APPLIED COGNITIVE PSYCHOLOGY
Appl. Cognit. Psychol. **17**: 215–224 (2003)
Published online in Wiley InterScience 28 November 2002
(www.interscience.wiley.com) DOI: 10.1002/acp.850

Cognitive Procedures for Correcting Proxy-response Biases in Surveys

ALEXANDER TODOROV*

New York University, USA

SUMMARY

To reduce survey costs, major surveys rely on self- and proxy-responses. The use of proxies can reduce data quality introducing biases in the survey estimates. This paper identifies one source of systematic differences between self- and proxy-reports: proxies' higher reliance on inferences. Using data from the National Health Interview Survey on Disability (NHIS-D), proxy-response biases were modelled by independently collected measures of cognitive inferences. Conditional likelihood judgements about a number of disabilities (e.g. likelihood that a person has a disability given another disability) predicted the conditional disability reports for proxy- but not for self-respondents (e.g. the proportion of respondents who reported difficulty learning after reporting difficulty communicating). A model of self/proxy differences was estimated on data from the 1994 NHIS-D and tested against 1995 data. The correlation between predicted and actual differences was 0.76. The correlation between predicted and actual proxy-reports was 0.95. Such research can be used to estimate and correct for systematic proxy-response biases. Copyright © 2002 John Wiley & Sons, Ltd.

Most major national surveys rely on both self- and proxy-reports. For example, the National Health Interview Survey (NHIS) is an annual household survey. Households are sampled and household members who are present at the time of the interview provide responses for household members who are absent—proxy-responses. This survey strategy dramatically reduces the survey cost (White and Massey, 1981), but may also reduce the data quality. In fact, a number of studies have shown consistent differences between self- and proxy-reports (Bassett *et al.*, 1990; Epstein *et al.*, 1989; Kovar and Wright, 1973; Mathiowetz and Groves, 1985; Rothman *et al.*, 1991). For instance, Todorov and Kirchner (2000) have shown that proxy- and self-reports of disabilities in the Disability Supplement of the NHIS (NHIS-D) from 1994 and 1995 differed reliably and systematically. Although both self- and proxy-reports are susceptible to reporting biases, self-reports tend to be more accurate (Loftus *et al.*, 1992), and even can be a stronger predictor of morbidity than physician evaluated morbidity (Ferraro and Farmer, 1999).

The main problem of studying differences between self- and proxy-responses in the context of surveys is that respondents are not randomly allocated to self- or proxy-status (Moore, 1988). Because respondent status (self versus proxy) is confounded with many

*Correspondence to: Alexander Todorov, Department of Psychology, New York University, 6 Washington Place, 7th Floor, New York, NY 10003, USA. E-mail: att206@nyu.edu

background variables, differences between self- and proxy-responses can reflect demographic confounds rather than proxy-response biases. The standard strategy for disentangling these sources of differences is to introduce statistical control. For instance, if differences between self- and proxy-responses remain reliable after adjusting for demographic variables such as age and gender, one can argue that these differences can be attributed to proxy-response biases (e.g. Todorov and Kirchner, 2000).

An additional strategy is to start with specific assumptions about the nature of the differences between self- and proxy-reports and to model these differences based on the assumptions. This paper pursues the latter approach. From a cognitive point of view, there are two major differences between self- and proxy-respondents: (a) the information available for self- and other-judgements is different; and (b) the cognitive strategies of forming these judgements can be different. Applying cognitive theories to the study of context effects on survey responses has led to a better understanding and prediction of these effects (Schwarz, 1999; Schwarz and Sudman, 1996; Sudman *et al.*, 1996; Tanur, 1992; Todorov, 2000a,b; Tourangeau and Rasinski, 1988). However, relatively little work has been done in applying such theories to predicting and understanding differences between self- and proxy-responses.

When people have to make a judgement about themselves, they can recall not only a variety of relevant behaviours but also feelings and beliefs. For example, a disability can have different expressions ranging from observable behavioural difficulties to unobservable subjective pain. The differential availability of information may lead to different judgements about what constitutes a disabling difficulty for oneself and what for another person. Moreover, the set of relevant behaviours is not accessible to the same extent to self- and proxy-respondents. One's behaviours are more salient and more accessible in memory than another's behaviours (Ross and Sicoly, 1979; Schwarz and Wellens, 1997). In fact, Todorov and Kirchner (2000) have shown that informational measures of disabilities—how observable are the disabilities and to what extent they are detectable in social interactions—accounted for 70% of the variance of differences between self- and proxy-reports of disabilities in the NHIS-D.

