

FILTER QUESTIONS AND QUESTION INTERPRETATION

PRESUPPOSITIONS AT WORK

BÄRBEL KNÄUPER

Freie Universität Berlin

Researchers frequently ask respondents to report how often they engage in a certain behavior or how often they have experienced a particular event. In these cases, filter questions are often used to sort out respondents for whom more detailed follow-up questions may not apply and should therefore not be asked. This filtering is done by asking, in a yes/no format, whether or not the respondents experienced the particular event in question. Follow-up questions are then only presented to respondents who endorse having experienced the event. This way the burden on respondents who have no experience with the issue is reduced and interviewing time is saved. This article argues that such filter questions present presuppositions that may change respondents' interpretations of the question.

For example, respondents who are asked whether they have witnessed a crime need to determine what the researcher wants to include in the category "crime"—does this include graffiti and vandalism, or only major crimes like robbery or car theft? To infer the relevant meaning, respondents may turn to features of the question (see Schwarz [1996, chap. 5]; Sudman, Bradburn, and Schwarz [1996, chap. 3], for reviews). In the case of filter questions ("Did you experience X?"), asking whether or not a crime has been experienced during the reference period acknowledges that no relevant instances may have occurred, thus conveying that the researcher is interested in relatively rare events. In most domains, however, rare events are more extreme and serious than frequent events (see Schwarz et al. 1988). In contrast, a direct-frequency question of the type, "How often have you experienced X?" conveys that most people are likely to experience X in the respective time period, suggesting that the

Address for correspondence: Freie Universität Berlin, Psychologie (WE 10), Habelschwerdter Allee 45, D-14195 Berlin (Germany); E-mail: bknauper@zedat.fu-berlin.de. This research was conducted while the author was at the Institute for Social Research, University of Michigan, and was funded by an award fellowship granted by the Max Planck Society, Germany. The author would like to thank Norbert Schwarz, Bob Belli, Eleanor Singer, Britta Renner, and two anonymous reviewers for helpful comments on an earlier draft of this article.

Public Opinion Quarterly Volume 62:70-78 © 1998 by the American Association for Public Opinion Research
All rights reserved. 0033-362X/98/6201-0005\$02.50

researcher includes more frequent, and probably less serious, instances in the definition of the concept.

Suppose, for example, respondents are asked about the frequency with which they have experienced sexual assaults in either of the following ways: (a) "In the past five years, have you been sexually assaulted?" (yes/no) "If yes, how many times have you been sexually assaulted?" (b) "In the past five years, how many times have you been sexually assaulted?"

To answer such questions, respondents have to infer what qualifies as an instance of sexual assault. "Sexual assault" may have many possible meanings, some relatively benign, some quite awful. Do verbal remarks count as a sexual assault or only physical assaults? Asking, as in version (a), whether or not a sexual assault happened may accentuate the serious meaning of the term. When a question comes to mean something more serious, fewer instances will qualify, and the item is less likely to be endorsed.

Version (b), however, conveys the presupposition that it is quite likely that the respondent experienced at least some form of sexual assault in the respective time period. This may change the interpretation of the term "sexual assault" to include more minor instances. As a consequence, a higher frequency of sexual assaults might be reported.

Hence, filter questions may elicit a different interpretation of the issue than direct-frequency questions. Depending on what the researcher has in mind, filter questions may result in an unintended exclusion (and, hence, underreporting) of relevant events of minor seriousness. The present study tests this reasoning. Following the presentation of the data, implications for related findings (Sterngold, Warland, and Herrmann 1994) will be discussed.

Method

As part of a self-administered questionnaire, 92 University of Michigan undergraduates, randomly assigned to conditions, were asked to report the frequency of two crime-related experiences in the past 10 years: witnessing a crime and being the victim of a crime. In specific, in the filter-question version of the witnessing question, respondents were asked:

"Did you witness a crime in the past 10 years?" (yes/no)
"If yes, how many times?" (_____ times)

In the direct questioning version they were asked:

"In the past 10 years, how many times did you witness a crime?" (_____ times)

After reporting the frequency of witnessing a crime, subjects were asked in an open-response format to describe the kinds of crimes they had witnessed:

What kind(s) of crime(s) was/were this/these? Please describe.

Subjects were given several lines for the descriptions. The purpose of this question was to directly assess the severity or seriousness of the type of crimes that were on people's minds when answering the frequency question. This approach was chosen rather than the more indirect approach of asking respondents post hoc what they believe was meant by the question or what they thought the researcher's expectations were. The assumption was that examples convey more directly what type of experiences were on subjects' minds while searching for an answer. According to the reasoning outlined above, it was assumed that respondents report mainly severe examples of crimes when being asked the filter question, whereas both mild and severe crimes come to mind and are reported when being asked the direct question. Hence, on average, more severe crimes should be reported in the filter than in the direct-question version.

