

Cognitive Dynamics of Proxy Responding: The Diverging Perspectives of Actors and Observers

Norbert Schwarz and Tracy Wellens

Psychological research demonstrated that people tend to explain their own behavior in terms of situational influences, but the behavior of others in terms of their dispositions (personality or likes and dislikes). We apply this literature to the cognitive processes underlying self and proxy reports. Several experiments demonstrate that proxy reports on related issues show a higher internal consistency than self reports, reflecting that related judgments are derived from the same dispositional information. Hence, the internal consistency of proxy reports should not be taken as evidence for their accuracy. Moreover, proxy reports underestimate the variability of behavior over time. Accordingly, proxy reports and self reports show low convergence for short and recent reference periods, for which the actor can draw on episodic information in providing self report. As the actor's access to episodic information decreases, the actor has to rely on dispositional information as well. Hence, convergence increases for distant events and lengthy reference periods. Implications for future research, respondent selection rules, and questionnaire construction are discussed.

Key words: Self reports; proxy reports; behavioral reports; autobiographical memory; response effects; response errors.

1. Introduction

Many surveys conducted by government agencies ask respondents to provide information about themselves as well as about other members of their household. For example, in the U.S. as well as in other countries, labor force surveys typically require one household member to report about the labor force participation of all household members. Other surveys (e.g., the Health Interview Survey and National Crime Survey in the U.S.) allow proxy responding when the target respondent is not available. Survey researchers have generally assumed that self reports are more complete than proxy reports, and by and large the available literature supports this assumption (see Moore 1988; Sudman, Bradburn, and Schwarz 1996, Chapter 10, for reviews). However, most of the available studies are seriously flawed, due to possible selection biases. Specifically, proxy respondents have typically been used when the target respondent is not available, resulting in a self-assignment to self and proxy conditions which undermines systematic comparisons.

A study that avoided selection biases was conducted by the Survey Research Laboratory

Acknowledgment: The reported research was conducted under a Joint Statistical Agreement between the Zentrum für Umfragen, Methoden und Analysen, ZUMA, Mannheim, Germany and the U.S. Bureau of the Census during the authors' tenure at ZUMA. Norbert Schwarz is now with the Institute for Social Research, University of Michigan and Tracy Wellens with the Center for Survey Methods Research at the U.S. Census Bureau. Address correspondence to Norbert Schwarz, Institute for Social Research, University of Michigan, Ann Arbor, MI 48106-1248 U.S.A.; e-mail: NSchwarz@umich.edu

of the University of Illinois (see Menon, Bickart, Sudman, and Blair 1995; Sudman, Bickart, Blair, and Menon 1994, for reports). Telephone interviews with 200 pairs of partners in the same household in Champaign County, Illinois, allowed the collection of a wide variety of attitudinal and behavioral reports and the systematic assessment of response convergence under self and proxy conditions. The results showed modest agreement between self and proxy reports (Sudman et al. 1994). For topics that were frequently discussed in the household, the behavioral frequency reports provided by self and proxy respondents correlated $r = .49$ under telephone interview conditions and $r = .32$ under face to face conditions. For topics that were less frequently discussed, the correlations averaged $r = .41$ under telephone and $r = .20$ under face to face conditions. Similarly, self and proxy respondents' behavioral reports showed a convergence of $r = .47$ (telephone interview conditions) and $r = .60$ (face to face conditions) for frequently discussed items, and of $r = .38$ and $.49$, respectively, for less frequently discussed items. Although the convergence of responses varied considerably across topics, these findings indicate that the convergence of self and proxy reports is far from perfect, even under conditions where one would expect the proxy respondent to know about the partner's behavior, due to frequent discussion of the respective activity or attitude topic in the household. This conclusion is further supported by the observation that the convergence of self and proxy reports of behaviors averaged $r = .62$ and $.78$ under telephone and face to face conditions, respectively, even under conditions where self and proxy respondents reported a high degree of joint participation in the respective activity.

These findings indicate that the convergence of self and proxy reports for a specified reference period is moderate, even under the theoretically optimal condition of joint participation in the activity. Given the widespread use of proxy respondents in large scale surveys, it is therefore important to explore the cognitive processes that underlie this divergence of self and proxy reports. To do so, we draw on the extensive literature in social cognition that is devoted to examining how people store, integrate and utilize information about self and others and bring this literature to bear on self and proxy reports in surveys.

2. The Present Research

We provide a selective review of the relevant literature and derive hypotheses which we tested in laboratory experiments and two experimental surveys of couples, conducted under a Joint Statistical Agreement with the U.S. Census Bureau. At the outset, we note that our goal was the development of an appropriate *conceptual* framework and the collection of data that allow targeted tests of crucial theoretical hypotheses. Central to the experimental strategy used is the random assignment of respondents to experimental conditions, not the representativeness of the sample. This design decision reflects that the key issue at the present stage of research is the identification of basic cognitive processes under controlled conditions.

Community Couples Survey. Respondents of several community surveys in the Mannheim, Germany, area were asked to volunteer for the couples survey. Cooperation was relatively low and the screening of some 2,000 respondents resulted in a list of 312 volunteer couples, of whom 271 completed the survey. The couples completed

self-administered questionnaires in their homes, while the interviewer was waiting. This procedure was chosen to ensure that each member of the couple completed the questionnaire without consulting with his/her partner.

The goal of recruiting a heterogeneous sample has been achieved, as a short description of demographic data may illustrate. The male respondents had a mean age of 43 years (range: 17 to 77 years) and the female respondents of 41 years (range: 21 to 80 years). The couples had been living together for an average of 16 years (range: 0 to 62 years) and knew each other for an average of 19 years (range: 1 to 65 years). 79 per cent of the male respondents and 62 per cent of the female respondents were employed at the time of the interview and the mean household income per month was DM 5,380 (range: DM 1,000 to DM 21,000). The sample was well educated, with 49.5 per cent of the male and 39.5 per cent of the female respondents having completed 13 years or more of schooling. As will become apparent below, this sample showed a relatively high agreement in their proxy reports compared to other samples of couples (e.g., Sudman et al. 1994).

