$ (see Table 1 for details of the distribution of behavioral frequencies). However, for the irregular behavior, a significantly greater percentage of respondents reported engaging in the behavior when they responded using the high-frequency response alternative (50%) than the low-frequency response alter-

TABLE 1
Experiment 1: Behavioral Frequency Distribution

<i>Low Response Alternatives</i>	<i>Regular Behavior</i>	
	<i>Percentage Responding</i>	<i>High Response Alternatives Percentage Responding</i>
Up to 1 time ^a	0.0	Up to 7 times ^a 76.7
2-3 times ^a	3.6	8-9 times ^b 23.3
4-5 times ^a	28.6	10-11 times ^b 0.0
6-7 times ^a	46.4	12-13 times ^b 0.0
> 7 times ^b	21.4	> 13 times ^b 0.0
<i>Low Response Alternatives</i>	<i>Irregular behavior</i>	
	<i>Percentage Responding</i>	<i>High Response Alternatives Percentage Responding</i>
Up to 1 time ^a	14.3	Up to 7 times ^a 50.0
2-3 times ^a	21.4	8-9 times ^b 23.3
4-5 times ^a	32.1	10-11 times ^b 6.7
6-7 times ^a	7.1	12-13 times ^b 3.3
> 7 times ^b	25.0	> 13 times ^b 16.7

Note. The categories in the low versus high response alternatives conditions sharing the same letter superscripts are the comparable categories.

native; 25%, $\chi^2(1) = 3.845, p < .05$. Therefore, these results successfully replicated the findings of Menon et al. (1995).

Matching the behaviors on frequency while varying them on regularity.

Further, an analysis of the frequencies in the open-ended control condition showed that the two behaviors were matched on frequency, with exactly 21.9% of the participants reporting that they had engaged in each of the behaviors "7 times" in the last week. The mean frequency for washing one's hair ($M = 5.6$) was also not different from that of making unplanned stops, $M = 5.7$; $t(31) = .12, p = .90$. Therefore, as desired, the two behaviors were matched on frequency.

Hypothesis 1

H1 predicts that estimates of past expense affect judgments of future expense (H1a) but do so less when frequency estimates are elicited using response alternatives versus an open-ended format (H1b). To test this hypothesis, we computed an interaction term between past expense and two levels of response alternatives (dummy-coded: 0 = no response alternatives, 1 = low- and high- response alternatives combined). We expect this interaction term to have a significant positive beta-coefficient. The regression using the full model comprising the interaction term between past expense and response alternatives and the two main effects was significant (model $R^2 = .57, p < .001$). As expected, the effect of past expense on estimates of future expense was significant ($B = .69, t = 7.66, p < .01$), as was the interaction term ($B = .12, t = 2.44, p < .001$). We also obtained a significant main effect of response alternatives ($B = 9.35, t = 2.43, p < .02$). Separate regressions, one each for the condition in which response alternatives were used and not used, showed that the coefficient associated with past expenses is higher when no response alternatives are provided, that is, in the open-ended condition ($R^2 = .69, p < .001; B = .95, t = 11.87, p < .001$) than when response alternatives are provided ($R^2 = .51, p < .001; B = .81, t = 10.12, p < .001$; test of significance between B's: $z = 1.25, p < .10$). Therefore, although past expenses are used to estimate future expenses, they are used to a greater extent when respondents are not exposed to response alternatives. In sum, the data are consistent with H1.

Hypothesis 2

H2 stated that when response alternatives were present, estimates of future expense would be lower if low (vs. high) frequency response alternatives were used to elicit behavioral frequencies; but this would only be true for irregular behaviors and not for regular behaviors. The mean estimates of past and future expenses by experimental condition are presented in Table 2.

