When More Is Less: A Study on Effects of Providing Definitions to Everyday Terms on Data Quality

Ting Yan
Joint Program in Survey Methodology, University of Maryland

Abstract

This study investigated the possible effects of providing such definitions on data quality. I provided unnecessary (and not very informative) definitions for everyday terms like "poultry" and "fat" in one condition, contrasting this with a condition in which respondents did not get any definitions. Instead of being confused by the apparently redundant definitions, the results seemed to suggest that respondents incorporated the definitions into their interpretation of the survey questions. The data provided weak support for predictions based on Grice' maxim of manner; a larger percentage of respondents inferred that the definitions were intended for someone else or the terms were used in a more technical sense when they were given the redundant definitions than when they were not given them. Respondents also seemed to use those redundant definitions; their responses to (some of) the key items had significantly smaller variances when redundant definitions were present than when they were not. The covariance structures of responses to key items and to related items were changed as well when redundant definitions were provided.

Keywords: measurement error, definitions, data quality, pragmatic effects,

1. Introduction

According to the survey response process framework, survey responding starts with comprehension of survey questions (Tourangeau, Rips, & Rasinski, 2000). Incorrect understanding of survey questions will affect – directly or indirectly – the retrieval of relevant information, the estimation and judgment strategies used, and even the mapping of an answer to one of the response options (Conrad & Schober, 2000; Schober & Conrad, 1997). Survey researchers have known for some time that, despite their best attempts to write clear questions, respondents have problems comprehending survey questions (Belson, 1981; Conrad & Schober, 2000; Fowler, 1992; Schober & Conrad, 1997).

There are at least two types of comprehension problems, both of which may affect the validity of the survey data. First of all, different respondents may interpret the same question very differently. In a well-known study, Belson (1981) showed that respondents vary in their interpretation of even simple words such as 'you.' For instance, while 84.6% of the respondents believed that the term 'you' referred to the individual alone, 1.9% thought it meant the individual and his family. For another 3.8% of respondents, 'you' was the individual and the spouse. The other 3.8% considered 'you' as a combination of the individual plus at least one other family member. What is more alarming is that 5.8% of respondents simply overlooked the term (Belson, 1981). Such a lack of consensus on the meaning of key survey terms (and survey questions as a whole) could lead to systematic variations, affecting the comparability of data across respondents, particularly if the differences of interpretation are large and coincide with boundaries of subgroups defined by culture, race, age or other characteristics (Martin, Campanelli, & Fay, 1991). It could also reduce survey researchers' ability to reach valid conclusions about relationships based on the data because of the error in measurement (see Fuller, 1991, on the effects of measurement error on regression coefficients and other statistics).

Second, respondents could be answering based on the same interpretation of the question, but one that does not fit the survey researchers' definitions. Respondents' interpretations may disagree with a survey definition by either being too broad or too narrow compared with the meaning intended in the survey. Suessbrick, Schober, and Conrad (2000) reported, for instance, that about 46% of respondents considered 'smoking' in the first question in the Tobacco Use Supplement to the Current Population Survey ("Have you smoked at least 100 cigarettes in your entire life?") as "only puffs inhaled," whereas the survey's definition included all puffs, whether inhaled or not. Such a difference between respondents' interpretations and survey's definitions could contribute to bias in the survey estimates.

Both variability across respondents and systematic misinterpretation jeopardize survey quality. Offering respondents definitions to clarify unclear or ambiguous terms is one possible solution. There is empirical evidence that providing definitions helps improve respondents' comprehension. For example, Fowler

(1992) revised seven questions from health surveys conducted by government agencies or academic survey organizations by offering definitions to clarify unclear terms. He found that the revised questions improved comprehension; the rates of requests for clarification and rates of inadequate answers declined (Fowler, 1992).

