

## DESIGN AND RESULTS OF THE WOMEN'S HEALTH STUDY

Roger Tourangeau, National Opinion Research Center,  
Jared B. Jobe and William F. Pratt, National Center for Health Statistics,  
and Kenneth Rasinski, National Opinion Research Center

Roger Tourangeau, NORC, 1350 Connecticut Avenue, Suite 500, Washington, DC 20036

**KEY WORDS:** Computer Administration, CASI, Mode Effects, Sensitive Questions

### 1. Introduction

This study investigates some sources of error in surveys that collect information on sensitive topics, topics that involve illegal or embarrassing activities. More specifically, the study tested procedures to improve the accuracy of data collected in the National Survey of Family Growth (NSFG). Since its beginning in 1971, the NSFG has obtained detailed information on fertility and reproductive health. In each of the four cycles of the NSFG conducted to date, the survey has explored a broad range of sensitive questions, concerning topics such as contraceptive practices, pregnancy histories (including fetal and infant deaths), unplanned and unwanted pregnancies, sexually transmitted diseases, and infertility.

From the outset, there have been concerns about the sensitive nature of the questions in the NSFG. For example, the first two cycles of the NSFG (carried out in 1973 and 1976) generally excluded women who had never been married, because it was believed that many unmarried women would not answer questions about pregnancy and contraceptive practice truthfully. Despite these concerns, the sample was expanded in the 1982 NSFG to represent all women regardless of marital status. Changes to the content of the NSFG questionnaires have also increased the sensitivity of the interview over time. For example, the 1988 NSFG added questions on risk factors for AIDS. Despite the increasingly intimate information being sought in the NSFG, response rates have remained high. Around 80% of the cases selected for the NSFG complete the interview and only about one-third of the nonrespondents are outright refusals. Moreover, nonresponse to individual questions has generally been less than 1%. Of course, the fact that respondents answer the questions offers no assurance that their answers are truthful.

#### 1.1 Sensitive Questions in Surveys

From the point of view of survey methodology, this study concerns a very general problem--how to collect data on topics that most people are likely to

regard as private. Many surveys include questions about private or potentially embarrassing matters, asking respondents about their annual income, their employment status, and so on. With the coming of the AIDS epidemic, the need for data on such sensitive topics as sexual behaviors and illicit drug use has dramatically increased. But though the need for such data is clear, it is not clear whether the data collected are accurate.

Findings from surveys on sexual behavior illustrate the problems in collecting sensitive data in surveys. Within a closed population, equal numbers of opposite-sex sexual partners should be reported by men and women; the same sexual pairings are being reported by respondents of both sexes. As Smith (1992) has demonstrated, however, men consistently report more opposite-sex sexual partners than women do, a difference that persists even when differences in the population sizes are taken into account. The most plausible account of the discrepancy is that men overstate their partners and that women overlook theirs. A recent review of the methodological problems in AIDS research described the situation this way: "Most sex research is based on self-reported sexual behavior of unknown validity" (Catania, Gibson, Chitwood, & Coates, 1990, p. 339). Much the same judgment would apply to research on illicit drug use, abortion, and other sensitive topics.

**Improving reporting on sensitive topics.** A hypothesis guiding much of the survey literature on reports about sensitive topics is that a major source of error is more or less deliberate misreporting. Questions about sensitive topics create conflicts for the respondents, who generally want to cooperate by giving correct answers, but who also want to avoid embarrassment or, when the behavior in question is illegal, legal repercussions. Much of the methodological research designed to improve answers to sensitive questions has concentrated on techniques that reduce the perceived threat of the questions by increasing the privacy of data collection. More recently, methodological studies have begun to examine the effect of computer-assisted data collection techniques on reporting of sensitive behaviors.

Increasing the privacy of data collection is widely believed to improve the accuracy of the answers. One of the most practical methods for increasing perceived privacy is to use self-administered questionnaires (SAQs) rather than face-to-face interviews to collect the data. In most surveys, the data are not entirely confidential because at least the interviewer is aware of the respondent's answers; further, when interviewers administer the questions and record the answers it is possible for other household members to overhear what the respondent is saying. Surveys that employ SAQs (in which respondents record their answers without the mediation of an interviewer) overcome these threats to confidentiality. Self-administered questionnaires generally obtain higher levels of reporting of sensitive behaviors than do face-to-face interviews, with telephone interviews falling somewhere in between the other two modes in levels of reporting (see Bradburn, 1983, for a review). The advantages of SAQs have been demonstrated for a number of sensitive topics, including sexual behavior (Boekeloo, Schiavo, Rabin, Conlon, Jordan, & Mundt, 1994), illicit drug use (Aquilino & LoSciuto, 1990; Schober, Caces, Pergamit, & Branden, 1992; and Turner, Lessler, & DeVore, 1992), alcohol consumption (Aquilino & LoSciuto, 1990; Hochstim, 1967), and abortion reporting (London & Williams, 1990; Mosher & Duffer, 1994; Mott, 1985).