TABLE 2
Experiment 1 Results

Dependent Variable	Regularity of the Behavior	Response Alternatives					
		Low Frequency ^a		High Frequency ^b		None (Open-Ended) ^c	
		M	SD	M	SD	M	SD
Estimates of Past Expense	Regular	10.86	7.03	14.17	12.21	12.65	9.69
	Irregular	30.39	21.22	32.73	38.36	29.31	26.31
Estimates of Future Expense	Regular	12.68	9.54	15.33	9.57	13.91	10.46
	Irregular	36.64	20.99	57.83	37.73	39.97	27.34

^a*n* = 28. ^b*n* = 30. ^c*n* = 32.

A 2×2 multivariate analysis of variance treating regularity as a within-subjects factor and response alternatives as a between-subjects variable reveals that the main effects of regularity of the behavior, $F(1, 56) = 65.57, p < .001$, and response alternatives, $F(1, 56) = 7.44, p < .008$, are qualified by the expected interaction between these two terms, $F(1, 56) = 5.10, p < .03$). The interaction is of the hypothesized pattern: Although the range of response alternatives affects estimates of future expense for an irregular behavior; $M = \$36.64$ versus $\$57.83$; $F(1, 56) = 6.85, p < .01$, it does not for a regular behavior; $M = \$12.68$ versus $\$15.33, F(1, 56) = 1.12, p = .29$.

Therefore, H2, arguing that the higher the range of response alternatives the greater the expense estimate for an irregular behavior, was supported.²

Discussion

The data support the model that estimates of future expense are based on estimates of past expense as well as estimates of past behavioral frequencies related to the expense category. These frequency estimates, however, are themselves influenced

²As can be seen in the mean expenditures reported in Table 2, the mean dollar amounts are higher for "movies" versus "shampoo." This issue of different absolute expenditure levels can be handled statistically through the creation of an expenditure index that is the difference in the estimates of future and past expenses divided by the past expense. The ensuing results are robust and do not change, demonstrating that the effect is real and not explainable through differential absolute expenditure levels.

by the response alternatives used to elicit the report. Although the use of both these pieces of information to estimate future expenses varies as a function of the regularity of the behavior, the extent to which estimates of past expenses are relied on is also determined by the availability of response alternatives to elicit frequencies.

Experiment 1 found that future expenses were based on estimates on past expenses (H1) and context-based information if the expenses were related to an irregular behavior (H2a). However, if context-based information were itself to affect estimates of past expense, then it may affect estimated future expenses in both a direct and indirect manner. This would imply that the order in which the two expense estimates (past and future) are elicited should affect the extent to which future expenses are based on contextual information. Experiment 2 examines this issue.

EXPERIMENT 2

The purpose of Experiment 2 is to examine the effect of question sequencing on the manner in which past expense information versus the range of response alternatives are used to construct future expense estimates.

Research Hypotheses

Experiment 1 showed that future expenses are based on (a) past expense estimates (if rendered accessible by the context) and (b) response alternatives used to elicit frequency judgments. However, if past expense estimates (like future expense estimates) are also affected by response alternatives used to elicit frequency judgments, then the range of response alternatives can affect estimated future expense directly and indirectly (through also affecting past expense estimates). Note that any response alternative effect should occur only for irregular behaviors, as estimates based on regular behaviors are less context dependent (Menon, Raghbir, et al., 1995, Experiment 1). For irregular behaviors, whether the range of response alternatives affects future expense estimates directly or indirectly is dependent on question order: whether past expense estimates are elicited prior to or subsequent to future expense estimates, with both being elicited subsequent to behavioral frequency questions using response alternatives.

If past expense estimates are elicited after behavioral frequency estimates, then the range of response alternatives can affect past expense estimates in a similar manner as it does estimates of future expense. Or, extending H2:

- H3: The effects of response alternatives used to elicit behavioral frequencies on past expense estimates will be moderated by the regularity of the behavior such that estimates of past expenses will be lower when the

response alternatives used to elicit frequencies is low (vs. high) for an irregular behavior, but not a regular one.

However, given that future expenses are based on past expenses, if past expense estimates are elicited before future expense estimates, the effect of response alternatives on past expense estimates should carry through to future expense estimates. This would be an indirect effect. However, the response alternatives may also affect estimates of future expenses directly, that is, a direct effect. Again, such effects should only manifest for irregular behaviors. Therefore:

H4: Estimated past expense mediates the effect of response alternatives on future expense estimates for irregular behaviors, but not for regular behaviors.

