Application of Computer-assisted Interviews to Sexual Behavior Research

Patricia Kissinger,[1] Janet Rice,[2] Thomas Farley,[3] Shelly Trim,[1] Kayla Jewitt,[1] Victor Margavio,[1] and David H. Martin[1]

Collection of sensitive data with the use of video-enhanced, computer-assisted, self-administered interviews (V-CASI) has the potential to reduce interview bias and improve the validity of the study. The purpose of this study was to compare responses to sensitive questions elicited by V-CASI and by face-to-face interview (FTFI) methods. Women attending a New Orleans, Louisiana, public family planning or sexually transmitted disease clinic from July 1995 to July 1996, diagnosed with a *Chlamydia trachomatis* infection responded to eight close-ended behavioral questions (four socially undesirable, two socially desirable, and two neutral behaviors) using both FTFI and V-CASI techniques in a randomized crossover design. Of the 280 women included, the mean age was 23 years, 95 percent were African American, and 71 percent felt comfortable using computers. While kappa scores indicated good-to-excellent agreement between interview techniques, women tended to admit to socially undesirable behaviors more often on V-CASI compared with FTFI. Thirty percent of the women gave a discrepant response between V-CASI and FTFI toward social desirability. Women who reported a socially undesirable behavior in V-CASI (i.e., more than two sex partners and infrequent condom usage) were more likely to have a discrepant response. Utilization of the same logistic regression model to predict condom use yielded different results when data from V-CASI were used compared with data from FTFI. The V-CASI technique can reduce social desirability bias and improve validity in research requiring information on sensitive sexual behaviors. *Am J Epidemiol* 1999;149:950–4.

interviews; sex behavior

A major challenge of survey-based research is to reduce bias in order to produce more valid results. When the behaviors being reported are sensitive, reporting biases become more likely because the respondents may deliberately under- or overreport the frequency of these behaviors.

Research suggests that computer-assisted, self-administered interviews (CASI) may produce more valid reports of sensitive behaviors than would more traditional survey techniques such as face-to-face interviews (FTFIs) (1, 2). The use of CASI has been used in clinical as well as research settings (3–5) and provides several potential methodological and logistic advantages. This technique allows standardization of the way in which questions are asked and who is ask-

ing them, and it eliminates interviewer interpretation of responses. The perceived anonymity of this type of interview may make respondents feel more at ease in reporting behaviors that are socially undesirable and less likely to embellish responses for socially desirable behaviors. The inclusion of a video component to the CASI (i.e., V-CASI) can also personalize the interview and facilitate responses by persons with a low literacy level. The ability to program the data collection immediately into a database can remove interviewer errors that may occur with complex skip questions. These advantages can reduce error and improve the validity of responses to sensitive questions.

It is possible, however, that V-CASI may be confusing or cumbersome for study participants. Complicated or lengthy questions may require explanation and clarification by an interviewer and may not be appropriate for the V-CASI method.

While researchers are using CASI more frequently, comparisons of this technique with the more traditional FTFIs have rarely been conducted. The purpose of this study was to determine whether sensitive information is better elicited using V-CASI compared with FTFI. Three hypotheses were examined specifically: whether women would answer yes to less socially desirable behaviors more often on V-CASI compared

with FTFI, which factors were associated with giving a discrepant response between V-CASI and FTFI to maintain social desirability, and whether multivariable models based on the two data sources would generate different results.

MATERIALS AND METHODS

Women attending a New Orleans, Louisiana, public family planning or sexually transmitted disease clinic from July 1995 to July 1996, who were enrolled in an infertility prevention cohort study and who provided consent were included. Enrollment criteria for the cohort study were ages 15–50 years; *Chlamydia trachomatis*-positive by GenProbe (San Diego, California), Chlamydiazyme (Abbott Laboratories, North Chicago, Illinois) or urine ligase chain reaction test; and no prior treatment for that episode of *C. trachomatis* infection. Women were administered a closed-ended survey via both FTFI and touch-screen V-CASI in a crossover design. Eight questions (four socially undesirable, two socially desirable, and two neutral behaviors) were chosen for analysis. Women were randomized to receive either FTFI or V-CASI first by using a randomized systematic allocation method. Upon completion of that interview, the other type of interview was immediately administered.