Several studies, for example, have shown that self-administration increases reporting of illicit drug use, alcohol consumption, or both. The first of these, by Aquilino and LoSciuto (1990), compared drug use data collected by interviewers over the telephone with data collected in a self-administered form as part of a personal interview. It found substantially higher reporting of both drinking and drug use with the self-administered questions. Two subsequent comparisons between face-to-face interviewing and self-administered questionnaires also found greater reporting of cocaine and marijuana use in the self-administered condition (Schober et al., 1992; Turner et al., 1992). Finally, an early comparison of face-to-face data collection with data collection by mail and telephone revealed lower levels of reported alcohol consumption in face-to-face interviews (Hochstim, 1967).

Self-administration also appears to reduce survey respondents' reluctance to admit that they have had an abortion. Mott (1985) reports evidence that self-administration greatly increased the number of abortions reported, and similar results have been obtained in studies of abortion reporting by London and Williams (1990; see also Mosher & Duffer, 1994).

Boekeloo and his colleagues demonstrate that self-administration also increases reporting on other sexual topics; respondents were more likely to admit to unprotected sexual intercourse and a history of sexually transmitted diseases in a self-administered questionnaire than in a face-to-face interview (Boekeloo et al., 1994).

Another method that increases the apparent confidentiality of survey responses is the randomized response technique (Warner, 1965). In this technique, a random device rather than the interviewer determine what question the respondent answers (e.g., the respondents spins a dial to determine which of two questions to answer); in this way, the interviewer cannot know for sure what the respondent's answer means. The randomized response technique method has been shown to increase the proportion of women reporting that they have had abortions (Abernathy, Greenberg, & Horvitz, 1970; I-Cheng, Chow, & Rider, 1972; Shimizu & Bonham, 1978). However, the procedure is difficult to use in a large survey and greatly complicates the analysis of the results.

**Other variables affecting levels of reporting.** Answers to threatening questions also appear to be affected by the format and wording of the questions (Bradburn, 1983), although the results for these variables are not so well documented as those for self-administration. Whether the items use an open or closed response format appears to have an effect on reporting of sensitive behaviors. For example, Bradburn, Sudman, and their colleagues (1979) found that, compared to closed questions, open questions produced increases in reporting that ranged from 14 percent for frequency of sexual intercourse during the past month to 108 percent for frequency of masturbation; over a number of sensitive items, the average increase in the level of reporting was 52 percent for the open as compared to the closed versions of the questions. However, the format of the questions did not affect whether respondents report engaging in the behaviors at all.

Longer questions may also yield fuller reporting. Reports about the frequency of behavior and amount of consumption are subject to memory errors even when there are no motivational obstacles to truthful reporting (see Jobe, Tourangeau, & Smith, 1993, for a review). Particularly if the behavior is frequent and episodes are not highly differentiated (as with frequent use of illicit drugs), respondents may not remember how many times they have engaged in the behavior during a particular reference period. By giving respondents more cues and more time to search their memories, longer questions can produce more complete reporting (Marquis & Cannell, 1971). In

their study of sensitive behaviors, Bradburn and his coworkers (1979) found that longer questions produced consistently higher levels of reported behavior.

A promising new technique that may increase perceived privacy and produce more accurate data on sensitive behavior is the computer-assisted self-administered interview, or CASI. A study by Waterton and Duffy (1984) found that a computer-administered questionnaire produced greater reports of alcohol use than a conventional face-to-face interview. This study confounds the effects of computer assistance and those of self-administration, as do several other studies on CASI (Locke et al., 1992; Lucas et al., 1977; Robinson & West, 1992).

There is, however, some evidence that computer assistance by itself can enhance the reporting of sensitive behaviors. In an analysis a test of computer-assisted personal interviewing (CAPI), Baker and Bradburn (1991) found that CAPI respondents were more likely than respondents to a paper-and-pencil interview to report having used birth control methods in the past month.

## 1.2 Variables for This Study

This study was based on the assumption that survey reports about abortion and other sensitive topics might be improved through several means. We selected three of these strategies for investigation. The first strategy--increasing the privacy of the data collection process--was already well-established in the survey methods literature. We chose this strategy over some of the other possibilities in the literature because we believed that privacy was the single most powerful variable affecting reporting on sensitive topics and that manipulating this variable would have the largest impact on the survey estimates.

In examining the privacy variable, we decided to test the improvements produced by self-administration; we ruled out the major alternative, the randomized response technique, because of the practical and statistical difficulties associated with that procedure. However, in addition to the use of self-administered questions, we sought to test the impact of moving the interview outside the respondent's home (and away from other family members). Relatively few studies have recorded whether face-to-face interviews involving sensitive topics were conducted in private or with other household members present or able to overhear the respondent's answers. As a result, the effects of the privacy of the setting in which the interview is carried out are unclear. We hypothesized that moving the interview to a neutral

site away from other family members might increase the respondent's sense of privacy and thus improve reporting.