Student and Staff Couples Survey. In addition, we conducted a smaller survey of 20 couples recruited among students and staff at the University of Heidelberg. The male respondents of this study had a mean age of 36.2 years and the female respondents of 32.3 years; they knew one another for an average of 1.9 years and had been living together for an average of six months.

Additional experiments. Additional data were collected as part of ongoing experiments with students at the University of Heidelberg and the University of Mannheim, as well as with a convenience sample of pedestrians in a Mannheim shopping district. Demographic data were not collected for these subjects.

3. The Diverging Perspectives of Actors and Observers: Lessons from Social Cognition Research

As a vast body of literature in person perception indicates, individuals tend to explain behavior in terms of dispositions of the actor, that is, in terms of the actor's personality and likes and dislikes. Moreover, they largely neglect situational influences that may explain the actor's behavior. This overemphasis on personal dispositions, rather than situational influences, has been termed the "fundamental attribution error" (see Ross 1977, for a review). Importantly, this tendency is more pronounced in explanations that we provide for *others'* behaviors rather than in explanations that we provide for our own behavior. This asymmetry in causal focus is known as the *actor-observer difference*, a term introduced by Jones and Nisbett (1971). These authors observed, "The actor's view of his behavior emphasizes the role of environmental conditions at the moment of action. The observer's view emphasizes the causal role of stable dispositional properties of the actor. We wish to argue that there is a pervasive tendency for actors to attribute their actions to situational requirements, whereas observers tend to attribute the same actions to stable personal dispositions" (Jones and Nisbett 1971, p. 2). In a nutshell, others do things because they are "that kind of person," whereas we do things because they are the appropriate thing to do in this kind of situation. Several mechanisms have been offered to account for these differences in causal explanation and these mechanisms may operate either in conjunction or separately to produce the actor-observer effect.

First, the actor is likely to have a richer representation of his or her behavior available in memory, including not only the present behavioral episode, but also previous behaviors or accompanying thoughts and feelings to which the observer has no access. Thus, the actor may, for example, recall doubts and mixed feelings experienced during the behavior and may be aware that he or she showed a different behavior on other occasions. In contrast, the observer is unlikely to have access to this information and may hence overestimate the consistency of the actor's behavior over time and underestimate the actor's second thoughts. If so, observers are likely to overestimate the extent to which the behavior reflects stable dispositions of the actor.

Second, in personal interactions, the actor is usually in the focus of the observer's attention, whereas the visual orientation of the actor excludes him- or herself. Given that we tend to attribute causality to whatever is in the focus of our attention, this difference in attentional focus further contributes to observers' tendency to explain the actor's behavior in terms of dispositions of the actor, as well as to the actor's tendency to explain his or her behavior in terms of dispositional factors. Consistent with this assumption, actors are more likely to emphasize dispositional factors in explaining their own behavior when they are exposed to a videotape taken from an observer's perspective (see Watson 1982, for a review).

Finally, individuals are motivated to perceive their environment as predictable and their own behavior as an adequate response to environmental demands. These motivations have different implications for actors and observers. On the one hand, observers are more likely to perceive predictability when they attribute the actor's behavior to his or her disposition, which implies that similar behaviors may be expected in the future. On the other hand, actors' motivation to see their behavior as an adequate response to environmental demands fosters explanations of their behavior in terms of situational requirements. Again these differences elicit dispositional explanations from observers, but situational explanations from actors.

Each of these accounts of the actor-observer difference in causal explanation has received some support in experimental research (for a detailed review of this literature see Watson 1982) and the operation of these different processes is not mutually exclusive. For our purposes, the two key points to be taken from this literature are, (1) that actors have access to a richer set of episodic information in memory, and (2) that actors and observers differ in the causal explanations they provide. We address each of these points in turn.

With regard to episodic information, it is not surprising that actors are more likely than observers to have access to information related to specific behavioral episodes. To the extent that the question pertains to a recent reference period, or to a rare and salient behavior, they can rely on this information in generating a self report. If the question pertains to a distant reference period, the actor's "usual" behavior, or to behaviors that are frequent and mundane, on the other hand, episodic information is unlikely to be available (see Schwarz 1990a, for a review). In these cases, actors need to employ inference strategies that draw on more general information to generate a behavioral report. Proxy respondents, on the other hand, may not have access to episodic information about the actor's behavior and may therefore have to rely on more general characteristics of the actor in most cases, unless the report pertains to a behavior that self and proxy were jointly involved in. Not surprisingly, Sudman and colleagues (1994) observed in this regard that agreement

between self and proxy reports is typically somewhat higher for behaviors in which the reporter participated, than for behaviors which the actor performed in the absence of the proxy reporter.

Nevertheless, proxy respondents seem perfectly willing to report the other's behavior under conditions where they did not witness it. Similarly, actors are willing to report on behaviors for which little episodic information may be available in memory. The key question is: How do respondents do that? Presumably, they draw on general knowledge about the other person, or about the behavior, to generate a response. If so, however, the information they draw on is likely to reflect the differences discussed above. Most importantly, proxy respondents are likely to draw on dispositional information about the actor, reflecting the tendency for dispositional information to figure prominently in the cognitive representations that we form of others. For that reason, proxy reports may be likely to reflect an overreliance on the target person's apparently stable traits and dispositions, resulting in systematic biases. Self reporters, on the other hand, may be more likely to draw on situational information, reflecting that this information figures more prominently in the actor's perspective.

These conjectures have important implications for self and proxy reporting and generate a variety of testable hypotheses. Next, we elaborate on these hypotheses and report relevant findings.

4. Recall Strategies

As a first prediction, the previously discussed differences in the type of information available and used, should be reflected in the use of different recall strategies under self and proxy conditions. Respondents who are asked to report on their own behavior would be expected to focus on situational factors to infer or recall their behavior, whereas respondents who have to provide a proxy report would be expected to focus on dispositions of the proxy. Similarly, self reports of attitudes should be likely to be derived from features of the attitude object, whereas proxy reports of attitudes should be more likely to be derived from the proxy's assumed disposition. The available data provide strong support for this assumption in the domain of attitude reports.