Similarly, Conrad and Schober demonstrated that uniformity of interpretation – and thus data quality – could be increased dramatically when respondents are provided with clarification about the meaning of the words in the questions (Conrad & Schober, 2000; Schober & Conrad, 1997; Schober, Conrad, & Fricker, 1999; Suessbrick, Schober, & Conrad, 2000). For instance, Schober and Conrad (1997) showed that allowing interviewers to provide definitions when the questions involved complicated mapping – that is, when the respondents' situations didn't map to the key terms in the questions in a straightforward way – greatly improved the accuracy of responses (87%) relative to interviewers who were not allowed to provide definitions (27% accuracy).

In another study, Conrad and Schober (2000) compared response changes in a reinterview due to different interviewing techniques. They observed that 22% of respondents changed their answers, on average, when the reinterview was conducted as a conversational interview in which interviewers provided definitions to respondents. By contrast, only about half as many changes occurred when the second interview was still standardized.

The existing research on the effects of definitions on survey responses has focused more on the mode of presenting definitions than on the content of the definitions themselves (see Tourangeau & Conrad, 2004, for an exception). These earlier studies have examined such aspects of offering definitions as the trigger for providing definitions (a request from the respondent, the interviewers' own initiative, or an automated feature in a web survey) and the accessibility of definitions (always shown, shown when the respondent click or rollover, and so on) in web surveys (see, for example, Bloom & Schober, 1999; Conrad, Couper, Tourangeau, & Peytchev, in press; Lind, Schober, & Conrad, 2001; Schober, Conrad, & Bloom, 2000; Schober, Conrad, & Fricker, 1999). These studies uncovered some problems with offering definitions to respondents to improve comprehension of the question.

One important finding is that the existing work also shows an increased interview time on average when definitions are requested, retrieved, or consulted (Conrad et al., in press; Conrad & Schober, 2000; Schober & Conrad, 1997). This seemed to provide further proof for the speculation that processing definitions requires efforts. It is also unclear whether respondents would be willing to spend the efforts to process the definition carefully even if they did access it.

Groves and his colleagues argue that there is a tension between explicitly defining terms in a question (in an attempt to eliminate ambiguity) and increasing the burden on the respondents to absorb the full intent of the question (Groves, Couper, Fowler, Lepkowski, Singer, & Tourangeau, 2004). Similarly, Tourangeau, Rips, and Rasinski (2000: Chapter 2) argue that attempts to clarify terms can lead to syntactically complex questions. The trade-off involves the amount of information to give in a definition so that it clarifies meaning without seeming redundant or adding too much complexity.

This tension becomes more acute when survey researchers define terms used everyday, terms such as 'you,' 'child,' 'poultry,' and so on. Defining terms that everyone already understands violates the maxim of manner (Grice, 1989). That maxim enjoins speakers to be brief, clear, and orderly, and to avoid unnecessary ambiguity and wordiness. Violations of the Gricean maxim tend to generate conversational implicature – an inference listeners work out to maintain the overarching assumption that the speaker is being cooperative. Defining everyday terms could suggest to respondents that the everyday terms in the survey questions are being used in some special or technical sense or that the definitions are intended for a subpopulation that needs them (e.g., non-native speakers). These implicatures may give rise more confusion than clarification.

Applied work in the field of computer-human interaction provides some empirical evidence for the confusions that can be caused by apparent violation of the Cooperative Principle. Young (1999) compared instruction descriptions generated according to Grice's CP to an "exhaustive plan," which gives the most detailed (and most redundant) instructions on every single step of the task, and to a "primitive plan," which describes only the lowest-level steps in the task. In other words, the three instructions differed sharply in the amount of information they contained. Subjects were asked to carry out a task described by the instructions within a computer simulation. Young (1999) found that subjects given the instructions produced according to the CP committed fewer errors and achieved more of their top-level goals than subjects who got either of the other two sets of

instructions. Specifically, the "exhaustive plan" led to the highest failure rate compared to the other two (Young, 1999). Giving more information than was necessary didn't improve performance.

A similar finding was reported by Gerber and her colleagues, who explored how to convey the notion of "residence" to respondents in an effort to improve the accuracy of their responses to roster questions like those used in the decennial census (Gerber, Wellens, & Keeley, 1996). Gerber and her coauthors found that providing definitional information about census rules on residence resulted in fewer correct answers for questions on simple and straightforward living situations. In two of five instances of straightforward living arrangements, the decreases were fairly substantial (15-18% decreases in correct answers). Gerber and her colleagues speculated that accuracy decreased because respondents regard the presentation of definitional rules which they already "know" as redundant; respondents may, as a result, reinterpret these rules in an effort to make sense of them (Gerber, Wellens, & Keeley, 1996).