The second approach we examined was that of placing the interview as a whole and, especially the questions on abortion, in a medical context. A medical context for the interview would, we thought, reinforce the need for accurate data for health planning purposes; in addition, we believed that respondents might be more accustomed to providing candid answers in the setting of a medical interview than in the survey setting. We attempted to foster a medical context in two ways. First, we used interviewers who were themselves medical practitioners--that is, we used nurses and nursing assistants as interviewers for some of the respondents. We know of little prior work investigating this approach for collecting sensitive data and sought to test its effectiveness in this project. Second, we attempted to foster a medical context by beginning the interview with a long series of questions about medical conditions and procedures.

A final strategy investigated in this project was the use of computer-assisted data collection. Early evaluations suggest that computer assistance may enhance either the apparent privacy of data collection or the perceived objectivity and importance of the study; either way, we hypothesized that computerization of the data collection process might increase respondents' willingness to report truthfully.

## 2. Methods

We conducted a large-scale field experiment in the city of Chicago. More than 1,000 women were interviewed, along with a small comparison sample of 100 men. The sample was selected from two sources. All of the men and most of the women respondents were selected from an area probability sample that had been screened to identify persons in the eligible age range (ages 15 through 35); the rest of the women were selected from rosters at cooperating health clinics and were known to have had abortions.

Questionnaires based on the one used in the National Survey of Family Growth were administered to the sample; the questionnaires included items on abortion, sexual behavior, and illicit drug use and took about an hour to complete (Rieger, Judkins, & Sperry, 1991). The experiment examined five variables: 1) whether the questionnaire began with a series of medical questions or with questions on pregnancy; 2) whether the interview was conducted by a nurse or a regular field interviewer; 3) whether the interview was done at the respondent's home or at a site outside the

home; 4) whether the interviewer or respondent administered the questions; and 5) whether the data were collected via computer or on paper. The analysis examined a number of outcome variables, including the response rates under the various experimental conditions, the level and accuracy of abortion reporting, and the level of reporting on other sensitive topics (such as the number of sexual partners). The focus here is on reports about sexual behavior (see Jobe et al., in press, for findings on the other topics).

## 2.1 Sample

**Area probability sample.** The area probability portion of the sample was a stratified, multistage sample of dwellings in the city of Chicago, selected using standard methods. In the first stage of selection, a sample of 85 area segments was drawn; each segment consisted of a single block or group of adjoining blocks, defined using data from the 1990 Census. After all of the blocks in the city of Chicago had been sorted by geographic area, we selected a systematic sample of 85 of them. Selection probabilities for each segment were proportional to the 1990 Census count of the number of housing units it contained. This method of sample selection assured that each area in the city of Chicago would receive proportionate representation in the sample. Each segment included at least 40 housing units (according to the census data); blocks that did not meet this size standard were linked to adjacent blocks until the combined unit included 40 or more housing units.

A subsample of dwellings on the 85 sample segments were designated to receive a short screening interview to identify persons eligible for the main experiment. In total, 6,325 occupied dwellings were selected for screening. Screening interviews were completed at 4,659 of these, for a response rate of 73.7%. The screening interview gathered information on the race, sex, age, and Hispanic background of person living at the dwelling. The screeners yielded information about 10,998 persons, of whom 3,141 were within the eligible age range (i.e., 15 to 35 years old at the beginning of the field period for the experiment).

**Clinic sample.** Two Chicago health clinics agreed to cooperate in the study by providing the names of women who had had abortions during the preceding year or so. The time frame was defined so that no one would be selected who had had an abortion during the three-month period prior to the beginning of data collection for the experiment. (Because the field period for the experiment was delayed, this window of eligibility in fact ended more

than nine months before the experiment began.) The clinic sample was also restricted to women who lived in the city of Chicago; the eligible age range, however, was expanded slightly relative to that for the area probability sample to include women between the ages of 15 through 40. The two clinics provided a total of 1,088 names.

To protect the confidentiality of the women selected from the clinic sample, the first author carried out the selection of both the clinic and area probability samples, and only he was aware of the sample from which the individual cases had been selected. In addition, we used an after-the-fact permission form procedure in which women who completed the interview were asked to sign a release form giving us access to their medical records at their sources of gynecological care. Women from the clinic sample who refused to sign the permission form were dropped from the analysis and their data were eliminated from the data files. A total of 48 members of the clinic sample were dropped for this reason.