A second question sequence directly followed the witnessing-question sequence. It asked how often the respondent had been the victim of a crime in the past 10 years. Subjects were again asked to describe the types of events they had in mind when answering the frequency question in an open-response format.

Two independent raters, who were blind to the experimental conditions, later rated the severity of the descriptions of crimes that the subjects had provided. They evaluated the severity of the described crimes on a 5-point scale ranging from 1 = not at all severe to 5 = very severe. The interrater reliabilities for the ratings of the two crime experiences was $r = .82$ (witnessing a crime) and $r = .83$ (being the victim of a crime).

Results

Because of skewed distributions, the reported frequencies were log transformed. Part A of table 1 presents the mean reported frequencies as well as the mean log-transformed frequencies by experimental condition. For this first analysis, responses of no in the filter-question condition were treated as a frequency of zero. As expected, respondents reported significantly fewer events when being asked the filter question than when being asked the direct-frequency question (witness: $t(88) = 3.41, p < .001$; victim: $t(90) = 2.29, p < .01$). This might reflect that respondents to the filter question considered mainly severe examples to be the target of the question, whereas respondents to the direct question may have interpreted the term "crime" as also referring to nonsevere events.

Table 1. Mean Frequency and Mean Severity Ratings as a Function of Question Type

	Filter (<i>N</i> = 53)	Direct (<i>N</i> = 39)	Test of Significance
A. Reported frequency—overall:			
Witness crime (<i>N</i>)	51*	39	
Mean frequency (SD)	15.9 (83.7)	29.4 (54.7)	
Log mean (SD)	1.13 (1.31)	2.19 (1.65)	$t(88) = 3.41,$ $p < .001$
Victim of crime (<i>N</i>)	53	39	
Mean frequency (SD)	.79 (1.04)	1.97 (3.45)	
Log mean (SD)	.44 (.52)	.75 (.76)	$t(90) = 2.29,$ $p < .01$
B. Saying no or zero (%):			
Witness crime (<i>N</i>)	41.2 (21 of 51)	17.9 (7 of 39)	$\chi^2 = 5.56, df = 1,$ $p < .02$
Victim of crime (<i>N</i>)	54.7 (29 of 53)	38.5 (15 of 39)	$\chi^2 = 2.38, df = 1,$ $p < .12$
C. Reported frequency—excluding no or zero:			
Witness crime (<i>N</i>)	30	32	
Mean frequency (SD)	27.0 (108.5)	35.8 (58.6)	
Log mean (SD)	1.92 (1.19)	2.67 (1.42)	$t(60) = 2.27,$ $p < .027$
Victim of crime (<i>N</i>)	24	24	
Mean frequency (SD)	1.75 (.85)	3.21 (3.95)	
Log mean (SD)	.97 (.29)	1.21 (.61)	$t(32, 95) = 1.76,$ $p < .088$
D. Rated Severity:			
Witness crime (<i>N</i>)	32	32	
Mean severity (SD)	4.23 (.46)	3.64 (.89)	$t(46.3) = 3.35,$ $p < .002$
Victim of crime (<i>N</i>)	24	24	
Mean severity (SD)	4.19 (.53)	3.67 (.70)	$t(46) = 2.91,$ $p < .006$

* Two respondents said yes to the witness filter question but subsequently failed to provide a frequency report. This leaves *N* = 51 for the overall frequency analysis (part A) and *N* = 30 for the frequency analysis excluding no (part C). These two respondents did, however, describe an example and are therefore included in part D.

Part B of table 1 shows that many more respondents answer no to the filter question than report a frequency of zero in response to the frequency questions that ask about witnessing a crime and being the victim of a crime (witness: 41.2 percent vs. 17.9 percent; victim: 54.7 percent vs. 38.5 percent). It seems that respondents in the direct-questioning condition felt inclined to include instances of crime in their frequency report that those in the filter-question condition did not feel qualified as instances of crime. Even after removing those respondents from the analysis who said no or zero, respectively, the mean frequency of reported crimes is reliably lower for the filter question than for the direct-frequency question (witness: $t(60) = 2.27, p < .027$; victim: $t(32.95) = 1.76, p < .088$; see pt. C of table 1).