The findings of Young (1999) and Gerber et al. (1996) raise the question of how much information to include in a definition or a set of instructions. This question is important since any definition can appear self-evident and redundant to those portions of the target population who share the same definition as the researcher. It will become a bigger issue especially as surveys are designed to accommodate populations that vary more in these linguistic backgrounds.

This study examines the potential costs of offering explicit definitions for everyday terms. It focuses on the effects of offering definitions on comprehension and data quality when the definitions fail to provide new information to respondents. This will often happen in practice since, for many respondents, their definitions of these terms will coincide with those of the survey. For these respondents, the definitions will seem unnecessary, violating the maxims of manner. The current study employs definitions that were designed not to provide new information.

The main hypotheses concern three dependent variables – responses to the questions, the response times, and respondents' inferences based on the definitions. When respondents are offered self-evident definitions for everyday terms, I predict that they will recognize the definitions aren't necessary; accordingly, they will become confused, try to work out an inference to explain why a definition was provided, and shape their responses based on the inference. I tested the following hypotheses:

1) Giving a definition to an everyday term leads to different survey responses from when no definition is offered. Specifically, the variance and covariance matrix for the items will differ.
2) Respondents who are given unneeded definitions will take longer on average to arrive at an answer than those who are not offered the definitions.
3) Respondents who are given definitions to everyday terms will infer that these terms are used in some technical sense or that the definitions are intended for some special population.

2. The Study

The experiment was one of several experiments embedded in a questionnaire administered via audio computer-assisted self-interviewing (ACASI). The questionnaire covered a range of topics, most of them political topics. The questionnaire included several experiments; most of them involved response order. We recruited 160 participants from the College Park, Greenbelt, and Silver Spring, MD area. We placed recruitment flyers at local libraries, advertised the study in local papers, and sent an e-mail invitation to the staff and graduate students at the University of Maryland. Since we desired a heterogeneous sample, we restricted the number of undergraduate students to less than 40 (one fourth of the targeted number of completes). Participants came to the Joint Program in Survey Methodology, where they first completed a 40-45 minute survey on a computer. The questionnaire was programmed in Blaise. Participants could listen to the questions (and response options) via earphones, read them displayed on the computer screen, or both. As I note below, the speed of the voice reading the questions was systematically varied. Participants indicated their answers by typing in the number corresponding to one of the answer options or by typing in text (in response to the open-ended questions). They then completed a paper-and-pencil questionnaire assessing the "Big Five" personality traits as well as their need for cognition. These questions took about 10 minutes. Participants were paid $25 upon completion of two questionnaires. The experiment ran from January 10, 2005, to February 28, 2005.

2.1 Experimental Manipulation

My experiment was placed in the middle part of the questionnaire. Participants were randomly assigned to one of the two experimental conditions. In one condition, redundant definitions for everyday terms were embedded in the question text and provided to

participants for all four key survey terms (poultry, fat, vegetable, and red meat). For example, here is the poultry item with its accompanying definition:

> "We will first ask you about how much poultry you eat. We define poultry as domestic fowl raised for meat. During the last 6 months, how much poultry did you typically consume?"

In the other condition, participants were asked the same questions with the same key terms. However, no definitions were given. An example of the wording is given below:

> "We are interested in studying Americans' consumption of poultry. We will first ask you about how much poultry you eat. During the last 6 months, how much poultry did you typically consume?"

The other experimental variable is related to the larger A-CASI study that included my experiment. We used synthesized voices that permitted us to systematically vary the rate of speech.[1] There were three speed conditions: slow speed, fast speed, and fast speed with pauses between the response options. Participants were randomly assigned to one of the three speed conditions. This randomization was independent of the random assignment to definition conditions. Although this manipulation is not directly relevant to my experiment, some of my analyses control for voice speed. Table 1 summarizes the number of participants in the two (definition or no definition) by three (slow speed, fast speed, or fast speed with pauses) experimental design.