**Selection of cases for the experiment.** Between the area probability and clinic samples, a pool of more than 4,200 persons was available for the experiment. We selected a subsample of 2,266 of these and randomly assigned them to a treatment cell. Within the area probability sample, the selection of persons for the experiment required several steps. In the first step, each household with eligible members was placed in one of six strata that were defined by sex, age (15 through 19 vs. 20 and over), and minority group membership. Households with members in more than one eligible group were randomly assigned to a single stratum, and, because it was impractical to interview more than one person from the same household, only one eligible household member was retained for the main study. Then, after each household had been assigned to a single stratum, a systematic sample was selected; the use of a systematic procedure assured that the members of the sample were drawn from all of the area segments. Altogether, 1,564 cases were selected for the experiment from the area probability sample.

The selection process for the clinic cases was considerably simpler than that for the area probability cases. Once each woman on the clinic lists had been classified by age category and minority status, a random sample was selected from each group. Initially, 544 women were selected for the experiment; subsequently, a supplemental sample including an additional 188 women was selected. Table 1 shows the number of cases selected for the experiment by source and stratum.

Table 1. Initial Sample Sizes

| Stratum                | Source       |            | Total        |
|------------------------|--------------|------------|--------------|
|                        | Area         | Clinic     |              |
| Younger Minority Women | 237          | 82         | 319          |
| Older Minority Women   | 549          | 398        | 947          |
| Other Younger Women    | 52           | 22         | 74           |
| Other Older Women      | 372          | 230        | 602          |
| Younger Men            | 29           | --         | 29           |
| Older Men              | 325          | --         | 325          |
| <b>TOTAL</b>           | <b>1,564</b> | <b>732</b> | <b>2,296</b> |

**Response rates.** A few names provided by the clinics turned out to be duplicates; in addition, the screening data regarding a person's age were sometimes in error and some members of the sample had moved outside of Chicago before the field period began. After these losses, 1,914 women and 350 men remained eligible for the study. After the sample was fielded, it became necessary to subsample males as a cost-saving measure; ultimately, only 100 men were interviewed. Table 2 shows the response rates for the study; the overall response rate for women was 55.2%. More than two-thirds of the nonrespondents were cases who were never contacted (primarily because they could not be located during the field period); of

Table 2. Response Rates

| Group    | Eligible | Complete | Response Rate |
|----------|----------|----------|---------------|
| Males    | 350      | 100      | --            |
| Females  | 1,914    | 1,059    | 55.2%         |
| Black    | 1,053    | 652      | 61.7%         |
| White    | 569      | 270      | 47.5%         |
| Hispanic | 191      | 88       | 46.1%         |
| Other    | 101      | 49       | 48.5%         |
| Area     | 1,191    | 705      | 59.1%         |
| Clinic   | 723      | 354      | 48.8%         |

the women who were contacted, about 85% completed an interview.

Of the 354 completed clinic cases, 48 refused to sign permission forms and the permission forms for another six women were lost; data for these cases were dropped, leaving 300 clinic cases for the analysis.

## 2.2 Experimental Design

We manipulated five variables in this experiment, in a completely crossed design. Two of the variables, *interviewing staff* and *version* of the questionnaire were attempts to enhance the medical context of the interview; we thought that respondents might be more willing to discuss sensitive topics in a survey if the context reinforced the health-related purposes of the study and if medical practitioners administered the questions.

Accordingly, we varied the *interviewing staff*, comparing nurses and nursing assistants with regular field interviewers. We hypothesized that nurses would elicit more reports of sensitive behaviors than regular field interviewers. The two *versions* of the questionnaire included the same items but varied the order in which two sets of abortion questions appeared. In one version, a series of pregnancy history questions came first; in the other, a set of questions about medical procedures questions first. In the pregnancy first version, the topic of abortion was initially raised during a series of questions about the respondent's pregnancy history; in the medical conditions first version, abortion was first mentioned in connection with a series of medical procedures affecting reproduction. We hypothesized that more abortions would be reported by respondents receiving the medical procedures questions first. We also believed that the combination of the two sets of abortion questions would yield more reported abortions than either set of questions alone.

We varied the *mode of data collection*, comparing paper-and-pencil to computer-assisted interviews, and the *method of administration*, comparing interviewer-administered to self-administered interviews. Crossing the mode of data collection and method of administration resulted in four groups: Interviewer-administered paper-and-pencil interviews (PAPI); computer-assisted personal interviews (CAPI); paper-and-pencil self-administered questionnaires (SAQ); and computer-assisted self-administered questionnaires (CASI). We hypothesized that respondents in both self-administered conditions (those completing the SAQ or CASI questionnaires) would report higher levels of sensitive behaviors.

We varied the *site* of data collection, conducting interviews either in the respondent's home or at a neutral site. We hypothesized that levels of reporting would be higher in the neutral site interviews, where other members of the household could not overhear the answers. A variety of sites was used for the interviews conducted outside the home, with NORC offices and neighborhood restaurants being the most frequent.