The questionnaire was developed by a multicenter group of investigators from New Orleans, Louisiana; Birmingham, Alabama; Seattle, Washington; San Francisco, California; and Indianapolis, Indiana (6), working on the Centers for Disease Control and Prevention Infertility Prevention Study. The purpose of that study was to determine factors associated with recurrent *C. trachomatis* infection. According to the SMOG technique (7), the survey was at a fourth-grade reading level.

Face-to-face interviews

The FTFI was administered by a study staff composed of six female interviewers aged 21–30 years, of whom four were African American and two were Caucasian. All of the interviewers received training using a multicentered protocol, and the quality of interview technique was monitored through weekly meetings and data integrity checks.

Computer interview

Toolbook development software (8) was used to create the video-assisted questionnaire with questions used verbatim from the FTFI questionnaire. Possible responses were printed on the screen and read by an audible voice with a video picture of the reader (a

young African-American woman) in the top right corner of the screen, and responses were highlighted as read to assist the viewer.

The respondents chose their answers either by touching the touch-sensitive video monitor or by using the computer mouse. An introductory question was included for training purposes, allowing the participant to practice operating the computer before beginning the real questionnaire. A response was required for each question before the subject could go to the next question. Study personnel were available to respond to any questions about the video questionnaire.

Analysis

The percentages of affirmative responses to socially undesirable, desirable, and neutral questions on V-CASI were compared with FTFI responses by using matched odds ratios and McNemar's chi-square test. Socially desirable responses were defined as a negative response to having had more than two sex partners or a new sex partner in the last 6 months, a negative response to having a history of gonorrhea or douching, and a positive response to either of two condom use questions (i.e., used a condom the last vaginal sex act and used a condom half of the time or more). Kappa scores were calculated to estimate the extent of agreement between the two interview modalities.

The first set of analyses compared the two survey modes and used conditional methods of analysis. In the next set of analyses, predictors of discrepant versus consistent responses were explored. Unconditional methods for these analyses were used since we were comparing two groups of respondents. Women were considered concordant responders only if they gave consistent responses to all behavioral questions. Covariates considered for the logistic regression model were age and behavioral variables (i.e., questions on sex partner and condom use, douching, and history of gonorrhea). The behavioral variables used came from the V-CASI because it was assumed from data in the table 1 that these answers may be closer to reality. Confounders included in the model were interview modality administered first and comfort with computers.

Finally, to determine whether responses from V-CASI would yield different effect measures and results than those from FTFI, two logistic regression analyses were conducted. The hypothesis was that self-report bias would add enough error to the findings that measures of association would be altered. The outcome variable was the use of condoms less than or equal to half of the time. Predictors considered were more than two sex partners, age, and history of gonorrhea infection. These variables were chosen for illustration

TABLE 1.   Assessment of socially desirable, undesirable, and neutral questions by interview modality (n = 280), New Orleans, Louisiana, 1995–1996

	V-CASI† (% responding yes)	FTFI† (% responding yes)	Crude OR‡	Kappa§
Socially undesirable				
>2 sex partners¶	6.1	4.4	1.42	0.70
New sex partner¶	12.5	11.4	1.12	0.72
History of gonorrhea	20.7	17.4	1.23	0.86
Douched	52.3	49.6	1.12	0.93
Socially desirable				
Used condom last vaginal intercourse	27.4	28.3	0.96	0.69
Used condom more than half of the time	28.3	39.9	0.60*	0.49
Neutral				
Pregnancy	13.2	12.8	1.04	0.95
Symptomatic at this visit	43.1	46.3	0.88	0.54

* McNemar's test, $p < 0.01$.
† V-CASI, video-enhanced, computer-assisted, self-administered interviews; FTFI, face-to-face interviews.
‡ Matched pairs odds ratios (OR).
§ κ > 0.75 denotes excellent reproducibility, 0.40 ≥ κ ≤ 0.75 denotes good reproducibility, and <0.40 denotes poor reproducibility.
¶ In the previous 2 months.

rather than for an actual analysis of condom usage. The V-CASI model used data from V-CASI only, and the FTFI model used data from the FTFI only.