Table 1. Number of Participants Assigned to Experimental Conditions

	Given Definition	No Definition	Total
Slow Speed	27	25	52
Fast Speed	22	30	52
Fast Speed with Pauses	26	30	56
Total	75	85	160

2.2 Target Questions

Respondents were asked about four food categories (poultry, fat, vegetable, and red meat). For each food category, respondents were asked about their typical consumption (e.g., "During the last 6 months, how much poultry did you typically consume?") and whether they tried to consume or avoid that food category (e.g., "Thinking about the food you eat, is poultry something you actively try to include in your diet, something you actively try to avoid, or something you do not think about either way?"). For some categories, the respondents were also asked to judge whether a specific food belonged to that food category (e.g., "Do you consider Cornish hens to be poultry?").

2.3 Follow-up Questions

I assessed respondents' inferences about the food categories with two follow-up questions. The first asked respondents whom they thought the study was intended for and the second asked whether respondents believed these terms were used in a technical sense or their ordinary sense.

3. Results

The analyses focus on three outcome variables – the interrelations among the responses, response times, and respondents' inferences. I begin my analyses with the effect of definitions on responses to the target questions.[2]

3.1 Effects on the Responses

Providing self-evident definitions to everyday terms breaches the maxim of manner. If respondents noticed the violation and used it to make an inference about the meaning of the questions, then their answers might be affected. I first compared the variance in responses to the four key survey questions when definitions were offered versus when they were not. Table 2 shows an apparent trend for smaller variances when definitions were offered than when they were not offered. *F*-tests showed that providing definitions significantly reduced the variance for two of the four target questions (consumption of poultry and consumption of red meat). The reduction in variance indicates only that responses were more uniform across respondents; without true values to compare the answers to, it is hard to tell whether the reduction in variance also represents a reduction in measurement error.

I also looked at whether the definitions helped respondents classify borderline instances of a category. I asked respondents about three borderline instances—whether eggs are poultry; whether potatoes are vegetables; and whether ham is red meat. I also asked

[1] The ACASI voice was a synthesized voice, generated by the AT&T Natural Voices® software, developed by AT&T Laboratories.

[2] Answers to the 14 target questions didn't vary by whether self-evident definitions were offered or not. Responses to one (out of 14) question were affected significantly by voice condition. The two experimental factors had significant interaction effects on responses to two other questions.

about more prototypical exemplars of the target categories – Cornish hen, broccoli, and steak – and nearly everyone classified these correctly with or without a definition. For the borderline instances, responses were somewhat evenly divided (see Table 3). The definitions didn't systematically move respondents either way. Furthermore, the definitions didn't significantly influence the speed with which respondents responded to the questions.

For each of the four key target terms, I also examined the relation of responses to the target question and several related questions. I fit multiple regression models with the target question as the dependent variable and the related questions as the predictors. If unnecessary definitions confuse respondents, their answers may be more error prone, reducing their correlation with the predictors. Thus, the focus of these analyses is the interaction effects between the definition and the other predictors.

Overall, four out of the 24 interactions between the predictors and the definition variable were significant, indicating that the definition altered the relation between the predictor and the target dietary consumption variable. However, of the four significant interactions, two are positive and two are negative. Thus, there didn't seem to be a consistent trend as to the size and direction of the significant interaction effects. Although giving definitions to everyday terms did have some effect on the underlying covariance matrix, these effects were neither consistent nor strong.

3.2 Response Times

I hypothesized that the total response time would be longer for respondents given definitions to the everyday terms than for those who didn't get definitions. We recorded the time from the moment the audio started to read the question until the moment when the respondent chose an answer. Thus, our measure of the response time encompasses the reading time for the question by the audio plus the time respondents took to answer the question. This measure poses a problem since the respondents' actual response time is confounded with the voice speed for the audio. The reaction times would almost necessarily be longer in the slow-speed voice condition than in the fast-speed voice conditions. Table 4 presents the results.