If future expenses are elicited before past expenses, then the effect of response alternatives on future expense estimates will be less evident when these are elicited immediately after the behavioral frequency question. This is because there will only be a direct effect of response alternatives on future expense estimates. The indirect effect will not be manifested due to the question order. Again, any effect of response alternatives on estimates of expense is contingent on the regularity of the target behavior. In sum:

H5: The effects of response alternatives on estimates of future expenses is moderated by the order in which the past and future expenses are elicited and the regularity of the behavior:

- a. For a regular behavior, the response alternatives will not have an effect, irrespective of the order in which the expense estimates are elicited.
- b. For an irregular behavior, the effect of the response alternatives will be less when future expenses are elicited before past expenses.

We now test whether the range of response alternatives affects estimates of past expense related to an irregular behavior in a manner similar to that of future expense (H3) and whether this effect carries through to estimates of future expense (H4), thereby exacerbating the overall impact of response alternatives on estimates of future expense (H5). Null effects are hypothesized for the regular behavior.

Method

Participants

Sixty-two participants enrolled in an introductory marketing course at New York University participated in this experiment for partial course credit.

Design

We used a $2 \times 2 \times 2$ (Response Alternatives: Low vs. High \times Regularity of the Behavior: Regular and Irregular \times Question Order for Past vs. Future Expenditure Estimates: Past-Future and Future-Past) full factorial mixed design. Although response alternatives and question order were manipulated between subjects, the regularity of the behavior was manipulated within-subjects and counterbalanced.

We used the same manipulations for response alternatives and regularity of the behavior as in Experiment 1. Further, in this experiment, we rotated the order in which past and future expenses were asked after behavioral frequencies were elicited using the response alternatives manipulation, to enable us to test hypotheses 3, 4, and 5.

Procedure

Participants were randomly assigned to the four Response Alternatives \times Question Order conditions. They were first asked their behavioral frequency using a low or a high set of response alternatives. They were then asked to estimate past or future expenses for the last/next 3 months on the same three product categories as in Experiment 1. A 5-min filler task was administered between the past and future expense questions. Manipulation check information was then collected.

Results

Manipulation Checks

Regularity of behaviors. Following the procedures of Menon (1993) and Menon, Raghurir, et al. (1995), we asked participants to rate the regularity of the two behaviors on a 7-point semantic-differential scale ranging from 1 (*not at all regular*) to 7 (*very regular*). As expected, “washing hair” was rated as being more regular ($M = 6.40$) than “making unplanned stops,” $M = 5.08$; $F(1, 61) = 31.15$, $p < .001$.

Difficulty associated with the judgment tasks. Further, following the method used by Menon, Raghurir, et al. (1995), we used a Difficulty Index of four scales to measure the task difficulty of responding to the frequency question (Cronbach’s $\alpha = 0.85$). These were 7-point scales ranging from 1 (*not at all difficult*) to 7 (*very difficult*); 1 (*no effort*) to 7 (*a lot of effort*); 1 (*no time*) to 7 (*a lot of time*); and 1 (*no thought*) to 7 (*a lot of thought*). We expected that the difficulty of answering the frequency question would be lower for the regular versus the irregular

behavior, given the accessibility of a rate of occurrence to answer this question (see also Menon, 1993). This is exactly what happened; M : regular = 2.93, irregular = 3.60; $F(1, 61) = 22.73, p < .001$. Therefore, the behavior manipulation worked in the manner intended.

Effects of the range of response alternatives on behavioral frequency judgments. Once again, we successfully replicated the Menon, Raghbir, et al. (1995) results regarding the effect of the range of response alternatives on behavioral frequency estimates for regular and irregular behaviors. The distributions of the behavioral frequency reports for the two sets of response alternatives are presented in Table 3. As before, the response alternatives (low vs. high) used to elicit behavioral frequencies affected the irregular behavior, such that the percentage reporting a frequency of more than seven times was lower when the low-response alternatives were used versus the high-response alternatives, 15.4% versus 52.0%, $\chi^2(1) = 7.69, p < .001$). The response alternatives had no effect on the regular behavior, 19.2% versus 24.0%, $\chi^2(1) = 0.17, p = .68$; see Table 3.