RESULTS

Of 280 subjects interviewed, 50 percent received the FTFI first and 50 percent received the V-CASI first. All women completed both interviews, and no woman refused to complete the survey after providing consent and being randomized. The study group was 95 percent African American, 5 percent Caucasian, and 1 percent other race; 20 percent age 14–17 years, 50 percent age 18–24 years, and 30 percent age 25–50 years; and 63 percent recruited at the family planning clinic and 37 percent recruited from a sexually transmitted disease clinic. Self-reported comfort level with computers was: very comfortable, 39 percent; comfortable, 32 percent; somewhat comfortable, 24 percent; and not comfortable at all, 5 percent. The subjects' age and race were not associated with comfort with computers. Median time to complete both interviews was 5 minutes per interview. Those randomized to V-CASI first were similar to those randomized to FTFI first with respect to age, race, and comfort level with computers.

Subjects tended to answer yes more often on V-CASI for socially undesirable behaviors and yes more often on FTFI for socially desirable behaviors; that is, for all questions, FTFI tended to elicit socially desirable responses more often than did V-CASI. Women were statistically more likely to report condom use more than half of the time on FTFI compared with V-CASI (odds ratio = 0.60) (table 1).

Thirty percent of the women had at least one discrepant response toward social desirability. Age and comfort with computers were not associated with giving a discrepant response. Logistic regression was conducted to determine factors associated with at least one discrepant response toward social desirability. Women who had more than two sex partners and who used condoms half the time or less were more likely to have a discrepant response toward social desirability after adjustment for age and comfort with computers (table 2).

TABLE 2.   Factors associated with discrepant responses toward social desirability in logistic regression (n = 280),† New Orleans, Louisiana, 1995–1996

	OR‡	95% CI‡
Age <18 vs. ≥18 years	1.41	0.69–2.90
V-CASI‡ first vs. FTFI‡ first	1.16	0.65–2.09
Comfortable with computers vs. not comfortable	0.61	0.32–1.15
Had >2 sex partners	5.31*	1.34–21.1
Had a new sex partner	1.58	0.59–4.22
Did not use a condom during last vaginal sex	0.98	0.45–2.12
Used condoms half of the time or less	2.65*	1.16–5.88
Had a history of gonorrhea	1.81	0.88–3.72
Douched	0.76	0.42–1.36

* $p < 0.01$.
† Predictor and confounder information was taken from V-CASI.
‡ OR, odds ratio; CI, confidence interval; V-CASI, video-enhanced, computer-assisted, self-administered interviews; FTFI, face-to-face interviews.

Results from the same logistic regression model with data from V-CASI and from FTFI were compared. While age was significant in both models, effect measures for all of the predictors were larger in the FTFI model. In addition, having had more than two sex partners was associated with using a condom half the time or less in the V-CASI model, but not in the FTFI model (table 3).

DISCUSSION

Our findings provide support that the V-CASI technique may reduce social desirability bias and improve validity in sexual behavior research. We found that 30 percent of the respondents tended to answer in a more socially desirable manner in FTFI compared with V-CASI. Women who reported behaviors that were less socially desirable were more likely to have this discrepancy. While it is not clear why respondents answered in a more socially desirable manner in V-CASI, it is likely that the perception of anonymity in the V-CASI may have removed respondents' inhibitions. Tourangeau et al. (2) summarized research finding from several studies and found that anonymity improves accuracy of response.