Results from a two-way ANOVA confirmed that voice condition made a big difference in the total response time (see Table 4). Attempting to tease apart the variation in response times caused by the experimental manipulation of the voice speed, I standardized the response times under each voice condition and analyzed these standardized times. Since response times are usually highly skewed (cf. Ratcliff, 1993), I replaced the standardized response times that were above three standard deviations with the value three standard deviations above the mean (cf. Ratcliff, 1993).

A one-way ANOVA on the total response time for this block of questions (with the factor being whether respondents were given a definition or not) revealed that respondents who were given the definitions turned out to be nonsignificantly faster (mean overall standardized response time=-0.030) than those who were not given a definition (mean=0.027) ($F(1,158)<1$, ns).

Additional analyses of the data showed that one respondent characteristic affected the response time – the need for cognition. According to Petty and Jarvis (1996), people with a high need for cognition (HNC) would normally process the questions more carefully than those with a low need for cognition (LNC respondents). Applying the conversational maxims requires cognitive effort; the HNC group is more likely to notice apparent violations of the conversational maxims and to draw implicatures based on apparent violations than their LNC counterparts. A study by McCabe and Brannon (2004) confirmed such a role for the need for cognition. They reported that HNC respondents applied the maxim of quantity to part-whole questions and displayed an attenuated correlation between the items, but not LNC respondents. Their finding suggested that the conversational norm to avoid redundancy is not automatically applied in the survey context; only those with a high need for cognition seemed to apply the maxim.

Based on this reasoning, I compared average response times for these two groups (HNC respondents vs. LNC respondents) with and without definitions. Figure 1 plots the mean standardized response times for the two groups.

The figure partially confirms the hypothesis. Within the HNC group, those who got a definition did take longer (mean standardized response time=-.19) than their counterparts who were not given a definition (mean=-.30). Even though the difference didn't approach significance ($F(1,71)<1$, ns), the direction was consistent with my hypothesis and replicated the finding by McCabe and Brannon (2004). On the other hand, the LNC group was slower on average (mean standardized response time=.19) than the HNC group

(mean=-.24). [3] The difference in response times between the two need for cognition groups was significant in the no definition condition ($F(1,73)=5.90$, $p<0.02$), but not in the definition condition ($F(1,83)=2.35$, $p=.13$). A three-way interaction between voice speed, definition condition, and need for cognition is marginally significant ($F(2,148)=2.22$, $p=.11$).

3.3 Inferences

I predicted that respondents who got the unnecessary definitions would conclude that the definitions were not intended for them or that the terms were not being used in their ordinary sense. I examined responses to the two follow-up questions to test for such inferences. One question specifically asked the respondents whom the survey was intended for. Respondents were somewhat more likely to think that the survey was intended for a special population rather than the general public when they got the unneeded definitions (7.0% of respondents given definitions vs. 3.6% of those without definitions), but the difference was not significant ($\chi^2=.72$, $p=.40$).

I also asked the respondents whether they believed the four key survey terms (poultry, vegetable, fat, and red meat) were used in their ordinary sense or in a special technical sense. All four of these follow-up items used the same scale, in which 1 meant the ordinary sense and 5 the technical sense. The mean responses are displayed in Figure 2. Higher numbers indicate ratings more in the direction of the technical sense.

It is clear from Figure 2 that when given definitions, respondents tended to regard the survey terms as used in a more technical way than their counterparts who were not shown definitions. The difference between the two groups of respondents was marginally significant for the term 'Fat' ($F(1,158)=3.04$, $p=0.08$). Thus, the study offers limited support for the notion that respondents found the definitions unnecessary and drew inferences to account for them.

4. Conclusions

This study investigated the effects of offering definitions for everyday terms in surveys. I predicted that defining terms that don't need definitions would violate the maxim of manner and create a linguistic anomaly; cooperative respondents recognize the anomaly and work out inferences to account for it. The results presented here lent at best weak support for the original hypotheses.