### 2.3 Instruments

At the beginning of each interview, the respondent was asked to note three or four important personal events on a calendar to help date events later in the questionnaire. Both versions of the questionnaire began by asking demographic questions. These were followed by the medical procedures and pregnancy history questions in counterbalanced order; both of these series of questions included items on abortion. The pregnancy history questions were the questions usually used on the NSFG, and asked the respondent to list all her pregnancies in order and to report certain data about each pregnancy, including its outcome (i.e., live birth, stillbirth, ectopic pregnancy, miscarriage, or abortion). The medical procedures questions were developed for this experiment and asked whether the respondent had had any of a number of medical procedures affecting reproductive health. Six of the procedures were methods for inducing an abortion: dilation and curettage (D & C) to end a pregnancy; dilation and evacuation (D & E) or suction curettage to end a pregnancy; injection of saline solution or prostaglandin to end a pregnancy; hysterectomy to end a pregnancy; hysterectomy during a pregnancy; and abortion, type unknown.

For the remaining topics, the two versions of the questionnaire were identical. Both versions contained numerous questions about the respondent's sexual behavior. Items asked when and with whom the respondents first had sexual intercourse, and whether it was voluntary; other items asked about the number of sex partners during the previous year, the previous five years, and in total. The questionnaires also contained items on whether respondents had had a sexually transmitted disease. In the section of questions on medical conditions, respondents were asked whether they had had chlamydia, gonorrhea, genital warts, genital herpes, or syphilis. Finally, there were items asking the respondents about their use of condoms in the last year and the last 30 days.

The questionnaires also included a series of items on illicit drug use. The initial drug question

asked whether the respondent had ever used any illegal drug, and follow-up questions asked about their use of marijuana, amphetamines, barbiturates, tranquilizers, psychedelics, cocaine, crack, and heroin. Another series of questions, for users of injectable drugs, asked how they cleaned their needles and related drug paraphernalia, and how often they shared them with other users.

## 3. Results

Our discussion of the results focusses on sexual behaviors. More specifically, we examined the average number of sexual partners reported as a function of the sex of the respondent and of the experimental variables; we then examined responses on the other sexual topics in the questionnaire, including sexually transmitted diseases and condom use. Because so few men completed the interview, we report mainly the results for the women, only occasionally discussing the results for the men.

### 3.1 Reported Sexual Partners

The data on the number of reported sex partners are counts and, as is common with such data, the distribution of the responses is highly skewed. To compensate for this departure from normality, we added .5 to the values and then carried out a logarithmic transformation prior to performing the analyses of variance. For ease of interpretation, we report untransformed values in presenting the group means. For respondents who had been sexually active for only one year, or for only five years, we used the number of sexual partners for that period as the value for questions about longer time spans.

**Experimental effects.** For all three time periods, women who completed self-administered questionnaires reported more sexual partners than women who responded to questions administered by an interviewer. There were significant effects for the method of administration for reported partners during the past year, the past five years, and the respondent's lifetime. For the past year, the women who answered self-administered questions reported a mean of 1.72 sexual partners versus 1.44 for those who answered questions administered by an interviewer ( $F(1,39) = 9.30, p < .01$ ). For the five-year period, women who completed self-administered questionnaires reported a mean of 3.87 sexual partners versus 2.82 for those who answered interviewer-administered questions ( $F(1,39) = 5.74, p < .05$ ). For the lifetime item, women who completed self-administered

questionnaires reported a mean of 6.51 sexual partners versus 5.43 for those who answered questions administered by an interviewer ( $F(1,39) = 9.54, p < .01$ ). No other main effects were significant.

Computerization seemed to interact with the site of the interview to affect the number of sexual partners reported. During home interviews, more sexual partners were reported by women interviewed using computer-assisted questionnaires than by those responding to conventional, paper-and-pencil questionnaires; for women interviewed outside the home, more sexual partners were reported on the pencil-and-paper questionnaires. Table 3 displays the relevant means. For the previous year, women interviewed at home reported fewer sexual partners on the paper-and-pencil questionnaires than on the computer-assisted ones (1.36 vs. 1.84), whereas the women interviewed outside the home reported more partners on the paper-and-pencil than on the computer-assisted questionnaires (1.68 vs. 1.43;  $F(1,39) = 7.72, p < .01$ ). Similarly, for the lifetime partners question, women interviewed at home reported fewer partners on the paper-and-pencil than on the computer-assisted questionnaires (5.06 vs. 7.48), whereas those interviewed outside their homes showed the opposite pattern, reporting more partners on the paper-and-pencil than on the computer-assisted questionnaires (6.26 vs. 5.08;  $F(1,39) = 5.89, p < .05$ ). The pattern is in the same direction but not significant for the five-year partners item. Overall levels of reporting are consistently higher using computer-assisted questionnaires, although not significantly so. Bringing computers into the respondents' homes may have fostered a sense of the importance or objectivity of the survey, promoting fuller reporting of sexual partners. Outside the home, especially in public places, the computer may make respondents feel conspicuous, inhibiting reporting.

