SAQS AND SEX: A RE-ANALYSIS OF THE NATIONAL HEALTH AND SOCIAL LIFE SURVEY

Mick P. Couper (University of Michigan and JPSM) & Linda Stinson (Bureau of Labor Statistics)
Mick P. Couper, Survey Research Center, P.O. Box 1248, Ann Arbor, MI 48106

Introduction and Background

The National Health and Social Life Survey (NHSLS) was designed to measure sexual behaviors and attitudes in the United States using a national probability sample and state of the art survey methods (Laumann et al., 1994; Miller, 1995). The study was conducted in 1992 by the National Opinion Research Center (NORC) at the University of Chicago.

The NHSLS was an interviewer-administered survey, but included a set of four self-administered questionnaires (SAQs) to elicit responses on the more sensitive questions. This paper focuses on the responses to these SAQs. SAQs are used to increase the accuracy or honesty on sensitive topics through increased privacy. However, this is predicated on the assumption that such enhanced privacy is indeed achieved in the interview. If interviewers assist respondents in the completion of SAQs or even complete them for some respondents, the data quality gains may be diminished. Hence, we explore the factors that may influence the self-completion decision, and the potential impact of this decision on the resultant data.

Couper and Rowe (1996) examined self-completion in the context of a computer assisted self interview (CASI). They found that in 79% of cases the respondent completed the CASI items themselves, while in 7% of cases they had interviewer assistance and in 14% of cases the interviewer completed the self-administered items on behalf of the respondent. Their findings suggest that who completed the self-administered items does appear to matter in terms of data quality. This paper replicates and extends their work in examining whether similar effects are found in paper and pencil SAQs.

There are three types of reasons why people may choose not to self-complete an SAQ, given that they have agreed to do the interview and provide data to the interviewer. We review each of these briefly in turn, and explore their effect empirically in the NHSLS.

1. Capacity

Some respondents may be unable to complete a SAQ because of illiteracy or health conditions (e.g., impaired vision) that physically constrain their ability to do so.

We have no measures of physical impairment in the NHSLS, and so use age as a weak proxy for visual impairment or other physical constraints on self-completion. Smith (1992) found age to be significantly related to item nonresponse on two SAQ items in the GSS, and similar effects were found for CASI self-completion by Couper and Rowe (1996). However, the NHSLS was restricted to persons age 18 to 59, diminishing the utility of age as an indicator of reduced capacity. In addition, age may also be related to the degree of perceived sensitivity of the questions on sexual activity. In summary, we would expect those with lower education, and older persons to be less likely to self-complete the SAQs.

2. Motivation

Even if respondents have the capacity to complete an SAQ, they may be unwilling to do so. We assert that the cognitive (and possibly physical) effort involved in self-completion is greater than that in responding to an interviewer's questions. Thus, allowing the interviewer to administer the items may be a form of satisficing (Krosnick, 1991). The task is still completed, but with minimal additional effort.

We thus expect that respondents who are less motivated to participate in the survey initially or who are less engaged in the interview are also less likely to self-administer the SAQs. We distinguish between two indicators of motivation, 1) indicators of initial reluctance (obtained from records of the data collection process) and 2) indicators of lack of engagement (obtained from interviewer reports at the end of the interview).

3. Sensitivity

Here the argument is that those for whom the SAQ items are sensitive would be more likely to self-complete, while those who regard the topics as less sensitive would be less likely to self-complete. However, this is confounded with the argument that self-completion improves reporting of sensitive information. Those who engage in sensitive behaviors might be more likely to self-administer the SAQs, but self-administration is hypothesized to increase the reporting of socially sensitive behaviors over interviewer-administration. Thus we cannot tell whether the act of self-completion leads to more honest reporting, or simply that those who have more to reveal about their behaviors are more likely to choose self-administration. To attempt to avoid this tautology, indicators of sensitivity are taken from the

interviewer-administered portion of the interview rather than the SAQs.