Not only is the available information for self- and other-judgements different, but also the cognitive strategies of forming these judgements can differ (Bickart *et al.*, 1994; Menon *et al.*, 1995). Making a judgement about another person may be based on inferences to a greater extent than making a judgement about oneself (Schwarz and Wellens, 1997). For example, if a person is asked how many times she visited a doctor for the last 12 months, she can try to recall all specific visits. However, if the person is asked the same question about a family member, she is likely to estimate the number of visits rather than try to recall them. In general, when proxy-respondents do not have sufficient factual information about a specific question, they may try to infer the response by supplying assumptions about the question. Thus, given that proxy-respondents have less information than self-respondents, they should be more likely to rely on inferences when responding to survey questions.

One source of inferences is the implicit theories that people hold about the domain of questions (Jennings *et al.*, 1982; Nisbett and Ross, 1991; Nisbett and Wilson, 1977). Consider a proxy-respondent who reports that a household member has difficulty communicating. Later the proxy-respondent is asked whether or not the household member has difficulty learning. People believe that difficulty learning and difficulty communicating are highly related and, correspondingly, proxy-respondents should over-report problems with difficulty learning given difficulty communicating. Indeed, in the NHIS-D among respondents who reported difficulty communicating, 49.4% of proxy-respondents reported that the reference person also had difficulty learning relative to 21.7% of self-respondents. On the other hand, problems with dizziness and difficulty communicating do not seem related. Under these conditions, proxy-respondents should under-report problems with dizziness given difficulty communicating. That was the case in the NHIS-D. Among respondents who reported difficulty communicating, 7.8% of proxy-respondents reported that the reference person also had problems with dizziness relative to 17.7% of self-respondents.

The objective of this paper is to demonstrate that differences between self- and proxy-reports are systematically related to measures of cognitive inferences. If proxy-respondents rely on inferences and, specifically, on lay theories about the relation between different disabilities, they should overestimate the likelihood that a reference person has a disability (e.g. difficulty learning) given that they previously reported a related disability (e.g. difficulty communicating). On the other hand, proxy-respondents should under-estimate the likelihood that a reference person has a disability if they previously reported an unrelated disability. To collect measures of relatedness of disabilities, participants were asked to estimate the likelihood of having a disability (e.g. difficulty learning) given another disability (e.g. difficulty communicating). For example, the mean likelihood judgement of having difficulty learning given difficulty communicating was 7.07 (on a 10-point scale). The likelihood judgement of having problems with dizziness given difficulty communicating was 2.37.

The likelihood judgements were used to predict the conditional self- and proxy-reports of disabilities in the NHIS-D.[1] These reports can be considered as conditional probabilities of disabilities given a previously reported disability. The NHIS-D is a nationally representative survey and provides the national estimates of the prevalence of disabilities. About 38% of the reports in this survey were proxy-reports (not including proxy-reports for persons younger than 18 years). If self- and proxy-reports differ systematically, the national estimates of disabilities will be biased. However, if the systematic sources of bias can be identified and measured, the estimates can be adjusted accordingly. Todorov and Kirchner (2000) have already identified one cognitive bias related to the different information available to self- and proxy-respondents. This paper attempts to identify another bias related to proxies' higher reliance on inferences.

The paper reports three sets of analyses. The purpose of the first set was to show that measures of inferences are a unique predictor of proxy-reports but not of self-reports. To the extent that conditional proxy-reports are grounded in reality, they should correlate highly with conditional self-reports. However, an additional source of proxy-reports is inferences based on lay theories linking different disabilities. Measures of such lay theories—the likelihood judgements—should predict conditional proxy-reports even after adjustment for the correlation with conditional self-reports. In contrast, such measures should not predict conditional self-reports after adjustment for the correlation with conditional proxy-reports.