Specifically, Bickart, Menon, Schwarz, and Blair (1994) had self and proxy sources make attitudinal reports about political groups, such as labor unions, women's rights groups, environmental groups, and the Ku Klux Klan. Analyses of think-aloud protocols, that were generated while respondents formed the judgments, suggest that self sources relied on features of the situation when making a judgment (the actor perspective), whereas proxy sources relied on features of the proxy when forming a judgment (the observer perspective). For the self reports, 42.1 per cent of the responses were based on features of the political group, whereas only 11.1 per cent of the responses under proxy conditions were based on these features. On the other hand, 37.5 per cent of the proxy reports relied on features of the proxy (i.e., personal dispositions) when forming the judgment, whereas only 14.0 per cent of the responses under self report conditions were based on features of the self. Thus, self and proxy reports were formed on the basis of different information. Moreover, the type of information used parallels the differences observed in the actor-observer literature in person perception: Whereas individuals tend to focus on the

situation or the target to explain their own behaviors or attitudes, they focus on features of the actor to explain others' behaviors or attitudes.

5. Inference Strategies and the Consistency of Related Responses

A second set of predictions pertains to the consistency of reports provided in response to substantively related questions. If self reports are based on more specific memories that take situational factors into account, whereas proxy reports are derived from general impressions of the other person, we may expect self and proxy reports to show different levels of differentiation. If respondents take their episodic memories into account in making self reports, but derive all related proxy reports from the same global impression, we may expect that judgments pertaining to related issues are more highly correlated in proxy than in self reports. We tested this hypothesis in two separate experiments.

Experiment 1. A first experiment was conducted with a convenience sample of $N = 132$ respondents (103 women and 29 men, mean age 52.3 years) in a Mannheim shopping district, randomly assigned to experimental conditions. Employing a between-subjects design, we had some respondents ($N = 65$) provide self reports and other respondents ($N = 67$) provide proxy reports for their spouses. Respondents were asked to indicate how much they, or their spouse, liked the three major political parties in Germany: the Christian Democrats, the Social Democrats, and the Greens (see Appendix for question wordings). It should be noted that the Christian Democrats are conservative, and the Social Democrats and the Greens are liberal parties. Moreover, recent surveys had indicated that Germans' party preferences at the time were not highly consistent – that is, many people agreed with different parties on different issues at a time when the country was undergoing major changes as a result of German unification.

As expected, our results indicated that self reports were indeed more differentiated than proxy reports. Specifically, respondents' self reported evaluations of the three political parties were uncorrelated, except for the correlation between the Greens and the Social Democrats, as shown in Table 1. On the other hand, proxy reports were highly correlated and clearly reflected the liberal versus conservative continuum – with strong negative correlations between liking of the liberal and conservative parties. Z-tests revealed that the correlations obtained under proxy report conditions exceeded the correlations obtained under self report conditions for CDU/SPD, $z = 3.24, p < .002$ and DCU/Greens, $z = 2.16, p < .03$, whereas the difference for SPD/Greens did not reach significance.

Table 1. Responses to related questions 1: liking of political parties

Self reports			
	CDU	SPD	Greens
		-.07	-.05
	SPD		.37*
Proxy reports			
	CDU	SPD	Greens
		-.57*	-.41*
	SPD		.46*

Note. Unpublished data (Wellens, de Graaf, and Schwarz). $N = 65$ for self reports and 67 for proxy reports. For correlations marked *, $p < .001$. Tests on differences reported in text.

Thus, when respondents reported about the party preferences of their spouses, their judgments were apparently based on a general impression of how liberal or conservative the spouse was. If the spouse is liberal, he or she must dislike the conservative party and vice versa for conservative spouses. On the other hand, judgments from the self perspective were uncorrelated, indicating that the judgments were probably based on other more specific and differentiating information, presumably bearing on specific political issues.

An important caveat needs to be addressed, however. Reflecting the population in the shopping district, most of the respondents in this experiment were women. Accordingly, the self reports reflect women's political attitudes, whereas the proxy reports reflect women's perceptions of their spouses' political attitudes. Unfortunately, the possibility cannot be ruled out that men's political attitudes are more consistent than women's, given that men tend to report higher political interest in most surveys (Erikson, Luttbeg, and Tedin 1988). If so, this gender difference may have contributed to the large differences obtained in the present experiment. However, data from a second experiment corroborate the main theoretical conclusion suggested by this first study.

Experiment 2. In this experiment, 20 heterosexual couples (recruited among the students and staff of the University of Heidelberg) were asked to provide reports of their own and their partner's attitudes and behaviors in response to general as well as specific questions. For example, a general attitude question would pertain to attitudes towards war in general, whereas a specific attitude question would pertain to the Persian Gulf War (see Appendix for question wordings). Similarly, a general behavior question would pertain to one's disposition to watch soccer, whereas the accompanying specific question would pertain to likelihood of watching the annual championship game. Each report was given along a seven-point rating scale. If proxy reports are based on dispositional information, whereas self reports are based on episodic information, we should find that the general and specific questions are more strongly correlated in proxy than in self reports.

The data shown in Table 2 support this prediction for six out of eight items, although the predicted difference is significant in only two cases and one significant reversal emerged. Treating the eight items as independent judgments, we can estimate the likelihood that the overall set of p -levels would have been obtained if the null hypothesis were true (Rosenthal 1984, p. 89). This test supports the hypothesis of stronger correlations under proxy report than self report conditions, $z = 1.96$, $p = .024$, one-tailed, although most of the individual differences failed to reach significance.

In sum, the obtained findings are consistent with the assumption that proxy respondents rely to a higher degree on dispositional information than self respondents. As a result, respondents' inferences about the proxy's specific behavior or attitude tend to be more in line with their perception of the proxy's general disposition with regard to the global domain, than is the case for self reports. From a methodological point of view, this implies that a high internal consistency of proxy reports cannot be used as an indicator of data quality. Instead, the apparent internal consistency may reflect proxy respondents' reliance on general dispositional information rather than the actual reliability of proxy respondents' memories.