First, offering definitions to respondents did seem to influence their responses. Providing definitions reduced the variances of the responses, but the effects on regressions involving the key items are not easy to interpret and do not show a consistent picture. There is no clear evidence about whether offering redundant definitions improves or reduces data quality.

Second, some respondents did apparently process the definitions. Respondents with a high need for cognition were slower in answering the target questions when they were given definitions than when they were not (see Figure 1). Still, these differences in response times were not significant.

Third, respondents who got the definitions were somewhat more likely to infer that the survey was targeted at a special population and that the terms were used in a technical sense than those who didn't get the definitions. Again, though the direction was right, the trend was not significant.

The major limitation of this study lies in its small sample size (n=160), which inevitably reduces the power of the study. Even though offering definitions along with survey questions could promote more uniform interpretations of the questions, we need to be careful with the level of detail and the amount of information to be included in definitions. Respondents in this experiment noticed the self-evident definitions and seemed to have worked out inferences to account for them; the definitions also affected the variability in their responses. More systematic research should be carried out to investigate the effects of offering definitions on survey data quality. In the meanwhile, a thorough pretesting of survey concepts should go beyond detecting simple cognitive problems to investigate potential pragmatic issues related to offering definitions.

Acknowledgements

I would like to extend my personal thanks to Roger Tourangeau for his support, comments, and suggestions.

References

Belson, W. A. (1981). *The design and understanding of survey questions.* Aldershot: Gower.

Bloom, J. E., & Schober, M. F. (1999). Respondent

[3] The longer response time by the LNC group could be explained by the positive correlation between education and need for cognition. The LNC group contains significantly more respondents who had less than a college degree. Thus, the LNC respondents are probably slower readers than the HNC respondents.

cues that survey questions are in danger of being misunderstood. In *Proceedings of the American Statistical Association, Section on Survey Research Methods* (pp. 992-997). Alexandria, VA: ASA.

Clark, H. H., & Schober, M. F. (1992). Asking questions and influencing answers. In J. M. Tanur (Ed.), *Questions about questions* (pp. 15-48). New York: Russell Sage.

Conrad, F. G., & Schober, M. (2000). Clarifying question meaning in a household telephone survey. *Public Opinion Quarterly*, 64, 1-28.

Conrad, F. G., Couper, M. P., Tourangeau, R. & Peytchev, A. (In press). Use and non-use of clarification features in web surveys. *Journal of Official Statistics.*

Fowler, F. J. (1992). How unclear terms affect survey data. *Public Opinion Quarterly* 56, 218-231.

Fuller, W. A. (1987). *Measurement error models*. New York: Wiley.

Gerber, E. R., Wellens, T. R. & Keeley, C. (1996). Who lives here? The use of vignettes in household roster research. In *Proceedings of the Section on Survey Research Methods, American Statistical Association* (pp. 962-967). Alexandria, VA.

Grice, H. P. (1989). *Studies in the way of words*. Harvard University Press.

Groves, R. M., Fowler, F. J., Couper, M. P., Lepkowski, J. M., Singer, E., & Tourangeau, R. (2004). *Survey methodology*. Hoboken, NJ: Wiley.

Krosnick, J. A. (1991). Response strategies for coping with the cognitive demands of attitude measures in surveys. *Applied Cognitive Psychology* 5, 213-236.

Lind, L. H., Schober, M. F., & Conrad, F. G. (2001). Clarifying question meaning in a web-based survey. In *Proceedings of the American Statistical Association, Section on Survey Research Methods*. Alexandria, VA: American Statistical Association.

Martin, E. A., Campanelli, P. C., & Fay. R. E. (1991). An Application of Rasch Analysis to questionnaire design: Using vignettes to study the meaning of 'Work' in the Current Population Survey. *The Statistician,* 40, 265-276.

McCabe, A., & Brannon, L. (2004). Application of conversational norms to the interpretation of survey results as a function of participants' need for cognition. *The Journal of Psychology*, 138, 91-94.