Taken altogether, the data suggest that there is a general shift in the reported frequency for the entire group of respondents toward reporting fewer events when asked the filter questions. Since severe crimes (like homicides) tend to happen less frequently than do minor crimes (like store thefts), the results reported so far can be interpreted as indirect evidence that respondents assume more severe events to be the target of the question in the filter version than in the direct-questioning version. In the direct-frequency question, they presumably assume that both minor and major crime-related experiences qualify as instances of crime.

More direct evidence that respondents consider differentially severe events to be the target of the question can be found in subjects' descriptions of crime instances. Following the frequency question, respondents were asked to describe "what kind(s) of crime(s)" these were. Two independent raters later judged the severity of the described instances of crime. The analysis of these severity ratings reveals, as expected, that respondents described more severe crimes in the filter-question condition than in the direct-questioning condition, as shown in part D of table 1 (witness: $t(46.3) = 3.35, p < .002$; victim: $t(46) = 2.91, p < .006$). These findings suggest that respondents who were asked the filter question based their frequency reports on more serious instances of crime than did those who were asked the direct-frequency question.

Discussion

In summary, the findings demonstrate that respondents are sensitive to the presuppositions conveyed in the questions. The data suggest that the questions conveyed different assumptions about the frequency of the target event. Asking whether a crime has been experienced may acknowledge that no relevant instances may have occurred, thus conveying that the researcher is interested in relatively rare, and thereby more serious, events. In contrast, the direct-frequency question may convey that the tar-

get event happens rather frequently to most people. Relating this information to their knowledge that frequently occurring events tend to be less serious (store thefts happen more frequently than homicides), respondents might assume that the researcher includes less serious instances in the definition of "crime." It's not surprising that this difference in interpretation resulted in higher frequency reports for the direct question than for the filter question.

In a similar vein, it can also be assumed that the length of the reference period that is used in a question can affect question interpretation. In the present example, respondents were asked to report crimes they had witnessed in the past 10 years. It is conceivable that a shorter reference period, such as the past month, would also shift the interpretation of the term "crime" toward referring to less serious crimes. This impact of conversational norms conveyed by the length of the reference period has recently been demonstrated by Winkelman and colleagues for questions asking about the frequency of anger in the past year, month, or week (Winkelman, Knäuper, and Schwarz 1997). Respondents referred to more serious episodes of anger when asked about the frequency of anger in a typical year rather than in a typical month or week.

It is informative to contrast the analysis reported here with other accounts of the impact of filter questions. In a related study, Sterngold, Warland, and Herrmann (1994) either asked respondents directly, "How concerned are you about [an issue]?" or presented a filter question ("Are you concerned about [an issue]?") before asking a question about respondents' degree of concern. The use of the filter question roughly doubled the percentage of respondents who said that they are "not concerned." The authors suggest that the direct question presupposes that respondents "are concerned about the issue, or that they should be concerned. Consequently, some respondents may accommodate this expectation by overstating their actual concerns" (p. 256). In contrast, the filter question may "encourage" respondents to admit their lack of concern. This social desirability account is consistent with theorizing about the operation of no-opinion filters, which are often assumed to "encourage" respondents to admit that they do not have an opinion (see, e.g., Bishop, Oldendick, and Tuchfarber 1983; Schuman and Presser 1981; but see Hippler and Schwarz [1989] for a different analysis). In a similar vein, one may propose that the direct-frequency question used in the present study invited or encouraged respondents to report crime experiences. Note, however, that this logic does not account for the observed shift in question meaning found in the present study in the examples that the subjects reported after answering the questions, nor is it obvious why having experienced a crime would be a desirable response.

Conversely, the present perspective may well account for Sterngold, Warland, and Herrmann's (1994) findings (as well as for the findings re-

ported by Schaeffer and Dykema [1994]): being asked how concerned one is indeed conveys that the researcher considers some level of concern likely, whereas the filter question does not convey this assumption. As a result, respondents may infer that different degrees of "concern" are the target of the question (e.g., "something that really bothers and preoccupies me" when asked the filter question vs. "something I'd prefer to be different" when asked the direct question). Also, the meaning of the public concern issue itself may change as a function of question type. For example, when asked about "air pollution," people may interpret the term as referring to (more severe) industry-related air pollution when asked the filter question and to (less severe) tobacco-smoke-related air pollution when asked the direct question. These shifts in meaning would result in the observed pattern of findings—yet, respondents' reports would pertain to different levels of "concern" or to different severe instances of the public concern issue. Hence, respondents who are asked the direct question would not "overstate" their concerns relative to respondents who are asked the filter question; instead, both groups would report on different levels of concern or on different question targets. This meaning-shift account suggests that both reports are subjectively meaningful, yet pertain to somewhat different questions, whereas Sterngold, Warland, and Herrmann's (1994) account suggests that the direct version elicits meaningless "overstatements."