Furthermore, sensitivity may interact with motivational factors. If one has nothing sensitive to report, or is not averse to the reporting of illicit or disapproved sexual behaviors to an interviewer, the added cognitive costs of completing the SAQ may not be worth it. But if one does have sensitive information to report, it may be worth the additional effort to avoid embarrassment.

We explore the relative effect of these three sets of factors in determining who completes paper and pencil SAQs. As we cannot examine interviewer effects directly here, we focus on capacity, motivation and sensitivity as possible factors in self-completion.

Design and Analysis

The NHSLS is based on a multistage area probability sample of households in the United States (Laumann et al., 1995). The population of inference was to all people age 18-59 with adequate English proficiency living in households in the United States. The sample consists of two parts, a cross-sectional sample and an oversample of Blacks and Hispanics. One sample person was selected from each household.

The results presented here are based on weighted analyses, where the weight is a combination of a sampling weight (to reflect the oversample), an eligibility weight (for household size) and a poststratification weight (primarily for differential nonresponse). Tests of statistical significance and standard error estimates are based on Taylor Series approximation, reflecting the complexity of the design. The models were estimated using SUDAAN (Shah, Barnwell, and Bieler, 1996).

The survey was conducted by trained NORC interviewers using paper and pencil methods. Most of the interviewers used for the study were women. An overall response rate of 78.6% was achieved, yielding a total of 3,432 completed interviews, 3,159 of which were from the cross-sectional sample, and the balance from the oversample.

The NHSLS instrument contained four SAQs, interspersed throughout the interview. The content of the SAQs is listed briefly as follows:

SAQ1: Personal and family income.
SAQ2: Number of sex partners, frequency of sex, etc.
SAQ3: Masturbation questions.
SAQ4: Separate versions for male and female respondents asked questions about oral sex, anal sex, paid sex, drug use, etc.

Interviewers were instructed to hand the SAQs to respondents at the appropriate points in the interview. On completion of each form, respondents were instructed to place it in a privacy envelope which was sealed at the end of the interview. Interviewers then completed the following check item for each SAQ:

R filled out SAQ   1
Interviewer assisted with SAQ 2
R refused SAQ   3

We used this classification to examine responses to individual SAQs. However, we also examined the completion pattern across the four SAQs. Finding no other clear pattern across the four, we created a combined measure with the following three categories: R did all SAQs, R did some SAQs, and R did no SAQs. The last category includes both those respondents who refused all four and who had the interviewer complete all four for them. In terms of the capacity hypothesis, we assume that those who completed <u>any</u> of the four are capable of completing an SAQ. For the motivation hypothesis, we consider those who completed <u>all</u> four to be highly motivated. In other words, we expect the capacity measures to be more closely associated with the third category of the combined variable, and the motivation measures to be associated with the first category. For the sensitivity hypothesis, it is more likely to depend on the nature of the questions, and so we may expect to see differential completion (some but not all SAQs). However, we also examine individual SAQs to explore the sensitivity issue.

To summarize, we pose three general questions. First, who does and who does not self-complete SAQ items? Second, why do some self-complete and others not? Finally, we explore the effect self-completion versus interviewer completion may have on the data. Would different conclusions be reached if these cases were included or excluded in the analyses?

Table 1 contains the distributions of who completed each SAQ. Across all 4 SAQs, 17.2% of respondents sought assistance from the interviewer at some point, while for only 3.3% was no SAQ information obtained from any of the four SAQs.

The distribution on the combined measure is as follows:

R did all SAQs without assistance   65.7%
R did some SAQs without assistance   19.8%
R did no SAQs without assistance   14.6%

We use this variable in the next few sections to explore the extent to which various types of respondents self-completed all SAQs.

Who Self-completes and Why?

Table 2 presents the percentage distributions for the combined measure of self-completion by various indicators of capacity, motivation and sensitivity. We discuss each of these in turn.