Petty, R, E., & Jarvis, W. (1996). An individual differences perspective on assessing cognitive processes. In N. Schwarz & S. Sudman (Eds.), *Answering Questions: Methodology for Determining Cognitive and Communicative Processes in Survey Research* (pp. 221-258). San Francisco: Jossey-Bass Publishers

Ratcliff, R. (1993). Methods for dealing with response time outliers. *Psychological Bulletin,* 114, 510-532.

Schober, M. F., & Conrad, F. G. (1997). Does conversational interviewing reduce survey measurement error? *Public Opinion Quarterly*, 61, 576-602.

Schober, M. F., Conrad, F. G., & Bloom, J. E. (2000). Clarifying word meanings in computer-administered survey interviews. In L. R. Gleitman, & A. K. Joshi (Eds.) *Proceedings of the 22nd Annual Conference of the Cognitive Science Society* (pp. 447-452). Mahwah, NJ: Lawrence Erlbaum

Schober, M. F., Conrad, F. G., & Fricker, S. S. (1999). When and how should survey interviewers clarify question meaning? In *Proceedings of the American Statistical Association, Section on Survey Research Methods* (pp. 986-991). Alexandria, VA: ASA.

Schwarz, N., & Oyserman, D (2001). Asking questions about behavior: Cognition, communication, and questionnaire construction, *American Journal of Evaluation*, 22, 127–160.

Suessbrick, A. A., Schober, M. F., & Conrad, F. G. (2000). Different respondents interpret ordinary questions quite differently. In *Proceedings of the American Statistical Association, Section on Survey Research Methods* (pp. 907-912). Alexandria, VA: ASA.

Tourangeau, R., & Conrad, F. G. (2004). Everyday concepts and reporting errors. Paper Presented at the 59th Annual Conference of the American Association for Public Opinion Research. Phoenix, AZ.

Tourangeau, R., Rips, L., & Rasinski, K. (2000). *The psychology of survey response*. Cambridge University Press.

Young, M. (1999). Cooperative plan identification: Constructing concise and effective plan descriptions. In *Proceedings of the National Conference of the American Association for Artificial Intelligence*, Orlando, FL.

Table 2. Variances of Responses to Four Target Questions By Definition Condition

	Definitions offered	Definitions not offered	*F* values
Consumption of poultry	0.488	0.853	$F(74,84)=1.75$, $p=.01$
Consumption of fat	0.907	0.727	$F(84,74)=1.24$, $p=.16$
Consumption of vegetable	0.728	0.766	$F(74,84)=1.05$, $p=.41$
Consumption of red meat	0.916	1.333	$F(74,84)=1.46$, $p=.05$

Table 3. Percentage of "Yes" Responses to Borderline Instances by Definition Condition

	Given Definition (%)	No Definition (%)	χ^2 test results
Eggs	34.1	46.0	$\chi^2=2.31$p=.13
Potatoes	63.5	61.3	$\chi^2=.08$ $p=.77$
Ham	47.6	44.0	$\chi^2=.21$ $p=.64$

Table 4. ANOVA Results on Total Response Time

	No Definition	Given Definition	ANOVA Results	
Slow speed	210.4	213.6	Definition	$F(1,154)=0.14$, n.s
Fast speed	177.3	178.2	Voice Condition	$F(2,154)=4.96$, p=0.01
Fast speed with pauses	199.6	185.5	Interaction	$F(2,154)=0.39$, ns

Figure 1. Mean Standardized Response Times for HNC and LNC Respondents

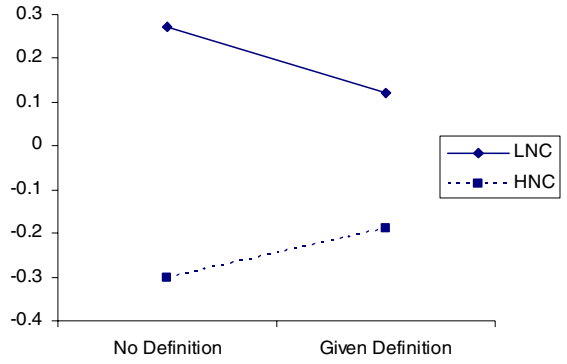

Figure 2. Mean Ratings of Whether Survey Terms Are Used in Technical Sense by Definition Condition