In the second set of analyses, the actual differences between proxy- and self-reports were predicted as a function of the likelihood judgements. The NHIS-D was conducted in two consecutive years, 1994 and 1995. In the final set of analyses, the coefficients of a regression model predicting the self/proxy differences as a function of the

[1]All analyses and interpretations are the sole responsibility of the author and not of the National Center for Health Statistics.

likelihood judgements were estimated on the data from 1994 and tested against the data in 1995.

METHOD

Data sources

Conditional reports of disabilities

The disabilities were measured in the Disability Supplement (DS) of the National Health Interview Survey (NHIS-D), 1994 and 1995 (National Center for Health Statistics, 1998a,b). The NHIS is a representative continuing nationwide household survey. Data are collected each week on a probability sample. Personnel of the US Bureau of the Census conduct the face-to-face interviews. The DS was administered after the core questions, which are asked every year in the NHIS.

Because only proxy-responses are obtained for respondents below 18 years old, the present analyses were limited to respondents who were 18 years of age or older. The sample size for that age group was 77,437 in 1994 and 67,570 in 1995 (145,007 combined). Respondents with unknown respondent status (8.8%, $N = 12{,}750$) were excluded from the analyses, leaving a sample of 132,257 respondents, 81,840 self-respondents and 50,417 proxy-respondents.

The DS was divided into several sections. The complete questionnaire forms can be found in Adams and Marano (1995). Analyses were performed on sections of questions covering sensory impairments, functional limitations, and mental health. Ten disabilities were measured in the sensory impairments section in the following order: difficulty seeing; difficulty hearing; difficulty communicating; difficulty understanding; difficulty learning; problems with dizziness; problems with balance; ringing, roaring, or buzzing in the ears; problems with sense of smell; and problems with sense of taste. The functional limitations section measured eight disabilities in the following order: difficulty lifting; difficulty walking up 10 steps; difficulty walking a quarter of a mile; difficulty standing for about 20 minutes; difficulty bending down; difficulty reaching up or reaching out; difficulty using fingers; and difficulty holding a pen or pencil. The mental health section measured seven disabilities in the following order: frequent depression; troubles with making or keeping friendships; problems getting along with other people; problems concentrating; problems with coping with day-to-day stresses; frequent confusion; and phobias.

The data for the analysis were conditional proportions of reporting a disability given a previously reported disability (that is, in response to an earlier question) for self- and proxy-respondents. For example, among self-respondents who reported difficulty seeing 20.8% reported difficulty hearing. This conditional proportion was 25.3% for proxy-respondents. Ninety-four conditional proportions were calculated from the survey data: 45 conditional proportions for sensory impairments, 28 proportions for functional limitations, and 21 proportions for mental health problems. The exact formula for the number of conditional proportions is: $n^*(n-1)/2$, where n is the number of disabilities in the respective section. For example, in the mental health section six mental health problems were conditioned on the report of frequent depression, five were conditioned on the report of troubles with making or keeping friendships, four were conditioned on the report of problems getting along with other people, three were conditioned on the report of problems concentrating, two were conditioned on the report of problems with coping with day-to-day stresses, and finally the report of phobias was conditioned on the report of frequent confusion.

Likelihood judgements

Thirty undergraduate students from the Department of Psychology at NYU participated in the study for a partial course credit. The disabilities were described in the same way as in the NHIS questionnaire. Each participant made 94 conditional judgements corresponding to the 94 conditional proportions described above. The judgements were made on an 11-point scale, ranging from 0 (*not at all likely*) to 10 (*extremely likely*). The order of the conditional judgements followed the order of the disability questions in the NHIS. For instance, the first disability question was about 'difficulty seeing'. Participants were told to imagine that a person has 'difficulty seeing' and were asked how likely is that this person has 'difficulty hearing'—the second disability question. They were also asked to make the same conditional judgement about the remaining eight sensory disabilities. Participants made 45 conditional judgements about sensory impairments, 28 judgements about functional limitations, and 21 judgements about mental health problems. The participants' judgements were highly consistent. Cronbach's alpha was 0.98 measured at the level of the judgements using as variables the participants. The mean score of the participants' judgements was used as a predictor of the self- and proxy-conditional reports of disabilities.