6. Assumed Stability Over Time

The observation that individuals tend to explain their own behavior in terms of situational

Table 2. Responses to related questions 2: general and specific questions

	Self	Proxy	<i>z</i>	<i>p</i>
Attitude questions				
War in general/ Persian Gulf War	.35	.49	.73	.23
Interest in politics/ following election campaigns	.40	.71	1.99	.02
Catholic church/ celibacy	.33	.48	.77	.22
Christian Democratic Party/ Chancellor Kohl	.70	.50	1.37	.09
Behavioral questions				
Watching soccer/ watching the championship game	.35	.52	.91	.18
Saving money/ unplanned purchases	.18	.47	1.44	.08
Supporting the poor/ giving money to beggars	.37	.58	1.18	.11
Being reliable at work/ missing days at work	.41	.39	.10	.46
Overall			1.96	.024

Note. Based on 20 couples. Correlations exceeding .35 are $p < .01$; the *z* and *p* columns show the results of *z*-tests on the difference between correlations, one-tailed.

requirements, but others' behavior in terms of their personality dispositions, has another important implication. It suggests that individuals should infer that others' behavior is relatively stable and predictable. In contrast, one's own behavior should be perceived as flexible and responsive to situational demands – and some researchers have advanced motivational explanations for this discrepancy (McArthur 1981; Taylor and Fiske 1978). As a result, respondents should underestimate the variation in others' behavior relative to the self's behavior. By implication, they should assume that the other's behavior in any given reference period is more similar to his or her usual behavior than is the self's behavior in any given reference period. Data of Experiment 3 bear directly on this prediction.

6.1. Reports of "Usual" and "Extreme" behavioral frequencies

Experiment 3. In a between-subjects design, 24 students at the University of Heidelberg were asked to estimate either their own or their partner's typical weekly behavioral frequency and the highest and lowest weekly frequency of the respective behavior during the preceding month. As shown in Table 3, the reported highest and lowest frequencies deviated more from the estimated "typical" frequency for self reports than for proxy reports for each of the three reports assessed. Pooling over the consistent pattern of the three questions asked, the mean reported highest frequencies exceeded the mean reported usual frequencies by 4.38 for self reports, but only by 2.87 for proxy reports, $F(1, 22) = 4.38$; $p < .05$, whereas the mean reported lowest frequencies fell short of the mean reported usual frequencies by -4.48 for self reports and -2.81 for proxy reports, $F(1, 22) = 4.49$; $p < .05$.

Table 3. Mean difference between typical and highest or lowest behavioral frequencies

	Self		Proxy	
	Highest	Lowest	Highest	Lowest
Library	+3.41	−4.32	+2.13	−1.91
Snacks	+7.13	−6.14	+4.67	−4.93
Being late	+2.61	−2.98	+1.82	−1.61

Note. Shown is the difference between the estimated typical weekly frequency and the highest and lowest estimated weekly frequency during the preceding month.

That respondents assume that the proxy’s extreme frequencies deviate less from the proxy’s typical behavior than is the case for the self, is once again consistent with the assumption that respondents rely on dispositional information in making proxy reports. Reliance on this source of information necessarily results in overestimation of the behavior’s consistency over time, reflecting insufficient adjustments for situational influences. Next, we offer further support for this finding.

6.2. Reports of “Usual” frequencies and frequencies during a specified reference period

Experiment 4. If respondents are less aware of the variation in the proxy’s behavior than of the variation in the self’s behavior, it follows that reports of a target’s usual behavior and his or her behavior during a specific reference period should be more highly correlated for proxy than for self reports. The data of Experiment 3 do not allow us to address this implication because we elicited estimates of the highest and lowest frequency to provide a direct test of the variability hypothesis, rather than estimates of the frequency during a specified reference period. However, data pertaining to a specified reference period were assessed as part of the survey of 20 couples described above (Experiment 2).

As part of the above mentioned survey of 20 couples (Experiment 2), respondents were asked to report their own (or their partner’s) average daily TV consumption and their own (or their partner’s) TV consumption on the preceding day in an open response format. As expected, the reported average TV consumption and the report pertaining to the preceding day correlated $r = .52, p < .001$, for self reports, but $r = .83, p < .001$, for proxy reports, resulting in a significant difference in correlations, $z = 2.63, p < .01$. This finding supports the assumption that respondents are more likely to be aware of deviations from their own usual behavior than of the variation in their partner’s behavior, as has been suggested by the variability reports of Experiment 3.

While these findings, by themselves, are not surprising, they do have a number of important applied implications, which are addressed next.

7. The Convergence of Self and Proxy Reports as a Function of the Reference Period Used

From an applied point of view, the findings of Experiment 3 and 4 suggest that the convergence between self and proxy reports should be higher for reports of “usual” behavior than for reports pertaining to a specified reference period. If so, the use of proxy

respondents would be more adequate if the target's "usual" behavior is of interest than if a specified reference period is of primary concern. This would have clear implications for respondent selection rules as a function of the time frame that is to be assessed, as we shall discuss below.

Two additional findings suggest the same prediction. Specifically, findings in the person memory literature indicate that there may be a shift in the actor-observer perspective over time: Whereas actors initially place more emphasis on situational attributions than observers do, they later tend to attribute their own behavior more to dispositional factors than to situational ones (e.g., Moore et al. 1979, Peterson 1980). Thus, with the passage of time, actors' perspectives tend to become more like observers' perspectives by relying more on dispositional factors when making judgments. In line with this observation from the experimental literature, Nigro and Neisser (1983) observed a related phenomenon in a study on autobiographical memory. They asked subjects to report on a distant childhood event or on an event that happened "yesterday." The majority of their subjects reported distant events from an observer's perspective, but recent events from an actor's perspective. This suggests that people are more likely to adopt an observer perspective as time goes by, a shift that is reflected in a greater reliance on dispositional factors in the recall of older memories than in the recall of more recent ones. Moreover, Suengas and Johnson (1988) observed that memory for thoughts and feelings – i.e., subjective experiences, which are primarily accessible to the actor him- or herself – decays faster than memories for visual details, which are more prominent for an observer. Perhaps this differential decay in the available information over time can explain the observed shift in perspective over time. Once information that figures prominently in the actor's perspective is no longer accessible in memory, actors may find themselves in a situation similar to observers and may have to employ general heuristics and inference strategies to generate a report about distant events, much as is the case for proxy reports under all temporal distance conditions.