Hypotheses 2 and 3

Table 4 presents the mean estimates of future and past expenses in each of the different experimental cells.

H2 predicted that the response alternatives would affect the future expense estimates for irregular but not for regular behaviors—that is, an interaction between response alternatives and the regularity of the behavior, with future expense estimates as the dependent variables. As predicted, this interaction was significant, $F(1, 59) = 28.76, p < .01$. For regular behaviors there was no difference in expense estimates irrespective of the range of response alternatives that had been used to elicit behavioral frequencies; mean expenses by response alternatives: low = \$15.10, high = \$13.97; contrast $F(1, 59) = 0.25, p = .62$. In contrast, for irregular behaviors, higher expenses were reported when respondents responded to the preceding behavioral frequency question using the higher range of response alternatives; means expenses by response alternatives: low = \$17.33, high = \$39.97; contrast $F(1, 61) = 44.07, p < .01$. As before, the overall main effect of response alternatives was significant, $F(1, 59) = 28.98, p = .01$, as was that of regularity, $F(1, 59) = 43.70, p < .01$. Therefore, the results of Experiment 1 were successfully replicated.

H3 predicted a similar effect on past expenses. The overall main effect of response alternatives is significant, $F(1, 60) = 8.17, p < .01$, as is the main effect of regularity, $F(1, 60) = 43.32, p < .01$. The latter was to be expected given the range in expenses

TABLE 3
Experiment 2: Behavioral Frequency Distribution

<i>Regular Behavior</i>			
<i>Low Response Alternatives</i>	<i>Percentage Responding</i>	<i>High Response Alternatives</i>	<i>Percentage Responding</i>
Up to 1 time ^a	15.4	Up to 7 times ^a	76.0
2-3 times ^a	30.8	8-9 times ^b	12.0
4-5 times ^a	11.5	10-11 times ^b	4.0
6-7 times ^a	23.1	12-13 times ^b	4.0
> 7 times ^b	19.2	> 13 times ^b	4.0
<i>Irregular Behavior</i>			
<i>Low Response Alternatives</i>	<i>Percentage Responding</i>	<i>High Response Alternatives</i>	<i>Percentage Responding</i>
Up to 1 time ^a	15.4	Up to 7 times ^a	48.0
2-3 times ^a	23.1	8-9 times ^b	16.0
4-5 times ^a	26.9	10-11 times ^b	12.0
6-7 times ^a	19.2	12-13 times ^b	4.0
> 7 times ^b	15.4	> 13 times ^b	20.0

Note. The categories in the low versus high response alternatives conditions sharing the same letter superscripts are the comparable categories.

TABLE 4
Experiment 2 Results

		<i>Regularity of the Behavior^a</i>							
		<i>Regular</i>				<i>Irregular</i>			
		<i>Order of Eliciting Expenditure Information</i>							
				<i>Past-Future</i>		<i>Future-Past</i>		<i>Past-Future</i>	
<i>Estimates</i>	<i>Response Alternatives</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Past expense	Low	13.77	7.35	15.09	9.93	23.92	21.68	17.73	11.83
	High	12.91	8.88	11.30	7.18	44.01	9.34	36.90	13.27
Future expense	Low	15.56	8.12	14.60	9.98	14.50	9.48	21.33	15.71
	High	15.31	9.86	12.43	6.82	45.25	11.73	33.93	14.25

Note. The dependent variable is the expenditure estimates.

^an = 62.

for the two behaviors. Further, as predicted, the main effects are qualified by a significant interaction effect between the two factors, $F(1, 60) = 16.61, p < .01$, with the response alternatives affecting estimates for irregular behaviors; mean expenses by response alternatives: low = \$21.08, high = \$40.67; contrast $F(1, 60) = 14.88, p < .01$, but not for regular behaviors; mean expenses by response alternatives: low = \$14.57, high = \$12.40; contrast $F(1, 60) = 1.03, p = .32$.