Preliminary analyses

The likelihood judgements and the self- and proxy-reports were positively skewed and scatter plots indicated that the relationship between judgements and reports was not completely linear. In order to approximate linearity, the self- and proxy-reports were logarithmically transformed. It should be noted that all analyses reported below were run on the untransformed variables and the findings were consistent with the findings of the analyses of the transformed variables. Relevant statistics from both analyses (on transformed and untransformed data) are reported below.

Because several disability reports could be conditioned on a single disability report, this introduced possible order dependencies in the data or auto-correlation effects. To control for such effects, indicator dummy variables were created to code for the disability on which the reports were conditioned. Reports were conditioned on 22 disabilities (nine for sensory impairments, seven for functional limitations, and six for mental health problems) and, correspondingly, 21 dummy variables were created. The main analyses reported below are regression analyses in which the self/proxy reports were regressed on the likelihood judgements and the dummy variables. One problem with the regression model which used 21 dummy variables is that within each type of disability one of the dummy variables coded for only one observation. For example, the last two questions in the sensory impairments sections were about problems with smell and taste. Thus, there was only one proportion conditional on the reports of problems with smell (e.g. the probability of having problems with taste given problems with smell). In the case of these three dummy variables (that coded for single observations), the regression model amounts to fitting a regression line to a single observation. Correspondingly, these observations were highly influential. In order to avoid this problem, each of the three single observations was added to the immediately preceding group of observations, that is, to the two reports of disabilities conditioned on the disability reported before the last two questions. This reduced model with 18 dummy variables had a better adjusted R^2 than the model with 21 variables and practically did not change the regression coefficients for the likelihood judgements and self (proxy) reports. Thus, all analyses were based on the reduced model.

RESULTS

The likelihood judgements correlated with both self- ($r = 0.49$, $p < 0.01$) and proxy-conditional reports ($r = 0.61$, $p < 0.01$). Because self- and proxy-reports were highly correlated ($r = 0.88$, $p < 0.01$), more important are the partial correlation analyses that control for the shared variance between self- and proxy-reports. As expected, the proxy-reports remained reliably correlated with the judgements after controlling for the correlation between self- and proxy-reports ($r = 0.44$, $p < 0.01$). In contrast, after controlling for the latter correlation, self-reports were not significantly correlated with the judgements ($r = -0.13$). The difference between these two correlations was significant, $z = 3.99$, $p < 0.001$.

Predicting self- and proxy-reports

Two regression models were fitted. In model 1, proxy-reports were regressed on likelihood judgements, self-reports, and the 18 indicator variables described above. In model 2, self-reports were regressed on likelihood judgements, proxy-reports, and the 18 indicator variables. As shown in Table 1, the likelihood judgements were a significant predictor of proxy-reports (standardized $\beta = 0.36$), $t(73) = 7.20$, $p < 0.001$, but not of self-reports (standardized $\beta = -0.12$), $t(73) = -1.27$, $p > 0.21$. These findings were consistent with analyses of the untransformed data. The likelihood judgements were not a significant predictor of self-reports, $t < 1.0$, but they did predict proxy-reports, $t(73) = 4.14$, $p < 0.001$.

The fitted regression models may seem complicated because of the many indicator variables. However, these models (a) have a simple interpretation and (b) were the only models that dealt with autocorrelation effects. The models fit a regression line with a constant slope for each disability on which reports were conditioned but with a different intercept. The coefficient for the indicator variable coding the respective disability is simply added to the initial intercept of the models. Second, only in these models the Durbin–Watson statistic, which measures autocorrelation effects, was not significant. In the simpler models—regressing proxy (self) reports on self (proxy) reports and judgements; and regressing proxy (self) reports on self (proxy) reports, judgements, and two indicator variables coding for the type of disability—there was systematic patterning of the regression residuals and the Durbin–Watson statistic was significant.