Again, this shift in actor-observer perspective over time has a number of interesting implications for self and proxy reports. It suggests that self reports of distant events (e.g., last year) should be more likely to reflect the impact of dispositional information than self reports of more recent events. Therefore, reports of distant events should be more similar to what one "usually" or "typically" does, than reports of more recent events (e.g., last week), which should utilize more situational and episodic information. Of course, this reliance on dispositional information for distant events is exactly what would be expected when the episodic information becomes less accessible for oneself over time. However, this loss of episodic information puts the self in the same position as the proxy, who may not have this episodic information in the first place. Thus, one would also expect that the correlation between self and proxy reports would be higher for distant than for recent events, unless the proxy's assumptions about the target's disposition differ from the target's self-perception of disposition.

In summary, proxy reporters should generally rely more on dispositional information than self reporters, but with the passage of time (for events recalled) both self and proxy reports should evidence an increase in the effect of dispositional information. Accordingly, the effect of using proxy rather than self respondents on data quality should be less pronounced for reports of distant as opposed to recent events. Empirical support

for these predictions would imply that survey researchers should try to obtain self reports if recent behaviors during a specified reference period are of interest, but may accept proxy reports without much loss in data quality (relative to self reports) if the target’s usual behavior or more distant events are of interest. Experiment 5 provides a first test of these predictions.

Experiment 5. This experiment was conducted as part of the above-mentioned community survey of couples. Respondents were asked to provide self and proxy reports of the frequency of reading books for fun (in terms of the number of books in a usual month, last month, or last year), missing a day at work due to illness (in terms of the number of occurrences in a typical month, last month, or last year), and eating out in a restaurant (also in terms of the number of occurrences in a typical week, last week, or last month; see Appendix for question wordings). Depending on experimental conditions, their reports pertained to the “usual frequency” (i.e., a “typical week”); a specific *recent reference period* (such as, “last week,” or “last month” depending on the expected frequency of the behavior); or a specific *distant reference period* (also “last month” or “last year,” again depending on the expected frequency of behavioral occurrences). Table 4 shows the convergence of self and proxy reports as a function of these reference periods.

For reading and days missed at work, self and proxy reports show significantly higher convergence (z ’s = 5.28 and 7.31, $p < .001$, respectively) for the distant and longer reference period of last month than for the recent and shorter reference period of last week, as predicted on theoretical grounds. Presumably, reports pertaining to last month are more difficult to compute on the basis of episodic information, putting the self in a situation that is similar to the situation of the proxy. As a result, both the self and the proxy refer to dispositional information, resulting in considerable convergence. On the other hand, the self has access to more episodic information for reports pertaining to last week than does the proxy, resulting in less convergence for this shorter and more recent reference period.

Note, however, that this does not hold for the “eating out” item. Specifically, the convergence of respondents’ reports of eating out was (nonsignificantly) higher for the reference period of one week than for the other reference periods. We conjecture that this reflects that eating out was a shared activity, which was hence better represented in memory, allowing proxy respondents to draw on episodic information for last week.

Finally, the reports of self’s and proxy’s “usual” behavior show significantly *lower* convergence than the reports pertaining to last month (z ’s = 3.89 and 8.29, $p < .001$,

Table 4. Convergence of self and proxy reports as a function of reference periods

	Recent	Distant	Usual
Read	.14 (48)	.84* (54)	.40* (48)
Miss work	.17 (36)	.95* (50)	-.05 (36)
Eat out	.66* (50)	.59* (58)	.52* (52)

Note. N ’s pertain to couples; different N ’s reflect missing data from at least one respondent in the couple. Correlations marked *, $p < .005$; other correlations $p > .20$.

for reading and missing days at work, respectively), contrary to the theoretical expectations. Moreover, the obtained convergence under “usual” conditions is not significantly higher than under “recent” conditions for the “reading” item, $z = 1.34, p = .18$, and even nonsignificantly lower for the “days missed” item, again contradicting predictions. On theoretical grounds, reports of the “usual” behavior should have most clearly reflected dispositional information. That convergence under this condition is lower than under the “last month” condition is particularly surprising. One possible explanation for this pattern is that convergence is highest when respondents rely on dispositional information and adjust the resulting estimate on the basis of some episodic information. Such a combination of informational sources may be most likely for reference periods like “last month,” where respondents may assume that they can hardly be expected to remember all episodes, rendering a “best guess” on the basis of episodic and dispositional information legitimate. Such a combination of sources is unlikely to be used, on the other hand, when the question is clearly dispositional in nature, as is the case when it refers to the “usual” behavior, thus discouraging the use of episodic information. Conversely, questions that pertain to reference periods like “last week” may carry the expectation that one should indeed remember the episodes, thus discouraging the use of dispositional information. These conjectures deserve more direct testing in future research.

From an applied point of view, the finding that deserves most attention is that, for reading and missing days at work, the convergence of self and proxy reports was higher for the reference period of one month than for the reference period of one week. Although our data do not allow us to determine if higher convergence reflects higher accuracy, convergences of $r = .84$ and $.95$ suggest that it may make little difference if self or proxy reports are obtained when one wants to assess the frequency of a behavior over an extended time period. Apparently, the judgmental situation of self and proxy reporters is quite similar under these conditions, resulting in high convergence and similar errors. If so, using proxy rather than self reports would be least problematic under these circumstances.

8. The Effect of Situational and Dispositional Cues

On theoretical grounds, we may expect that proxy respondents’ reliance on dispositional information may be reduced by an increased accessibility of situational information, which may be manipulated by appropriate questions. On the other hand, the reliance of self reports on dispositional information may increase if this information is rendered accessible by preceding questions. In general, the convergence of self and proxy reports should increase when self and proxy rely on the same type of information and should decrease when self and proxy draw on different types of information. Experiment 6 bears on these predictions.

Experiment 6. To explore the role of both sources of information, some respondents in the community couples survey were asked to describe relevant dispositions of the target (self or proxy), whereas other were asked to report relevant situational information. A third group of respondents received neither of these cueing questions. A comparison of the convergence of self and proxy reports under different cueing conditions bears on the theoretical rationale proposed here. Specifically, the convergence of self and proxy reports should be higher under conditions where self and proxy reporters rely on the same source of information.