Table 3. Sexual Partners by Mode and Site

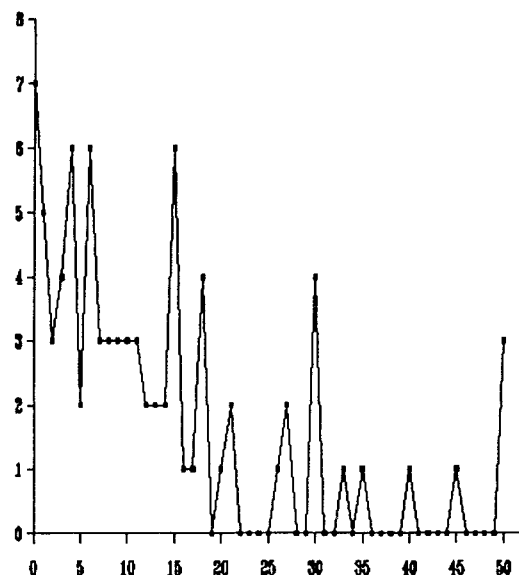
|            | At Home |          | Outside the Home |          |
|------------|---------|----------|------------------|----------|
|            | Paper   | Computer | Paper            | Computer |
| One Year   | 1.36    | 1.84     | 1.68             | 1.43     |
| Five Years | 2.81    | 4.51     | 3.33             | 2.74     |
| Lifetime   | 5.06    | 7.48     | 6.26             | 5.08     |

Note: Means based on untransformed counts.

**Males vs. females.** As has been observed in earlier surveys on sexual behavior, the men reported more opposite-sex sexual partners than the women did. This was true for the past year (4.19 for the men versus 1.58 for the women), the past five years (12.47 vs. 3.34), and lifetime (23.96 vs. 5.97); all three differences are highly significant ( $F$  values all greater than 10;  $p$  values all less than .001). In the analyses that include the data for men, the main effect of self-administration remains significant and that variable does not interact with sex. (The sex of the respondent did occasionally enter into higher-order interactions with the experimental variables, but none of these interactions was readily interpretable.)

**Rounding of values.** Morris (1993) has argued that the discrepancy between men and women in the reported number of sexual partners largely reflects differences within the subgroup of respondents with a relatively large number of partners to report; within this subgroup, the differences between men and women may reflect differences in rounding behavior (with the women rounding their answers down and the men rounding theirs up). Figure 1 shows the distribution of the number of lifetime sexual partners reported by the men in our study. (The results for the women, which are not shown, are quite similar.) The preponderance of reported values that are exact multiples of five strongly suggests that respondents of both sexes are reporting their answers in round numbers. Over 57.2% of the 145 respondents who reported 11 or more sexual partners gave an answer that is a multiple of five.

Figure 1. Distribution of lifetime sexual partners



### 3.2 Other Sexual Topics

Women who answered self-administered questions reported more sexually transmitted diseases (STDs) than those answering questions administered by an interviewer (22.0% vs. 17.0%). This effect of the method of administration was only marginally significant ( $\chi^2_1 = 2.93, p < .10$ ). No other main effects or interactions were significant. Results were the same for both the logistic regression models and chi-square tests.

We analyzed the ratio between two items concerning condom use and sexual intercourse in the last 30 days; the ratio represented the percentage of time the respondent used a condom in the past month. We performed an analysis of variance to examine this variable. Women who reported that they had not had sexual intercourse in the last 30 days were dropped from this analysis; data from 641 women were included in the analysis. Significantly more condom use was reported with self-administered questionnaires (average reported use 47% of the time) than with interviewer-administered questions (35%). The main effect for the method of administration variable was significant ( $F(1,39) = 8.18, p < .001$ ). We found no other significant effects on the condom use variable.

### 4. Discussion

**Effects of self-administration.** The variable with the most consistent impact on the level of reporting was the method of administering the questions. Women who completed self-administered questionnaires reported more sexual partners, more sexually transmitted diseases, and greater use of condoms than those who responded to questions read by an interviewer. These findings are summarized in Table 4, which displays the ratio between the levels of reporting under the self-administered and interviewer-administered conditions. As the table shows, the levels of reporting are substantially higher--from 19 to 37% higher--when the questions are self-administered. The effects of self-administration are similar for men. The lack of effects for the site of the interview suggests that respondents may be more concerned about the reactions of the interviewer than about the threat of other family members overhearing.