In sum, the data support H2 and H3.

Hypothesis 4

H4 argued that past expenses will mediate the use of response alternatives in arriving at future expenses for irregular behaviors, but not for regular behaviors. To test this hypothesis, we examine the condition in which past expenses were elicited before estimates of future expenses. To conduct the mediational analysis, the three regression models (M) needed are (cf. Baron & Kenny, 1986):

- MA: Regress the mediator (i.e., past expense estimates) on the independent variable (i.e., response alternatives).
- MB: Regress the dependent variable (i.e., future expenses estimates) on the independent variable (i.e., response alternatives).
- MC: Regress the dependent variable (i.e., future expenses estimates) on both the independent (i.e., response alternatives) and mediator variables (i.e., past expenses estimates).

Because we were predicting differences as a function of the regularity of the behavior, we did the mediational analyses separately for the two behaviors. However, first we ran an overall regression in which we included a full model of regularity, response alternatives, estimates of past expense, and the consequent interaction terms, and we used estimates of future expenses as the dependent variables. We used the reported dollar estimates for past and future expenses and dummy-coded the response alternatives condition (0 = low and 1 = high). This model was significant overall ($R^2 = .76$) as was the coefficient associated with the interaction term between past expenses and regularity ($B = -.69, t = -2.64, p < .01$), supporting our hypothesis.

Next, we tested each of the aforementioned Models A, B, and C for each level of regularity. As predicted, for the regular behavior, neither Model A nor Model B is significant (Model A: $R^2 = .00, p = .82$; Model B: $R^2 = .00, p = .93$). However, as one would expect, Model C is significant ($R^2 = .62, p < .001$), with the beta-coefficient associated with past expenses significant ($B = .88, t = 6.00, p < .001$) and the one associated with response alternatives not significant ($B = -1.23, t = -0.60, p = .55$). Therefore, as hypothesized, for regular behaviors response

alternatives do not have a direct or an indirect effect on estimates of future expenses. Further, as expected, past expenses are a significant predictor of future expenses.

For the irregular behavior, on the other hand, we had predicted that past expenses act as a mediator to the effects of response alternatives on future expenses. Models A and B are significant (Model A: $R^2 = .13$, $p < .031$ with $B = 14.95$, $t = 1.94$, $p < .03$; Model B: $R^2 = .69$, $p < .01$ with $B = 34.25$, $t = 8.52$, $p < .01$). Model C is also significant ($R^2 = .71$, $p < .001$), with the beta-coefficient associated with past expenses being significant ($B = 0.24$, $t = 2.24$, $p < .02$) and the one associated with response alternatives being significant, though lower than in Model B ($B = 26.92$, $t = 6.25$, $p < .001$; $z = 1.24$, $p < .10$). According to Baron and Kenny (1986), the decrease in the coefficient associated with the independent variable in Model C compared to Model B, in the presence of its being significant in Model A, is an indication of partial mediation. Therefore, the hypothesis that both direct and indirect effects of response alternatives manifest for irregular behaviors is borne out.

Thus, the mediational hypothesis predicted by H4 is supported.

Hypothesis 5

H5 predicted that although response alternatives have no effect on estimates of future expenses for regular behaviors, the magnitude of the effect on estimates of future expenses for irregular behaviors is contingent on the order in which past and future expenses are elicited after respondents have been exposed to the response alternatives. In other words, we expected a three-way interaction between regularity, order of eliciting past and future expenses, and response alternatives in the omnibus analysis of variance using future expenses as the dependent measure. This is indeed what happened, $F(1, 57) = 3.58$, $p < .03$; see bottom half of Table 4 for means and standard deviations.