1. Capacity

We have only two indicators of capacity in the NHSLS data, education and age. We expected a positive relationship between education level and self-completion. However, we find a curvilinear effect, with the proportion self-completing all SAQs peaking at some college, then declining slightly. One explanation may be that capacity constraints (literacy) operate at the lower education levels, but motivation may be more important at the higher education levels. Education could also be confounded with age or other variables.

In terms of age, we expected older respondents to be less inclined to self-administer because of vision or cognitive impairments associated with age. However, the NHSLS limited the sample to those age 18-59, and we do not expect to find such effects at these ages. The bivariate results show a monotonic but non-significant (p>.05) decline in self-response (the "All" category) with age. This provides weak support for the capacity hypothesis.

2. Motivation

We explore two sets of motivational factors, namely indicators of initial reluctance and measures of lack of engagement. In terms of initial reluctance, indicators include the following: whether the case was transferred to another interviewer (a common refusal conversion strategy); whether telephone conversion of the case was attempted; and an interviewer rating of how difficult it was to get the case. For each variable, we expect greater reluctance to be associated with lower probability of self-completion of all SAQs.

We can see from Table 2 that there are consistent effects across these indicators (as well as others not included in the table). In each case, those who showed greater initial reluctance are less likely to self-complete the SAQs.

To examine lack of engagement, we use 3 post-interview ratings by interviewers, on the respondents' comprehension, cooperativeness and frankness respectively. Given the positively skewed responses to these items, we collapsed each rating into a dichotomy. These are also presented in Table 2. It can be seen that all 3 indicators of engagement in or commitment to the interview are strongly related to the likelihood of self-completing the SAQs. Again, this suggests that motivational factors may be important in determining who self-completes and who does not.

3. Sensitivity

Finally, we examine sensitivity as a possible factor in the decision to self-complete. Here we postulate that those for whom the SAQ items are sensitive would be more likely to self-complete the SAQs.

A distinction can be made between engagement in "sensitive behaviors" and the respondent's attitudes about the sensitivity of such behaviors. For example, one would expect a drug user who is sensitive about his/her drug use to be more likely to self-administer an SAQ on the topic than a drug user who is not sensitive about his/her drug use. Both these groups should still have higher likelihood of self-administration than non-drug users.

To measure the likelihood of engagement in sensitive behaviors, and the perceived sensitivity of such behaviors, we used measures external to the SAQs themselves. The interviewer-administered portion of the NHSLS contains a series of items on the appeal of various sexual behaviors, including oral sex, anal sex, group sex, and so on. One indicator of sensitivity is the mean responses to a series of 14 questions on the appeal of various behaviors. Our assumption is that those who admit to finding fewer "disapproved" sexual behaviors appealing, may be less likely to self-complete the SAQs, as they have less to hide. The combined measure is rescaled to the original item response categories (1=very appealing to 4=not at all appealing, so that a high score indicates greater sensitivity).

A further indicator of sensitivity may be the number of sex partners reported by the respondent in the interviewer-administered portion of the interview. We could assume that those in a monogamous relationship or those with no sex partners would find reporting the self-administered items to an interviewer less sensitive.

Another factor that may increase the sensitivity of the items is the presence of others during the interview. We hypothesize that with others present, respondents may be more likely to self-administer the SAQs. The relationship between these 3 indicators of sensitivity and the distribution of completion of the 4 SAQs are also presented in Table 2.

Using these indicators of sensitivity, we find little support for the sensitivity hypothesis. Those who completed all SAQs themselves exhibited significantly less sensitivity to these interviewer-administered items (i.e., reported more of them to be appealing) than those who answered all or some of the SAQs themselves. The significance of the sex partners variable appears largely driven by the missing cases, suggesting that item-missing data in the interviewer-administered portion of the interview is associated with non-self-completion of the SAQs, which suggest motivational issues more so than sensitivity. These data suggest, albeit with weak measures of sensitivity, that the content of the SAQs had little effect on the likelihood of their completion by the respondent.