Table 4. Reported sexual behavior

|                           | Method of Administration |             |       |
|---------------------------|--------------------------|-------------|-------|
|                           | Self                     | Interviewer | Ratio |
| <b>Sexual Partners</b>    |                          |             |       |
| Past Year--Women          | 1.72                     | 1.44        | 1.19  |
| Past Year--Men            | 4.52                     | 3.88        | 1.16  |
| Past 5 Years--Women       | 3.87                     | 2.62        | 1.37  |
| Past 5 Years--Men         | 14.72                    | 10.43       | 1.41  |
| Lifetime--Women           | 6.51                     | 5.43        | 1.20  |
| Lifetime--Men             | 22.76                    | 25.00       | 0.91  |
| <b>Condom Use (Women)</b> |                          |             |       |
| Past 30 Days              | 46.7%                    | 35.3%       | 1.32  |
| Past Year                 | 23.8%                    | 17.9%       | 1.33  |
| <b>STDs (Women)</b>       | 22.0%                    | 17.0%       | 1.29  |

**Effects of site and medical context.** In contrast to the clear results for self-administration, we observed few effects for the site of the interview. In addition, we observed no effects for either of our attempts to induce a medical context for the questions.

Several studies have attempted to observe the impact of the presence of other family members on reports of sensitive behaviors. For example, in two studies on illicit drug use reporting, interviewers noted whether other family members were present during the interview (Schober et al., 1992; Lessler et al., 1992); neither study found an effect of this variable on reported drug use. Mosher and Duffer (1994), on the other hand, report an effect for the site of the interview on abortion reporting. It may be that the effects of this variable are hard to observe consistently. As we already suggested, respondents may be worried less about the reactions of other household members than about those of the interviewer. In addition, respondents may live alone, or with others (e.g., infants) whose presence is not a cause for concern. Such circumstances will reduce the impact of the site of the interview and make it difficult to demonstrate the effect of this variable.



Neither the version of the questionnaire nor the type of interviewer collecting the data had any discernible effects on reporting. These variable may have made little impression on the respondents. The nurses did not wear distinctive uniforms and, although they introduced themselves as nurses, this fact probably did not remain very salient to the respondents as the interview progressed. It is also quite possible that respondents see nurses and other medical personnel as authority figures and are no more willing to make embarrassing revelations to them than to ordinary survey interviewers. Several studies demonstrate that respondents admit more sensitive behaviors in a self-administered questionnaire than they do in interviews conducted by medical personnel (see, for example, Boekeloo et al., 1994; Locke et al., 1992); these results suggest that respondents withhold sensitive information from medical personnel, just as they do with field interviewers.

**Computerization.** We found that computerization by itself had no consistent effects on levels of reporting among the respondents. Instead, the effects of computer assistance seemed to vary somewhat by the topic of the question and the site of the interview (see Table 3). In reports on sexual partners, computer assistance seemed to increase the number of partners reported when data collection took place in the home but to reduce the number reported when data collection took place outside the home. We have no compelling explanation for this mode by site interaction.

Past investigations of computerized interviewing have tended to emphasize its effects on item nonresponse, timeliness, and cost rather than on the answers that are obtained. Only a few studies have reported effects of computer-assisted data collection on levels of reporting. The experiment comparing CAPI with conventional paper-and-pencil data collection on the National Longitudinal Study of Labor Market Behavior/Youth Cohort found that more respondents reported using birth control under CAPI than under paper-and-pencil interviewing (Baker and Bradburn, 1991). Several other studies have shown effects on reporting for computer-assisted self-administration, but in these studies, it is impossible to disentangle the effects of computerization from those of self-administration (e.g., Waterton and Duffy, 1984). We suspect that computerization by itself has little effect on the answers respondents give, a conclusion consistent with much of the previous literature on computer-assisted telephone interviewing (Groves & Mathiowetz, 1987).

## 5. References

- Abernathy, J., Greenberg, B., & Horvitz, D. (1970). Estimates of induced abortion in urban North Carolina. *Demography*, 7, 19-29.
- Aquilino, W., & LoSciuto, L. (1990) Effect of interview mode on self-reported drug use. *Public Opinion Quarterly*, 54, 362-395.
- Baker, R. P., & Bradburn, N. M. (1991). CAPI: Impacts on data quality and survey costs. Paper presented at the 1991 Public Health Conference on Records and Statistics, Washington, DC.
- Boekeloo, B., Schiavo, L., Rabin, D., Conlon, R., Jordan, C., & Mundt, D. (1994). Self-reports of HIV risk factors at a sexually transmitted disease clinic: Audio vs written questionnaires. *American Journal of Public Health*, 84, 754-760.
- Bradburn, N.M. (1983). Response effects. In P. Rossi, J. Wright, & A. Anderson (Eds.), *Handbook of Survey Research*. New York: Academic Press.
- Bradburn, N., Sudman, S., & Associates. (1979). *Improving interview method and questionnaire design: Response effects to threatening questions in survey research*. San Francisco: Jossey-Bass.
- Catania, J., Gibson, D., Chitwood, D., & Coates, T. (1990). Methodological problems in AIDS behavioral research: Influences on measurement error and participation bias in studies of sexual behavior. *Psychological Bulletin*, 108, 339-362.
- Groves, R., & Mathiowetz, N. (1987). A comparison of CATI and non-CATI questionnaires. In O. Thornberry, *An Experimental Comparison of Telephone and Health Interview Surveys*. Vital and Health Statistics. Series 2, No. 106.
- Hochstim, J. (1967). A critical comparison of three strategies of collecting data from households. *Journal of the American Statistical Association*, 62, 976-989.
- I-Cheng, C., Chow, L.P., & Rider, R.V. (1972). The randomized response techniques as used in the Taiwan outcome of pregnancy study. New York: A Publication of the Populations Council 33, No.11: 265-269.