Contrary to expectations, priming the same sources of information did *not* consistently increase the convergence of self and proxy reports, as shown in Table 5. This suggests that the priming questions used brought different information to mind, reflecting that self and proxy reporters had differential knowledge about situations and perhaps differential assumptions about their own and their partner's dispositions. As a result, the effect of the cues activated by the priming questions is mixed across the behaviors assessed. Although theoretically disappointing, these findings suggest that cueing questions may not increase convergence. Rather, convergence was consistently high when respondents were left to their own devices. We conclude that more research is needed to explore the nature of adequate cueing questions before any recommendations can be offered. At present, it seems better to allow respondents to use whatever strategy they are spontaneously using, rather than inducing them to use specific cues, which may not reflect their own subjective organization of the relevant information.

9. Getting from Dispositions to Behavioral Frequencies: The Effect of Response Alternatives

So far, we argued that proxy respondents are less likely to have access to relevant episodic information, and are therefore forced to rely on other judgmental strategies to generate a response. Given that dispositional information figures prominently in our cognitive representations of others, proxy respondents are highly likely to rely on this information, as the findings reviewed above indicate. Nevertheless, getting from a target's disposition to the frequency of a specific behavior, for example, still poses a complex task. When we know that someone is a "heavy smoker," how many cigarettes a day does that actually imply? A growing body of literature (see Schwarz 1990a, b, 1996, Schwarz and Hippler 1991, for reviews) suggests that respondents solve this task by referring to the frequency range of the response alternatives provided to them by the researcher.

Specifically, respondents assume that the average or typical behavior is reflected by values stated in the middle range of the response alternatives and that the extremes of the list reflect the extremes of the distribution. Based on this assumption, respondents

Table 5. Convergence of self and proxy reports as a function of relevant cues

	Cue for self/proxy				
	d/d	d/s	s/d	s/s	n/n
Behavior					
Late	.96 (39)	.51 (35)	.23@ (33)	-.06# (42)	.93 (30)
Newspaper	.49 (59)	.61 (56)	.41 (58)	.57 (54)	.73 (58)
Beer	.69 (57)	.79 (54)	.09# (55)	.79 (56)	.81 (51)

Note. N's pertain to couples; differences in N reflect missing data from at least one member of a couple; d = dispositional cue, s = situational cue, n = no cue. All correlations $p < .001$, except marked @, $p < .10$, and #, $p > .20$.

use the range of the response alternatives as a frame of reference in estimating their own behavioral frequencies and report higher frequencies on scales that present high rather than low frequency response alternatives (e.g., Schwarz, Hippler, Deutsch, and Strack 1985). Not surprisingly, they are more likely to rely on this estimation strategy, the less other relevant information is accessible in memory (see Menon, Rhagubir, and Schwarz 1995, Schwarz and Bienias 1990). This is reflected in a more pronounced effect of scale range on self reports of respondents with a chronically low accessibility of self related information (Schwarz and Bienias 1990, Experiment 3). Conversely, the effect of response scales is eliminated when respondents can draw on rate-of-occurrence information (e.g., “once a week,” Menon et al. 1995) that obliterates the need to compute an estimate based on scale range.

These considerations suggest that the effect of scale range should be more pronounced for proxy than for self reports. The relevant findings, however, are mixed. Schwarz and Bienias (1990, Experiments 1 and 2) observed that the effect of scale range on frequency reports of TV consumption was most pronounced when student respondents reported about a “typical undergraduate” of their university, less pronounced when they reported about a close friend, and least pronounced for self reports. Note, however, that the proxy reports obtained by Schwarz and Bienias (1990) pertained to a friend or to a “typical undergraduate” of the respondents’ university, not to a member of the same household. Hence, respondents’ knowledge about the proxy’s behavior may have been particularly poor. Qualifying the Schwarz and Bienias findings, Menon, Bickart et al. (1995) observed that scale effects were more pronounced for proxy reports than for self reports when the proxy was dissimilar to the respondent, but not when the proxy was similar. In the latter case, respondents could draw on information about the self to arrive at a proxy report, thus attenuating the effect of response alternatives to the level obtained for self reports. Hence, it is an open issue if the findings obtained by Schwarz and Bienias (1990) extend to proxy reports for members of the same household. To address this issue, we included Experiment 7 in the survey of couples drawn from a community sample.

Experiment 7. Specifically, we asked respondents to provide self and proxy reports of watching TV and taking the trash out along scales that presented either high or low frequency response alternatives.

As in previous research, respondents reported higher frequencies along scales that presented high rather than low frequency response alternatives, as shown in Table 6. For self reports, this resulted in a difference of 24 percentage points on the TV question and 6.1 percentage points on the trash question. The respective differences in proxy reports, however, are only slightly more pronounced, with 25.9 percentage points on the TV and 13.4 percentage points on the trash question. Contrast analyses on proportions (see Rosenthal and Rosnow 1985) revealed significant main effects of the response alternatives for the TV, $z = 8.46$, $p < .001$, as well as the trash question, $z = 3.48$, $p < .001$. However, the impact of response range did not vary as a function of self vs proxy report, as reflected in nonsignificant interaction terms of scale and report type for both questions ($z = .322$ and 1.3 , respectively). These findings are consistent with the results of Menon, Bickart et al. (1995) and suggest that the increased effect of scale range under proxy report conditions may pose less of a problem than the initial findings of Schwarz and Bienias

Table 6. Behavioral frequencies as a function of response alternatives and target person

Response alternatives	Target person			
	Self		Proxy	
	High	Low	High	Low
TV	54.5% (268)	30.5% (266)	60.1% (268)	34.2% (256)
Trash	30.5% (262)	24.4% (270)	37.2% (258)	23.8% (265)

Note. *N*'s pertain to individual respondents. Shown is the percentage of respondents who reported a TV consumption of more than ten hours per week or taking the trash out more often than three times per week. *N*'s given in parentheses.

(1990) suggested. Hence, we conclude that response alternatives do indeed bias behavioral frequency reports, but that the size of that bias is roughly the same for self and proxy reports pertaining to a member of the same household, in particular when the proxy's behavior is observable for the respondent.