Predicting differences between self- and proxy-reports

The above analyses showed that the likelihood judgements were a significant predictor of proxy- but not of self-reports. In the second set of analyses, the likelihood judgements

Table 1. Unstandardized regression coefficients of likelihood judgements, self-reports, proxy-reports and explained variance in regression analyses

Dependent variable	Judgements	Proxy-reports	Self-reports	Variance (R^2)
Proxy-reports	0.14*		0.57*	93%
Self-reports	−0.04	0.88*		85%

*$p < 0.01$.
Note: The analyses were performed on the logarithmic transformations of the self- and proxy-reports.

were used to predict the actual differences between self- and proxy-reports. A measure of these differences is the ratio of proxy- and self-reports. This ratio is greater than 1 for proxies' over-reporting and smaller than 1 for proxies' under-reporting. The ratio was regressed on the likelihood judgements and the 18 indicator variables. Although in this model the judgements were a significant predictor of the ratio, $\beta = 0.07$ (standardized $\beta = 0.39$), $t(74) = 3.22$, $p < 0.002$, the scatter plot of the regression residuals against the predicted values showed substantial heteroscedasticity. To stabilize the variance, the likelihood judgements were divided by the self-reports and both ratios of proxy- to self-reports and judgements to self-reports were logarithmically transformed. Then the transformed proxy/self ratio was regressed on the transformed judgements/self ratio and the indicator variables. This regression model accounted for 65.4% of the variance and satisfied the regression assumptions. The judgements/self ratio was a significant predictor of the proxy/self ratio, $\beta = 0.47$ (standardized $\beta = 0.73$), $t(74) = 7.01$, $p < 0.001$.

Cross-validation of regression models

The NHIS-D was conducted in two consecutive years, 1994 and 1995. This allows for a cross-validation of the regression model described above. The coefficients of the regression model were estimated from the 1994 data and the model was tested against the 1995 data. The correlation between the predicted and actual untransformed proxy/self ratios was 0.77, $p < 0.001$. Knowing the predicted proxy/self ratio and the self-reports for 1995, one can predict the proxy-reports for 1995. This corresponds to a situation in which one does not use any proxy-respondents but predicts proxy-responses from self-responses and a model of proxy-reporting bias. The correlation between the actual and predicted proxy-reports for 1995 was 0.95, $p < 0.001$.

To identify the contribution of the likelihood judgements to the prediction of proxy-reports, two additional models were cross-validated and compared. In the simpler model, the logarithmically transformed proxy-reports were predicted by the transformed self-reports and the indicator variables. In this model, which did not include the likelihood judgements, the correlation between the predicted and actual values was 0.93 accounting for 86.5% of the variance. In the second model, which included the likelihood judgements as a predictor, the correlation between the predicted and actual values was 0.96 accounting for 91.2% of the variance. Although the increment in the accounted variance (4.7%) could seem small, the effect was highly significant, $F(1, 73) = 40.12$, $p < 0.001$. Further, the model which did not include the likelihood judgements accounted for 86.5% of the variance leaving only 13.5% unexplained variance. Thus, the inclusion of the likelihood judgements in the model reduced the unexplained variance with 34.8%.

DISCUSSION

If proxy-reports are systematically different from self-reports and one of the sources of these differences is the proxies' higher reliance on inferences, then measures of such inferences should predict proxy-reports. Specifically, likelihood judgements should be a significant predictor of proxy-reports when the analysis controls for self-reports, but they should not predict self-reports when the analysis controls for proxy-reports. Both the partial correlation analyses and the regression analyses confirmed this hypothesis. Lay judgements of the relatedness of disabilities predicted proxy-reports of disabilities measured in a nationally representative survey but did not predict self-reports. These judgements also predicted the differences between self- and proxy-reports. The higher the judged relatedness of disabilities, the more likely proxies' over-reporting of disabilities. The lower the judged relatedness of disabilities, the more likely proxies' under-reporting of disabilities.

These findings suggest that proxy-reports are systematically biased. When respondents are asked to report about other people but do not have sufficient information, they appear to rely on inferences grounded in lay theories about the domain of questions. In the case of disabilities, respondents rely on theories about how disabilities are related to each other. This will lead to over-reporting of disabilities seemingly related to a previously reported disability and to under-reporting of disabilities seemingly unrelated to the previously reported disability.

The conditional judgements used in this study are measures of inferences, which are based on lay theories. It should be noted that (a) the only information on which the judgements were based was a pair of disabilities and (b) no specific evidence pertaining to the target person was presented. Compare this impoverished judgemental situation with the situation of a proxy-respondent who makes a judgement about a family member and, presumably, has relevant information pertaining to the specific survey question. Nevertheless, lay conditional judgements uniquely predicted conditional proxy-reports of disabilities.