The purpose of the analysis shown in table 3 was to determine whether logistic regression models predicting factors associated with condom use half the time or less would yield different results (i.e., whether self-report bias for this sensitive question varied by modality). Higher effect measures were found for all but one predictor in the FTFI model. In addition, the finding that more than two sex partners was associated with less condom use (an important clinical and public health finding) was missed in the FTFI model. Since no "gold standard" has been established between the two methods, one cannot say which model is correct. However, one can appreciate that self-report bias does

cause enough error to alter effect measures. It is interesting that the FTFI has some stronger associations. We usually think error dampens associations.

Other studies comparing CASI with other techniques have examined these techniques using different groups (1, 4, 9). The use of the experimental crossover design in our study allowed for direct comparisons of the techniques on the same people. This reduces the chance that the differences were due to random variability. Because we randomized the order in which the interview type was received, it is unlikely that the differences we found were due to the order of the interviews.

V-CASI can remove the potential bias that is created by variable reading skills. Approximately 32 million adults in the United States report speaking a language other than English in their home (10), and another 44 million (or 23 percent of the adult population) have difficulty with rudimentary reading and writing (11). V-CASI can help to include these populations who might otherwise be eliminated from research. To accommodate a diverse study population, computer-assisted techniques can promote multilingual questionnaires asked in a standard manner.

Logistically, V-CASI can be appealing. It can decrease staffing requirements and eliminate the need for data entry. The price of the V-CASI technique is decreasing. Multimedia software and hardware equipment are becoming less expensive and more sophisticated. The touch-screen monitor (which was the most costly item in this study) was not necessary, since many of our respondents chose to use the computer mouse. This was particularly true for those who felt more comfortable with computers. In this study, video segments were recorded by study staff using a regular camcorder. While professional quality of the video segments was not required, good lighting and attention to background color and detail were important. Total time to produce the video segments was about 5 workdays, and total programming time was several weeks. Maintenance of the computer consisted of weekly calibration of the screen to adjust for touch sensitivity and data backup on diskette. While the video picture made the interview more personal, it also was not necessary. Several studies have used simple audio techniques or less sophisticated computer techniques with good results (9, 12–17).

Many of the participants remarked that they liked the innovation of the computer questionnaire. We conducted four focus groups to explore how the women felt about the questionnaire, and the response was overwhelmingly positive for this technique. In addition, the majority of study participants (71 percent) reported that they were very comfortable or were comfortable with computers. This indicates that computer

TABLE 3. Logistic regression results for factors associated with using a condom half the time or less by interview modality (n = 280),† New Orleans, Louisiana, 1995–1996

	V-CASI		FTFI	
	OR‡	95% CI‡	OR	95% CI
Age (years)				
14–17	2.86	1.25–6.55	5.20	2.30–11.78***
18–24	2.22	1.09–4.55**	4.25	2.13–8.50***
≥25	1.00		1.00	
Had >2 sex partners	0.49	0.24–1.02*	0.80	0.36–1.75
Prior gonorrhea	1.08	0.81–1.45	1.24	0.94–1.63

* p < 0.10; ** p < 0.05; *** p < 0.01.
† Video-enhanced, computer-assisted, self-administered interview (V-CASI) model used data from V-CASI only. Face-to-face interview (FTFI) model used data from FTFI only.
‡ OR, odds ratio; CI, confidence interval.

skill is unlikely to be a barrier to V-CASI use in this population.

V-CASI does not remove the requirement to design and pilot test a survey well. CASI may not be appropriate for all types of survey research. For example, some questions may require explanation or special probing techniques to help the respondent understand or remember. Eisenhower et al. (18) identify forgetting and misplacing events in time (telescoping) as two sources of recall bias. Dillman and Ternai (19) report that asking respondents to recall details associated with an event increases its recall. This type of interviewing may be too cumbersome for the V-CASI technique. In general, complicated questions or those that require probing may not be well suited to V-CASI.