10. Summary and Recommendations

Summing up, the reviewed findings indicate that proxy reports are more likely to be based on dispositional information than self reports, in line with what the person memory literature would suggest. This has several important applied implications:

First, proxy reports on related issues show a higher internal consistency than self reports, reflecting that related judgments are derived from the same dispositional information. Hence, the internal consistency of these reports should not be taken as evidence for their accuracy.

Second, proxy reports underestimate the variability of behavior over time. Accordingly, proxy reports and self reports show low convergence for short and recent reference periods, for which the actor can draw on episodic information in providing a self report. As the actor's access to episodic information decreases due to longer or more distant reference periods, the actor has to rely on dispositional information as well. As a result, the convergence of self and proxy reports increases. This implies that the use of proxy respondents is least problematic when long or distant reference periods are of interest, for which convergence was in the range of $r = .8$ to $.9$ in the present data.

Third, the expectation that priming similar cues for self and proxy respondents would increase convergence received no support. It is conceivable that this reflects that the same cueing question may bring different information to mind, as discussed above. Future research needs to address the types of cues that are helpful for different behaviors.

Fourth, although the frequency range of response alternatives influence respondents' reports, this influence was only slightly more pronounced for proxy than for self reports. Hence, the biases introduced by the use of response alternatives present a general problem, but this problem is not compounded when proxy respondents are used.

Finally, we emphasize that the present studies focused on the convergence of self and proxy reports and did not include an attempt to validate respondents' reports. Accordingly,

our findings bear on the conditions under which self and proxy respondents are likely to provide similar information, but they do not bear on the accuracy of that information. Future research will need to address the latter issue.

Appendix

Translated Question Wordings

The German self report versions used the term “you” (“Sie”), whereas the proxy report versions used the term “your partner” (“Ihr Partner/Ihre Partnerin”). Below, we present the translated self report versions of the questions used. The occasional awkwardness of the translated wordings results from our attempt to stay close to the German original.

Experiment 1

How do you [does your partner] feel about the following political parties?

Christian Democrats

not so favorable 1 2 3 4 5 6 7 8 9 very favorable

Social Democrats

not so favorable 1 2 3 4 5 6 7 8 9 very favorable

Greens

not so favorable 1 2 3 4 5 6 7 8 9 very favorable

Experiment 2

All questions were to be answered along seven-point rating scales and were presented in a random order, thus separating the general and specific question; six filler items, used as part of an unrelated wording experiment, were included in the randomized presentation.

Attitude questions

1. Do you generally feel that there are circumstances in which a war may be justified?
no, a war is never justified/yes, there are justifying circumstances

Do you feel that the Persian Gulf War was justified?

no, it was not/yes, it was

2. How interested are you in politics?

not at all/very much

How much do you follow election campaigns?

not at all/very much

3. Do you generally agree or disagree with the moral positions taken by the Catholic church?

agree/disagree

Do you agree or disagree with the Catholic church's requirements that priests have to

live in celibacy?
agree/disagree

4. How do you feel about the Christian Democratic Party?
not so positive/very positive

How do you feel about chancellor Helmut Kohl?
not so positive/very positive

Behavioral questions

1. How likely are you to watch soccer games on TV?
not so likely/very likely

How likely is it that you watch the annual soccer championship game on TV?
not so likely/very likely

2. Are you the kind of person who regularly puts some money into a savings account?
no, not at all/yes, very much

How often do you make unplanned purchases?
rarely/frequently

3. Do you think that it is important that all citizens contribute to supporting the poor?
no, not very important/yes, very important

How often do you give money to beggars?
rarely/frequently

4. Do you think of yourself as being very reliable at work or not so reliable?
not so reliable/ very reliable

How often do you miss a day at work?
rarely/frequently

Experiment 3

We are interested in a variety of activities that may vary in frequency. For each activity, we'd first like you to report how often you [your partner] engage in it in a *typical* week. Next, we'd like to think of last month and to report the lowest and the highest weekly frequencies during that month.

Library

How many items (books and journal articles) do you usually use in the library in a *typical* week during the semester?
_____ number of items

And what is the *lowest* number of items you used in a week last month?
_____ number of items

And what is the *highest* number of items you used in a week last month?
_____ number of items

Snacks

Thinking of a typical week, how often do you usually have a snack between regular meals?

_____ number of snacks

Thinking of times when you snack less often than usual, what was the lowest number of snacks you had in a week last month?

_____ number of snacks

And thinking of times when you snack more often than usual, what is the highest number of snacks you had in a week last month?

_____ number of snacks

Being late

Thinking again of a typical week, how often are you late for an appointment or social event?

_____ number of times

Let us now turn to a week when you were rarely late. What was the lowest number of times you were late in a week last month?

_____ number of times

And turning to a week in which you are particularly often late, what was the highest number of times you were late in a week last month?

_____ number of times

Experiment 4

How many hours do you [does your partner] spend watching TV on a typical day?

Please fill in : _____ hours

And how about yourself? How many hours did you [your partner] watch TV yesterday?

Please fill in : _____ hours

Experiment 5*Pleasure reading*

Usual: In a typical month, how many books do you usually read just for pleasure?

Recent: How many books did you read last month just for pleasure?

Distant: How many books did you read last year just for pleasure?

Days missed at work

Usual: In a typical month, how many days do you usually miss from work or school because of an illness or injury?

Recent: How many days did you miss from work or school last month because of an illness or injury?

Distant: How many days did you miss from work or school last year because of an illness or injury?

Eating out

Usual: In a typical week, how many times do you usually eat in a restaurant?

Recent: How many times did you eat in a restaurant last week?

Distant: How many times did you eat in a restaurant last month?

Experiment 6

Being late

Situational cue

Please list the circumstances that may cause you to be late to work or school.

Dispositional cue

Would you say that you are the type of person who is

☐ never late

☐ occasionally late

☐ often late

Target question

How many days did you arrive late to work or school last month?

Newspaper

Situational cue

Please list all of the places you are likely to read a newspaper.