Self- and proxy-reports differ in two fundamental aspects: the amount of information available to respond to a survey question and the cognitive strategies of generating a response. Todorov and Kirchner (2000) showed that lay judgements of informational properties of disabilities such as observability and impact on social interaction predicted differences between (unconditional) self- and proxy-reports of disabilities in the NHIS-D. In fact, these judgements accounted for 70% of the variance of the self/proxy differences. The less observable a disability, the larger the differences between self- and proxy-reports of the disability. The present study demonstrated that lay conditional judgements of the likelihood of a disability given another disability uniquely predicted conditional proxy-reports of disabilities. The findings from these studies are related. For example, proxy-respondents should be more likely to rely on inferences when the available information (e.g. observations of the target person) is insufficient or ambiguous. Research on the proxies' reliance on inferences suggests that this is one of the cognitive mechanisms underlying self/proxy differences which derive from differential availability of information.

The standard way of approaching differences between self- and proxy-respondents in field surveys is to introduce statistical control for confounding variables. If the differences remain reliable after statistical adjustment, the researcher may argue that these differences are due to proxy-response biases. Although statistical control is the first necessary step to address self/proxy differences, statistical control specifies neither the source nor the magnitude of the proxy-response bias. The main advantage of the procedures introduced in Todorov and Kirchner (2000) and in the current studies is that they are designed as a direct measure of sources of systematic proxy-response bias.

Two things should be noted for this strategy of studying self/proxy differences. First, the analysis is done at the level of group proportions and not at the level of individual respondents. Although this might be considered a weakness, in national surveys researchers are usually interested in population estimates which are derived from proportions. Second, the variance of self/proxy differences accounted for by such informational and inference measures can be specifically attributed to systematic proxy-response biases.

If these procedures prove to be reliable and valid, they can be used to estimate the unique variance attributable to proxy-response bias. For instance, the total variance of self/proxy differences can be partitioned into variance due to real differences, variance due to proxy-response biases, and error variance. Statistical analyses controlling for demographic differences can estimate the variance due to true differences, whereas analyses based on procedures as the ones in the current study can estimate the variance due to systematic proxy-response bias.

These procedures seem to work in the health domain. For example, Todorov and Kirchner (2000) estimated that the use of proxy-reports in the NHIS-D underestimated the number of people with disabilities in the USA with 1,609,000. However, before such procedures are accepted, they should be validated in other survey domains (e.g. political and economic surveys). Further, one should use representative samples to estimate inferences or lay theories related to self/proxy differences. In the present study, these inferences were estimated using a sample of college undergraduates clearly different in their demographic characteristics from the nationally representative sample. Although the estimated inferences successfully predicted proxy-responses, it is likely that the use of samples matched on their characteristics will increase the accuracy of prediction.

Cognitive rating procedures have been applied successfully not only to the prediction of differences between self- and proxy-reports in nationally representative surveys but also to the prediction of the size of context effects on survey reports (Todorov, 2000a). These procedures are grounded in cognitive theories and can be a useful supplement to recently developed survey techniques such as cognitive interviewing (DeMaio and Rothgeb, 1996). Most importantly, these procedures introduce a theory-based modelling approach to the study of proxy-response biases.

ACKNOWLEDGEMENT

I thank Dawn Porter for her comments on previous versions of this paper.

REFERENCES

Adams PF, Marano MA. 1995. Current estimates from the National Health Interview Survey, 1994. *Vital and Health Statistics* **10**(193).

Bassett SS, Magaziner J, Hebel JR. 1990. Reliability of proxy response on mental health indices for aged, community-dwelling women. *Psychology and Aging* **5**: 127–132.

Bickart B, Menon G, Schwarz N, Blair J. 1994. The use of anchoring strategies in constructing proxy reports of attitudes. *International Journal of Public Opinion Research* **6**: 375–379.

DeMaio TJ, Rothgeb JM. 1996. Cognitive interviewing techniques: in the lab and in the field. In *Answering Questions: Methodology for Determining Cognitive and Communicative Processes in Survey Research*, Schwarz N, Sudman S (eds). Jossey-Bass: San Francisco, CA; 177–195.