- Jobe, J. B., Pratt, W. F., Tourangeau, R., Baldwin, A., & Rasinski, K. (In press). Effects of interview mode on sensitive questions in a fertility survey. In L. Lyberg et al. (Eds.), *Survey measurement and process quality*. New York: Wiley.
- Jobe, J. B., Tourangeau, R., & Smith, A.F. (1993). Contributions of survey research to the understanding of memory. *Applied Cognitive Psychology*, 7, 567-584.
- Jones, E., & Forrest, J. (1992). Underreporting of abortions in surveys of U.S. women: 1976 to 1988. *Demography*, 29, 113-126.
- Locke, S.E., Kowaloff, H.B., Hoff, R.G., Safran, C., Popovsky, M.A., Cotton, D.J., Finkelstein, D.M. Page, P.L., and Slack, W.V. (1992). Computer-based interview for screening blood-donors for risk of HIV transmission. *Journal of the American Medical Association*, 29, 1301-1305.
- London, K., & Williams, L. (1990, May). A comparison of abortion underreporting in an in-person interview and self-administered questionnaire. Paper presented at the Annual Meeting of the Population Association of America, Toronto, Canada.
- Lucas, R.W., Mullen, P.J., Luna, C.B.X., & McNroy, D.C. (1977). Psychiatrist and computer interrogators of patients with alcohol-related illnesses: A comparison. *British Journal of Psychiatry*, 131, 160-167.
- Marquis, K., & Cannell, C. (1971). Effect of some experimental techniques on reporting in the Health Interview. *Vital and Health Statistics*. National Center for Health Statistics. DHEW Publication No. 1000, Series 2, No. 41. Washington, D.C.: U.S. Government Printing Office.
- Morris, M. (1993). Telling tails explain the discrepancy in sexual partner reports. *Nature*, 365, 437-440.
- Mosher, W. D., & Duffer, A. P., Jr. (1994, May). *Experiments in survey data collection: The National Survey of Family Growth Pretest*. Presented at the meeting of the Population Association of America, Miami.
- Mott, F. (1985). *Evaluation of fertility data and preliminary analytic results from the 1983 survey of the National Longitudinal Surveys of Work Experience of Youth*. A report to the National Institute of Child Health and Development by the Center for Human Resources Research, January, 1985.
- Rieger, S., Judkins, D., & Sperry, S. (1991). *National Survey of Family Growth: Cycle IV CATI Phase. Final Report*. Hyattsville, MD: National Center for Health Statistics.
- Robinson, R., & West, R. (1992). A comparison of computer and questionnaire methods of history-taking in a genito-urinary clinic. *Psychology and Health*, 6, 77-84.
- Schober, S., Caces, M. Pergamit, M. & Branden, L. (1992). Effects of mode of administration on reporting of drug use in the National Longitudinal Survey. In C. Turner, J. Lessler, & J. Gfroerer (Eds.), *Survey Measurement of Drug Use: Methodological Studies* (pp. 267-276). Rockville, Md.: National Institute on Drug Abuse.
- Shimizu, I., & Bonham, G. (1978). Randomized response technique in a national survey. *Journal of the American Statistical Association*, 73, 35-39.
- Smith, T. 1992. Discrepancies between men and women in reporting number of sexual partners: A summary from four countries. *Social Biology*, 39, 203-211.
- Turner, C., J. Lessler, & Devore, J. (1992). Effects of mode of administration and wording on reporting of drug use. In C. Turner, J. Lessler, & J. Gfroerer, (Eds.), *Survey Measurement of Drug Use: Methodological Studies* (Pp. 177-220). Rockville, Md.: National Institute on Drug Abuse.
- Warner, S. (1965). Randomized response: A survey technique for eliminating evasive answer bias. *Journal of the American Statistical Association*, 60, 63-69.
- Waterton, J. & Duffy, J. (1984). A comparison of computer interviewing techniques and traditional methods for the collection of self-report alcohol consumption data in a field survey. *International Statistical Review*, 52, 173-182.