We then examined the simple effects for regular and irregular behaviors. As expected, the two-way interaction between the order of elicitation and response alternatives was not significant for the regular behavior ($F < 1.0$) but was significant for the irregular behavior, $F(1, 59) = 8.25$, $p < .01$. As predicted by H5a, for the regular behavior, response alternatives had no effect irrespective of the order of elicitation ($ps > .50$). In the case of the irregular behavior, as per H5b, response alternatives had an effect on estimates of future expenditure when past expenses were elicited before future expenses, $F(1, 30) = 66.49$, $p < .01$. This effect was attenuated but remained significant when future expenses were elicited prior to past expenses, $F(1, 27) = 5.09$, $p < .015$.

Therefore, H5 was supported. Response alternatives did not affect estimates of future expenses of regular behaviors regardless of whether or not estimates of past expenses were accessible in memory. Response alternatives did affect estimates for

irregular behaviors to a greater extent when the context made past expense estimates accessible.

Discussion

The findings of Experiment 2 corroborate the results of Experiment 1 and extend the theory further. The results demonstrate that response alternatives used to elicit behavioral frequencies in a prior question may affect responses to later questions in the case of irregular behaviors, but not regular behaviors. We demonstrate this effect for both estimates of past and future expenses. Further, we demonstrate that the effects of such response alternatives on estimates of future expenses may be direct or indirect through other prior questions (in this case, estimates of past expenses) for irregular behaviors. Therefore, in the interest of unbiased data, question sequencing is critical for irregular behaviors. The results also testify to the robustness of information in memory for regular behaviors, especially because response alternatives used in a prior question had no impact, regardless of the order in which past and future expenses were elicited in subsequent questions.

GENERAL DISCUSSION

The two experiments reported in this article indicate that the results pertaining to future expense parallel the impact of response alternatives on behavioral frequency reports, providing evidence consistent with the proposition that estimates of future expenditure are partially based on estimates of behavioral frequency. The higher the range of response alternatives, the higher the behavioral frequency reports, and in turn, the higher the estimates of future expense. However, the effect of response alternatives on future expense estimates is contingent on the regularity of the behavior—with the effect only manifested for the irregular behavior, and not the regular behavior. This pattern continues when past expenses are elicited after behavioral frequency reports, with the range of response alternatives affecting estimated past expense, and through this exerting even more of an effect on future expense estimates—but, again, only for irregular behaviors. Theoretical and managerial implications of these findings are discussed later.

Theoretical Implications

Use of Memory-Based Versus Context-Based Information in Judgment Formulation

The effects of response alternatives have received considerable attention in research on questionnaire design in recent years (e.g., Bless et al., 1992; Menon, Raghurir,

et al., 1995; Schwarz & Bienias, 1990; Schwarz et al., 1985; see Schwarz & Hippler, 1991, for a review). In marketing research, such effects become particularly important, as many product decisions are based on estimates of consumer usage and purchase. Such estimates are used in making brand-share forecasts. The research presented here was guided by our thinking that the effects of response context extend to future estimates of expenses related to the behaviors they measure. We found evidence consistent with the notion that the accessibility of diagnostic information in memory renders the information provided by the response alternatives less influential when respondents are asked about their own behavioral frequency and related future expenses. As the results indicate, respondents rely on the information provided by response alternatives for irregular but not for regular behaviors. Accordingly, estimates of future expenditure are also biased by the response alternatives used to elicit frequency reports for an irregular, but not for a regular behavior.

Process of Estimating Future Expenses

This research contributes to our understanding of the cognitive processes involved in estimating future expenses and bears on key issues of questionnaire construction. Menon (1993) demonstrated that respondents are most accurate in formulating behavioral frequencies for a behavior that is regular and for which they have a rate-of-occurrence available in memory. The experiment presented in this article substantiates the robustness of the rate-of-occurrence as a judgmental heuristic. For a regular behavior, respondents rely on a rate-of-occurrence accessible in memory in formulating frequency judgments and thus are not affected by the response alternatives provided. More interestingly, behavioral regularity moderates the use of information inferred from response alternatives while estimating of future expense.