Epstein AM, Hall JA, Tognetti J, Son LH, Conant L. 1989. Using proxies to evaluate quality of life. *Medical Care* **27**(Suppl. 3): 91–98.

Ferraro KF, Farmer MF. 1999. Utility of health data from social surveys: is there a gold standard for measuring morbidity? *American Sociological Review* **64**: 303–315.

Jennings DL, Amabile TM, Ross L. 1982. Informal covariation assessment: data-based versus theory-based judgements. In *Judgements under Uncertainty: Heuristics and Biases*, Kahneman D, Slovic P, Tversky A (eds). Cambridge University Press: New York; 211–230.

Kovar MG, Wright RA. 1973. An experiment with alternate respondent rules in the National Health Interview Survey. *Proceedings of the Social Statistics Section*, American Statistical Association: Washington, DC; 311–316.

Loftus EF, Smith KD, Klinger MR, Fiedler J. 1992. Memory and mismemory for health events. In *Questions about Questions: Inquiries into the Cognitive Bases of Surveys*, Tanur JM (ed.). Russell Sage: New York; 102–137.

Mathiowetz NA, Groves RM. 1985. The effects of respondent rules on health survey reports. *American Journal of Public Health* **75**: 639–644.

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.

Moore JC. 1988. Self/proxy response status and survey response quality: a review of the literature. *Journal of Official Statistics* **4**: 155–172.

National Center for Health Statistics. 1998a. *Data File Documentation: National Health Interview Survey on Disability, Phase 1 and Phase 2, 1994* (machine readable data file and documentation, CD-ROM Series 10, No. 8A): Hyattsville, MD.

National Center for Health Statistics. 1998b. *Data File Documentation: National Health Interview Survey on Disability, Phase 1 and Phase 2, 1995* (machine readable data file and documentation, CD-ROM Series 10, No. 10A): Hyattsville, MD.

Nisbett RE, Ross L. 1991. *Human Inference: Strategies and Shortcomings of Social Judgement*. Prentice Hall: Englewood Cliffs, NJ.

Nisbett RE, Wilson TD. 1977. Telling more than we can know: verbal reports on mental processes. *Psychological Review* **84**: 231–259.

Ross M, Sicoly F. 1979. Egocentric biases in availability and attribution. *Journal of Personality and Social Psychology* **37**: 322–336.

Rothman ML, Hedrick SC, Bulcroft KA, Hickam DH, Rubenstein LZ. 1991. The validity of proxy-generated scores as measures of patient health status. *Medical Care* **29**: 115–124.

Schwarz N. 1999. Self-reports: how the questions shape the answers. *American Psychologist* **54**: 93–105.

Schwarz N, Sudman S (eds). 1996. *Answering Questions: Methodology for Determining Cognitive and Communicative Processes in Survey Research*. Jossey-Bass: San Francisco, CA.

Schwarz N, Wellens T. 1997. Cognitive dynamics of proxy responding: the diverging perspectives of actors and observers. *Journal of Official Statistics* **13**: 159–179.

Sudman S, Bradburn NM, Schwarz N. 1996. *Thinking about Answers: The Application of Cognitive Processes to Survey Methodology.* Jossey-Bass: San Francisco, CA.

Tanur JM (ed.). 1992. *Questions about Questions: Inquiries into the Cognitive Bases of Surveys*. Russell Sage: New York.

Todorov A. 2000a. The accessibility and applicability of knowledge: predicting context effects in national surveys. *Public Opinion Quarterly* **64**: 429–451.

Todorov A. 2000b. Context effects in national health surveys: effects of preceding questions on reporting serious difficulty seeing and legal blindness. *Public Opinion Quarterly* **64**: 65–76.

Todorov A, Kirchner C. 2000. Bias in proxies' reports of disability: data from the National Health Interview Survey on Disability. *American Journal of Public Health* **90**: 1248–1253.

Tourangeau R, Rasinski KA. 1988. Cognitive processes underlying context effects in attitude measurement. *Psychological Bulletin* **103**: 299–314.

White AA, Massey JT. 1981. Selective reduction of proxy response bias in a household interview survey. *Proceedings of the Social Statistics Section*. American Statistical Association: Washington, DC; 211–216.