Dispositional cue

When you read the newspaper, would you say that you typically:

☐ just skim over the headlines

☐ read only some of the articles

☐ read most of the articles

Target question

How many minutes did you spend reading the newspaper yesterday?

Beer

Situational cue

Please list all of the situations when you are likely to drink beer or wine.

Dispositional cue

Would you say you:

☐ never drink alcohol

☐ occasionally drink alcohol

☐ frequently drink alcohol?

Target question:

How many glasses of beer or wine did you drink during the last week?

Experiment 7

TV

In a typical week, how many hours of television do you usually watch?

Response alternatives:

Low

High

☐ up to 2 1/2 hours

☐ up to 10 hours

☐ 2 1/2 to 5 hours

☐ 10 to 15 hours

☐ 5 to 7 1/2 hours

☐ 15 to 20 hours

☐ 7 1/2 to 10 hours

☐ 20 to 25 hours

☐ more than 10 hours

☐ more than 25 hours

Trash

In a typical week, how many times do you usually take out the trash?

Response alternatives:

Low

High

☐ none

☐ less than 3 times

☐ 1 time

☐ 4 times

☐ 2 times

☐ 5 times

☐ 3 times

☐ 6 times

☐ more than 3 times

☐ more than 6 times

11. References

- Bickart, B., Menon, G., Schwarz, N., and Blair, J. (1994). The Use of Anchoring Strategies in Constructing Proxy Reports of Attitudes. *International Journal of Public Opinion Research*, 6, 375–379.
- Erikson, R.S., Luttbeg, N.R., and Tedin, K.L. (1988). *American Public Opinion*. New York: Macmillan.
- Jones, E.E. and Nisbett, R.E. (1971). *The Actor and the Observer: Divergent Perceptions of the Causes of Behavior*. Morristown, N.J.: General Learning Press.
- McArthur, L.Z. (1981). What Grabs You? The Role of Attention in Impression Formation and Causal Attribution. In E.T. Higgins, C.P. Herman, and M.P. Zanna (eds.), *Social Cognition: The Ontario Symposium*. Hillsdale, NJ.
- Menon, G., Bickart, B., Sudman, S., and Blair, J. (1995). How Well Do You Know Your Partner? Strategies for Formulating Proxy Reports and Their Convergence with Self Reports. *Journal of Marketing Research*, 32, 75–84.
- Menon, G., Raghubir, P., and Schwarz, N. (1995). Behavioral Frequency Judgments: An Accessibility-Diagnosticity Framework. *Journal of Consumer Research*, 22, 212–228.
- Moore, J.C. (1988). Self/Proxy Response Status and Survey Response Quality. *Journal of Official Statistics*, 4, 155–172.
- Moore, B.S., Sherrod, D.R., Lui, T.J., and Underwood, B. (1979). The Dispositional Shift in Attribution Over Time. *Journal of Experimental Social Psychology*, 15, 553–569.

- Nigro, G. and Nisser, U. (1983). Point of View in Personal Memories. *Cognitive Psychology*, 15, 467–482.
- Peterson, C. (1980). Memory and the “Dispositional Shift.” *Social Psychology Quarterly*, 43, 372–380.
- Rosenthal, R. (1984). *Meta-analytic Procedures for the Social Sciences*. Beverly Hills, CA: Sage.
- Rosenthal, R. and Rosnow, R.L. (1985). *Contrast Analysis: Focused Comparisons in the Analysis of Variance*. Cambridge: Cambridge University Press.
- Ross, L. (1977). The Intuitive Psychologist and His Shortcomings: Distortions in the Attributional Process. In L. Berkowitz (ed.), *Advances in Experimental Social Psychology* (Vol 10). New York: Academic Press.
- Schwarz, N. (1990a). Assessing Frequency Reports of Mundane Behaviors: Contributions of Cognitive Psychology to Questionnaire Construction. In C. Hendrick and M.S. Clark (eds.), *Research Methods in Personality and Social Psychology*. (Review of Personality and Social Psychology, Vol. 11; pp. 98–199). Beverly Hills, CA: Sage.
- Schwarz, N. (1990b). What Respondents Learn from Scales: Informative Functions of Response Alternatives. *International Journal of Public Opinion Research*, 2, 274–285.
- Schwarz, N. (1996). *Cognition and Communication: Judgmental Biases, Research Methods and the Logic of Conversation*. Hillsdale, NJ: Erlbaum.
- Schwarz, N. and Bienias, J. (1990). What Mediates the Impact of Response Alternatives on Frequency Reports of Mundane Behaviors? *Applied Cognitive Psychology*, 4, 61–72.
- Schwarz, N. and Hippler, H.J. (1991). Response Alternatives: The Impact of Their Choice and Ordering. In P. Biemer, R. Groves, N. Mathiowetz, and S. Sudman (eds.), *Measurement Error in Surveys* (pp. 41–56). Chichester, UK: Wiley.
- Schwarz, N., Hippler, H.J., Deutsch, B., and Strack, F. (1985). Response Scales: Effects of Category Range on Reported Behavior and Subsequent Judgments. *Public Opinion Quarterly*, 49, 388–395.
- Sudman, S., Bickart, B., Blair, J., and Menon, G. (1994). The Effect of Level of Participation on Reports of Behavior and Attitudes by Proxy Reporters. In N. Schwarz and S. Sudman (eds.), *Autobiographical Memory and the Validity of Retrospective Reports*. New York: Springer Verlag.
- Sudman, S., Bradburn, N., and Schwarz, N. (1996). *Thinking About Answers: The Application of Cognitive Processes to Survey Methodology*. San Francisco, CA: Jossey-Bass.
- Suengas, A.G. and Johnson, M.K. (1988). Qualitative Effects of Rehearsal on Memories for Perceived or Imagined Complex Events. *Journal of Experimental Psychology: General*, 117, 377–389.
- Taylor, S.E. and Fiske, S.T. (1978). Salience, Attention, and Attribution: Top of the Head Phenomena. In L. Berkowitz (ed.), *Advances in Experimental Social Psychology* (Vol. 11). New York, Academic Press.
- Watson, D. (1982). The Actor and the Observer: How are Their Perceptions of Causality Divergent? *Psychological Bulletin*, 92, 682–700.

Received July 1996

Revised January 1997