Additional Support of the Accessibility–Diagnosticity Framework

Our findings support the outcome of previous research that the use of information sources is a function of their accessibility (e.g., Menon, Raghbir, et al., 1995; Simmons et al., 1993). Given that more than one information source is accessible, their use is an inverse function of the accessibility and diagnosticity of the alternative sources of information (Feldman & Lynch, 1988).

We tested this in the context of future expense estimates. The two alternate sources of information were conceptualized as being past expenses and behavioral frequencies, with the latter affected by the range of response alternatives on which they were elicited for irregular behaviors. We successfully demonstrated the generalizability of the accessibility–diagnosticity framework by showing that the

use of past expenses as a source of information to forecast future expenses was contingent on the availability of easily accessible behavioral frequency data and the regularity of the behavior.

Practical Implications

When should market researchers pay attention to question-sequencing? Because eliciting information on future expenses is an integral part of consumer research, it is important to ensure that such reports are not affected by the manner in which they are elicited. There is a plethora of research on question order effects that has determined that responses to an earlier question undoubtedly affect judgments to later questions (e.g., Bickart, 1993; Menon, Raghbir, et al., 1995; Schwarz & Bless, 1992; Schwarz, Strack, & Mai, 1991; Simmons et al., 1993). The results of the two experiments reported in this article clearly provide additional corroborative support. Specifically, we demonstrated three scenarios in which these effects could manifest.

Effects of the Wording of Prior Questions on Responses to Later Questions

First, we demonstrated that the response alternatives used to elicit behavioral frequencies affects the responses to questions on expenditure that are elicited in an open-ended manner when the basic behavior is irregular in nature. What this implies is that although questions related to regular behaviors may be less susceptible to contextual influences (presumably because of the accessibility of memory-based information; see Figure 1), questions related to irregular behaviors need to be framed carefully, and special attention needs to be paid to the sequencing of such questions in a questionnaire.

“Long-Term” Effects of Prior Questions on Response to Later Questions

The results of Experiment 2 also reveal that the range of response alternatives used to elicit behavioral frequencies not only has an impact on responses to the question immediately following this one (in this case estimates of past expenses) but they also have a direct effect on the other subsequent questions (in this case the estimates of future expenses that followed the estimates of past expenses). Therefore, we demonstrate the relatively “long-term” effects of prior questions on responses to later questions.

Effects of Responses to Prior Questions on Response to Later Questions

Further, the results of Experiment 2 indicate that there is also an indirect effect of the response alternatives on estimates of future expenses through estimates of past expenses for irregular behaviors. Therefore, responses to prior questions could also affect responses to later questions. This particular issue is relatively underresearched and probably warrants more attention in the future given that accuracy of survey results may be at stake.

Future Research

The results reported in the two experiments are restricted to estimates of expense related to frequent behaviors. Although there is convincing evidence that the regularity of the behavior moderates the effect of response alternatives for such estimates, it would be interesting to study the case of expenses relating to infrequent behaviors such as airline travel, vacations, and so forth, where episodic memory may be highly accessible and accurate. In such cases it is possible that estimates of future expense may be made in a more piecemeal fashion (i.e., based on behavioral frequency estimates) rather than through an overall extrapolation from the past to the future. This would imply that such expenses may be less liable to the biases noted in this article.

At a more general level, what this implies is that different kinds of information may be accessible depending on the actual frequency of the behavior under investigation. Accordingly, while specific episodes may be highly accessible in the case of infrequent behaviors, more generalized scripts are accessible for frequent behaviors. The degree to which contextual influences prevail (such as the ones demonstrated in this article) will be highly contingent on the diagnosticity of this accessible information. This is an interesting avenue for future research.

ACKNOWLEDGMENT

Order of authorship is alphabetical. We thank the Editor, Paul Herr, and two anonymous *Journal of Consumer Psychology* reviewers for their very helpful and constructive comments.

REFERENCES

- Baron, R. M. & Kenny, D. A. (1986). The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology*, 51, 1173–1182.