What Effect Does This Have on Estimates?

We have seen that there is a substantial minority of respondents who choose for various reasons not to self-complete the SAQs. The obvious next question is whether this makes any difference to the data collected in the SAQs. This is difficult to answer because of the problem that those who have the interviewer complete the SAQs for them may have different answers than those who self-completed because of the method of completion or because those with different behaviors make different self-completion decisions. In the current design, we cannot distinguish between these two possibilities. Rather, we present the SAQ responses for both self-completers and interviewer-assisted completers to see whether any differences are present.

To do so, we focus on SAQ3 and SAQ4, containing the most sensitive questions. We should note that the category "neither completed" in Table 1 contains some cases for whom data are present in the SAQ. Our assumption is that for these cases, the interviewer check item on who completed the SAQ was not completed, so we cannot determine who actually did the SAQ.

SAQ3 contains a series of items on masturbation. Table 3 presents the percentages of respondents reporting selected activities on this SAQ by who completed the SAQ. SAQ4 contains a series of items on various heterosexual practices. These are presented in similar fashion in Table 3.

The percentages in Table 3 suggest differences in the reporting of these sexual practices by self-administration of the SAQ or whether the interviewer assisted. However, given the small number of cases in the self-completed category for these two SAQs (see Table 1), the effect on the overall estimates is not large. For males, none of the three differences reach statistical significance (p<.05). The effects are stronger for female respondents, with all differences between self-completers and interviewer-assisted reaching significance (p<.05). The effect in each case is to lower the estimates of engagement in the practice.

Discussion and Conclusions

We have seen that there are systematic differences in the likelihood of self-completion of SAQs by respondents. We find modest support for the capacity hypothesis, but did not expect strong effects given the variables at our disposal and the limited age range of the sample. Motivation factors appear to be an important determinant of who completes an SAQ. Those for whom it took greater effort to bring into the respondent pool appear less likely to self-complete, suggesting a possible relationship between nonresponse error and data quality. We find little support for the sensitivity hypothesis. This

is encouraging, in that respondents' decisions to do the SAQs themselves do not appear to be related to the behavior measured in the SAQs.

The findings we present here are similar to those reported by Couper and Rowe (1996), suggesting that the problem of non-self-completion is not confined to computerized instruments. Both capacity and motivation appear to affect the decision to self-complete, whether it be a paper and pencil or a computerized instrument. Audio-CASI appears to offer promise in reducing the effect of literacy on self-completion, but may in turn deter those with little or no experience with computers. Further, the use of audio-CASI is unlikely to have much effect on the motivational factors in the decision.

We have also found some evidence that self- versus interviewer-completion has some effect on the response reported in the SAQs. This is consistent with the argument for using SAQs in the first place. While the effect on overall estimates is not large, given the relatively small proportions of those who chose not to self-complete, the impact on estimates may be more pronounced in subgroup analyses.

To summarize, the is ample empirical evidence that self-administration increases the overall reporting of sensitive items. However, the results reported here, and those of Couper and Rowe (1996) suggest that not all respondents are willing or able to complete such items. Obviously efforts to maximize self-completion are desirable. We have been unable to explore interviewer-level effects here, but suspect there may be interviewer variation in efforts to encourage respondents to self-complete. This is worthy of further exploration. Similarly, we believe the work of Cannell, Miller and Oksenberg (1981) in attempting to gain explicit commitments from respondents to increase motivation are noteworthy. Our findings also raise questions about the relationship between nonresponse error and measurement error (see also Couper, 1997). At the very least, these findings suggest we can identify, even prior to or at the onset of the interview, those at risk for non-self-completion of SAQs. A variety of strategies could then be employed to reduce the likelihood of non-self-completion or even non-completion, including motivational strategies and incentives.