Although a contribution of social desirability is unlikely for the specific target issue used in the present study (witnessing or experiencing crimes), it is generally conceivable that social desirability tendencies can additionally contribute to filter-question effects, such as the one found by Sterngold, Warland, and Herrmann (1994). Filter questions may simultaneously induce social desirability tendencies and meaning shifts. If one would employ a very broad definition of socially desirable response behavior, one could even argue that the meaning shift proposed here is the result of social desirability. According to the conversational perspective employed in the present article (cf. Schwarz 1996), respondents to a survey try to be cooperative communication partners. They do their best to infer what the researcher means by the question and what kind of response he or she expects (cf. Schwarz 1996). The direct question may suggest to respondents that the researcher expects a nonzero number. In order to comply with that expectation (= social desirability), the respondent might eventually broaden the range of acceptable target events (= meaning shift) and may think of and include less serious events. In this broad sense, the effect observed here could be interpreted as a form of acquiescence or socially desirable response behavior.¹

It should be investigated in future studies whether there might be an

1. I thank an anonymous reviewer for this notion.

additional contribution of social desirability to the filter-question effect (as Sterngold, Warland, and Herrmann [1994] seem to propose) that goes beyond the rationale just provided. Further research is also needed in order to test the generalizability of the present findings. The findings presented here are based on just one sample of a student population, and the assumptions were tested for just one particular class of events (witnessing or experiencing crimes). These further studies should also attempt to assess respondents' understanding of what the researcher expected or intended in a more comprehensive way than the one employed in the present study.

Given that the use of filtered versus unfiltered question forms may change respondents' interpretation of what the question refers to, researchers must determine to what extent the information conveyed by each question form is compatible with their intended interpretation of the question. Depending on what the researcher has in mind, filter questions may actually filter out more respondents than intended by conveying a more extreme question meaning. Hence, asking a filter question can lead to an underreporting of events in which the researcher was interested. Obviously, such effects are limited to questions that refer to behaviors or experiences that vary along a dimension of extremity, as was the case for the present study. When the question pertains to a well-defined behavior (e.g., "Do you own a house?" vs. "How many houses do you own?"), no meaning shift is to be expected, and both question forms are likely to result in the same answers.

To avoid unwanted shifts in question meaning, surveyors can provide more concrete definitions of the target issue or can pretest filter questions to make sure that the intended definition matches what respondents understand. The present findings indicate that we need to pay close attention to what we want the respondents to consider as possible instances of the issues in question. As Smedslund (1988, p. 149) notes, "Only when we know how an item is understood can we know what the response to that item means."

References

- Bishop, George F., Robert W. Oldendick, and Alfred J. Tuchfarber. 1983. "Effects of Filter Questions in Public Opinion Surveys." *Public Opinion Quarterly* 47:528-46.
- Hippler, Hans-Jürgen, and Norbert Schwarz. 1989. "'No Opinion' Filters: A Cognitive Perspective." *International Journal of Public Opinion Research* 1:77-87.
- Schaeffer, Nora Cate, and Jennifer Dykema. 1994. "Open and Closed Questions about the Frequency of Events: Replication, Extension, Paradox." Working Paper no. 94-18, University of Wisconsin—Madison, Center for Demography and Ecology.
- Schuman, Howard, and Stanley Presser. 1981. *Questions and Answers in Attitude Surveys*. San Diego, CA: Academic Press.
- Schwarz, Norbert. 1996. *Cognition and Communication: Judgmental Biases, Research Methods, and the Logic of Conversation*. Hillsdale, NJ: Erlbaum.

- Schwarz, Norbert, Fritz Strack, Gesine Müller, and Brigitte Chassein. 1988. "The Range of Response Alternatives May Determine the Meaning of the Question: Further Evidence on Informative Functions of Response Alternatives." *Social Cognition* 6:107–17.
- Smedslund, J. 1988. "What Is Measured by a Psychological Measure?" *Scandinavian Journal of Psychology* 29:148–51.
- Sterngold, Arthur, R. H. Warland, and Richard O. Herrmann. 1994. "Do Surveys Overstate Public Concerns?" *Public Opinion Quarterly* 58:255–63.
- Sudman, Seymour, Norman Bradburn, and Norbert Schwarz. 1996. *Thinking about Answers: The Application of Cognitive Processes to Survey Methodology*. San Francisco: Jossey-Bass.
- Winkielman, Piotr, Bärbel Knäuper, and Norbert Schwarz. 1997. "Looking Back at Anger: The Interpretation of Emotion Frequency Questions Depends on the Time Frame Used." Manuscript under review.