In addition, complicated questions may also not be appropriate for the V-CASI format. The format of the question and the context in which it is asked can affect the response. We noted a greater difference between V-CASI and FTFI for the more conceptual condom question (i.e., how often do you usually use condoms?) compared with the more concrete condom question (i.e., the last time you had sex, did you use a condom?). This suggests that more complicated questions may require the explanation of an interviewer.

The findings of our study provide support for the idea that the V-CASI technique can help to remove social desirability bias and improve validity in research, eliciting information on sensitive sexual behaviors. More evaluation of this technique is needed.

ACKNOWLEDGMENTS

REFERENCES

1. Turner CF, Ku L, Rogers SM, et al. Adolescent sexual behavior, drug use, and violence: increased reporting with computer survey technology. Science 1998;280:867–73.
2. Tourangeau R, Smith T. Asking sensitive questions: the impact of data collection mode, question format, and question context. Public Opinion Q 1996;60:275–304.
3. Kohlmeier L, Mendez M, McDuffie J, et al. Computer-assisted self-interviewing: a multimedia approach to dietary assessment. Am J Clin Nutr 1997;65 (Suppl. 4):1275S–81S.
4. Lessler JT, O'Reilly JM. Mode of interview and reporting of sensitive issues: design and implementation of audio computer-assisted self-interviewing. NIDA Res Monogr 1997;167:366–82.
5. Kobak KA, Taylor LH, Dottl SL, et al. A computer-administered telephone interview to identify mental disorders. JAMA 1997;278;905–10.
6. Kent CK, Litchfield W, Handsfield H, et al. Factors associated with successful self-referral of male sexual partners for treatment of contact to chlamydia. Oral presentation at the 1996 National STD Prevention Conference, Tampa, FL, December 1996.
7. McGraw HC. US Health and Human Services. National Institute of Health. Make Health Programs Work: a planners guide. Baltimore, MD: Office of Education, pp. 77–9.
8. Toolbook software. Bellview, WA: Asymetrix, 1992.
9. Education and Social Stratification Branch, population division, US Bureau of Census. Languages spoken at home and ability to speak English for United States, regions, and states, 1990. Special tabulation 1990 CPH-L-133. Washington, DC: US GPO, 1993.
10. US Department of Education. National Adult Literacy Survey, 1992 Washington, DC: Department of Education, 1993.
11. Nebot M, Celentano D, Burwell L, et al. AIDS and behavioural risk factors in women in inner city Baltimore: a comparison of telephone and face to face surveys. J Epidemiol Community Health 1994;48:412–18.
12. Boekeloo BO, Schiavo L, Rabin DL, et al. Self-reports of HIV risk factors by patients at a sexually transmitted disease clinic: audio vs. written questionnaires. Am J Public Health 1994; 84:754–60.
13. Hasley S. A comparison of computer-based and personal interviews for the gynecologic history update. Obstet Gynecol 1995;85:494–8.
14. Maitland ME, Mandel AR. A client-computer interface for questionnaire data. Arch Phys Med Rehabil 1994;75:639–42.
15. Solomon GL, Dechter M. Are patients pleased with computer use in the examination room? J Fam Pract 1995;41:241–4.
16. Turner CF, Rogers SM, Hendershot TP, et al. Improving representation of linguistic minorities in health surveys. Public Health Rep 1996;111:276–9.
17. Locke SE, Kowaloff HB, Hoff RG, et al. Computer-based interview for screening blood donors for risk of HIV transmission. JAMA 1992;268:1301–5.
18. Eisenhower D, Mathiowetz N, Morganstein D. Recall error: sources and bias reduction techniques. In: Biemer P, Groves R, Lyberg L, et al. Measurement errors in surveys. New York, NY: John Wiley & Sons, Inc., 1991.
19. Dillman D, Tarnai J. Mode effects of cognitively designed recall questions: a comparison of answers to telephone and mail surveys. In: Biemer P, Groves R, Lyberg L, et al., eds. Measurement errors in surveys. New York, NY: John Wiley & Sons, Inc., 1991.