These results also suggest that who completes the SAQs is an important quality indicator, similar to item missing data rates or nonresponse rates. NORC should be commended for including these interviewer check-items on who completed the SAQs in the instrument, and making them available to other researchers. With these indicators researchers could consider the possibility of statistical adjustment to account for the non-self-completion of SAQs. One could at least use these

measures to judge the quality of the data collected in the SAQs and make decisions on treating such cases as missing data or considering statistical adjustment.

These preliminary analyses have shown that there are systematic differences in the completion of SAQs on sensitive items in interviewer-administered surveys. This fact, and the implications for data quality, deserve further attention.

References

Cannell, C.F., Miller, P.V., and Oksenberg, L. (1981), "Research on interviewing techniques." Chapter 11 in Leinhardt, S., *Sociological Methodology 1981.* San Francisco: Jossey-Bass.

Couper, M.P. (1997), "Survey introductions and data quality." *Public Opinion Quarterly,* 61 (2): 317-338.

Couper, M.P. and Rowe, B. (1996), "Evaluation of a computer-assisted self-interview (CASI) component in a CAPI survey." *Public Opinion Quarterly,* 60: 89-105.

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

Laumann, E.O., Gagnon, J.H., Michael, R.T., and Michaels, S. (1994), *The Social Organization of Sexuality: Sexual Practices in the United States.* Chicago: University of Chicago Press.

Laumann, E.O., Gagnon, J.H., Michael, R.T., and Michaels, S. (1995), *National Health and Social Life Survey, 1992 (Computer file; ICPSR version).* Chicago, IL: University of Chicago and National Opinion Research Center (producer). Ann Arbor, MI: Inter-university Consortium for Political and Social Research (distributor), 1995.

Shah, B.V., Barnwell, B.G., and Bieler, G.S. (1996), *SUDAAN User's Manual, Release 7.0,* Research Triangle Park, NC: Research Triangle Institute.

Smith, T.W. (1992), "A methodological analysis of the sexual behavior questions on the General Social Survey." *Journal of Official Statistics,* 8 (3): 309-325.

Table 1. Completion of the Four Self-Administered Questionnaires

	Respondent only	Interviewer assistance	Neither or not reported
SAQ1	79.7%	15.0%	5.3%
SAQ2	82.3%	13.3%	4.4%
SAQ3	79.6%	11.6%	8.8%
SAQ4	73.4%	11.0%	15.6%

Table 2. Respondent Completion of SAQs by Indicators of Capacity, Motivation and Sensitivity

	All	Some	None	(n)
Capacity Indicators				
Education:				
<12th grade (%)	54.2	23.5	22.3	(530)
High school (%)	65.0	19.9	15.1	(1,007)
Some college (%)	69.8	19.1	11.1	(1,115)
Completed college (%)	69.8	17.3	12.8	(541)
Graduate school (%)	64.7	19.8	15.6	(239)

$$\chi^2 = 36.0, \text{d.f.} = 8, p < .001$$

	All	Some	None	(n)
Age:				
18-24 (%)	69.7	20.7	9.6	(558)
25-34 (%)	65.9	18.6	15.5	(1,070)
35-44 (%)	65.8	20.1	14.1	(947)
45-54 (%)	62.9	19.8	17.3	(595)
55-59 (%)	61.8	20.9	17.3	(262)

$$\chi^2 = 15.4, \text{d.f.} = 8, \text{n.s.}$$

Motivation Indicators: Initial Reluctance

Case transferred to another interviewer:

	All	Some	None	(n)
Yes (%)	52.7	19.8	27.5	(515)
No (%)	68.1	19.7	12.1	(2,644)

$$\chi^2 = 22.6, \text{d.f.} = 2, p < .001$$

Telephone conversion of case:

	All	Some	None	(n)
Yes (%)	45.7	14.0	40.3	(335)
No (%)	68.0	20.4	11.6	(3,097)

$$\chi^2 = 51.7, \text{d.f.} = 2, p < .001$$

Interviewer rating of difficulty:

	All	Some	None	(n)
Very easy/easy (%)	68.9	19.1	12.0	(2,205)
All other (%)	59.8	20.9	19.3	(1,227)

$$\chi^2 = 19.0, \text{d.f.} = 2, p < .001$$

Motivation Indicators: Lack of Engagement

Respondent comprehension:

	All	Some	None	(n)
Excellent, good (%)	67.4	19.4	13.2	(3,148)
Other (%)	47.3	23.5	29.2	(284)

$$\chi^2 = 17.1, \text{d.f.} = 2, p < .001$$

Table 2 (continued)

	All	Some	None	(n)
Respondent cooperation:				
V. cooperative (%)	68.0	19.2	12.8	(2,897)
Other (%)	53.1	22.9	24.1	(535)
$\chi^2 = 33.1$, d.f. =2, p<.001				
Respondent frankness:				
Entirely frank (%)	68.5	18.7	12.8	(2,417)
All other (%)	58.9	22.3	18.8	(1,015)
$\chi^2 = 15.3$, d.f. =2, p<.001				

Indicators of Sensitivity

	All	Some	None	(n)
Appeal of sexual behaviors:				
Mean	3.08	3.07	3.20	(3,432)
(s.e.)	(0.02)	(0.02)	(0.02)	
F=13.09, d.f.=2, p<.01				
Number of sex partners in past year:				
None (%)	61.3	23.7	15.1	(408)
One (%)	69.7	18.8	11.5	(2,353)
Two or more (%)	64.7	23.8	11.5	(517)
Missing (%)	15.7	14.7	69.6	(154)
$\chi^2 = 40.5$, d.f. =6, p<.01				
Presence of others during interview:				
Yes (%)	66.0	22.9	11.1	(410)
No (%)	65.6	19.3	15.1	(3,022)
$\chi^2 = 3.28$, d.f. =2, p>.05				

Table 3. Response to SAQ items by who completed SAQ (percentages)

	R only	IWER assisted	Total
Frequency of masturbation (men):			
Once a week or more	26.2	11.5	24.3
Other	36.8	41.9	37.4
Not at all	37.0	46.6	38.6
Total	100.0	100.0	100.0
χ^2=20.25, d.f.=2, p<.01			

Table 3 (continued)

	R only	IWER assisted	Total
Frequency of masturbation (women):			
Once a week or more	6.9	4.3	6.5
Other	33.7	20.7	32.2
Not at all	59.4	74.9	61.3
Total	100.0	100.0	100.0
χ^2=12.22, d.f.=2, p<.01			
Among those who reported masturbating, % always or usually having orgasm:			
Men	89.4	83.9	88.8
χ^2=1.27, d.f.=1, n.s.			
Women	75.6	63.9	74.6
χ^2=3.04, d.f.=1, n.s.			
% always or usually feeling guilty after masturbation			
Men	60.6	28.9	57.2
χ^2=15.85, d.f.=2, p<.01			
Women	53.2	34.0	51.7
χ^2=4.99, d.f.=1, p<.01			
Percent reporting active oral sex:			
Men	74.9	62.6	73.3
χ^2=5.3, d.f.=2, n.s.			
Women	66.9	48.0	64.5
χ^2=17.3, d.f.=2, p<.01			
Percent reporting receptive oral sex			
Men	75.9	68.2	74.8
χ^2=3.5, d.f.=2, n.s.			
Women	71.9	54.4	69.7
χ^2=9.9, d.f.=2, p<.05			
Percent reporting active anal sex			
Men	23.6	19.3	23.0
χ^2=1.5, d.f.=2, n.s.			
Percent reporting receptive anal sex			
Women	19.3	10.9	18.3
χ^2=8.8, d.f.=2, p<.05			