- Bickart, B. (1993). Carryover and backfire effects in marketing research. *Journal of Marketing Research*, 30, 52–62.
- Blair, E., & Burton, S. (1987). Cognitive processes used by survey respondents to answer behavioral frequency questions. *Journal of Consumer Research*, 14, 280–8.
- Bless, H., Bohner, G., Hild, T., & Schwarz, N. (1992). Asking difficult questions: Task complexity increases the impact of response alternatives. *European Journal of Social Psychology*, 22, 309–12.
- Burton, S., & Blair, E. (1991). Task conditions, response formulation processes, and response accuracy for behavioral frequency questions in surveys. *Public Opinion Quarterly*, 55, 50–79.
- Feldman, J. M., & Lynch, J. G., Jr. (1988). Self-generated validity and other effects of measurement on belief, attitude, intention and behavior. *Journal of Applied Psychology*, 73, 421–35.
- Granbois, D. H., & Summers, J. O. (1975). Primary and secondary validity of consumer purchase probabilities. *Journal of Consumer Research*, 1, 31–8.
- Menon, G. (1993). The effects of accessibility of information in memory on judgments of behavioral frequencies. *Journal of Consumer Research*, 20, 431–40.
- Menon, G. (in press). Are the parts better than the whole? The effects of decompositional questions on judgments of frequent behaviors. *Journal of Marketing Research*.
- Menon, G., Bickart, B., Sudman, S., & Blair, J. (1995). How well do you know your partner? Strategies for formulating proxy-reports and their effects on convergence to self-reports. *Journal of Marketing Research*, 32, 75–84.
- Menon, G., Raghurir, P., & Schwarz, N. (1995). Behavioral frequency judgments: An accessibility–diagnosticity framework. *Journal of Consumer Research*, 22, 212–28.
- Schwarz, N. (1990). Assessing frequency reports of mundane behavior: Contribution of cognitive psychology to questionnaire construction. In C. Hendrick & M. A. Clark (Eds.), *Research methods in personality and social psychology* (pp. 98–119). Newbury Park, CA: Sage.
- Schwarz, N., & Bienias, J. (1990). What mediates the impact of response alternatives on frequency reports of mundane behaviors? *Applied Cognitive Psychology*, 4, 61–72.
- Schwarz, N., & Bless, H. (1992). Constructing reality and its alternatives: An inclusion/exclusion model of assimilation and contrast effects in social judgment. In L. Martin & A. Tesser (Eds.), *The construction of social judgment* (pp. 217–245). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.
- Schwarz, N., & Hippler, H-J. (1987). What response scales may tell your respondents: Informative functions of response alternatives. In H-J. Hippler, N. Schwarz, & S. Sudman (Eds.), *Social information processing and survey methodology* (pp. 163–78). New York: Springer-Verlag.
- Schwarz, N., & Hippler, H-J. (1991). Response alternatives: The impact of their choice and presentation order. In P. P. Biemer, R. M. Groves, L. E. Lyberg, N. A. Mathiowetz, & S. Sudman (Eds.), *Measurement errors in surveys* (pp. 41–56). New York: Wiley.
- Schwarz, N., Hippler, H-J., Deutsch, B., & Strack, F. (1985). Response categories: Effects on behavioral reports and comparative judgments. *Public Opinion Quarterly*, 49, 388–95.
- Schwarz, N., & Scheuring, B. (1988). Judgments of relationship satisfaction: Inter- and intraindividual comparisons as a function of questionnaire structure. *European Journal of Social Psychology*, 18, 485–96.
- Schwarz, N., Strack, F., & Mai, H-P. (1991). Assimilation and contrast effects in part-whole question sequences: A conversational logic analysis. *Public Opinion Quarterly*, 55, 3–23.
- Schwarz, N., & Sudman, S. (1994). *Autobiographical memory and the validity of retrospective reports*. New York: Springer-Verlag.
- Simmons, C. J., Bickart, B. A., & Lynch, J. G., Jr. (1993). Capturing and creating public opinion in survey research. *Journal of Consumer Research*, 20, 316–29.

Accepted by Paul Herr.