TECHNICAL PAPERS ON
HEALTH AND BEHAVIOR MEASUREMENT

TECHNICAL PAPER 28

Audio-CASI: The Impact of Operational Characteristics on Data Quality

Susan M. Rogers, Heather G. Miller, Barbara H. Forsyth,
Timothy K. Smith, and Charles F. Turner

Reference Citation

Rogers, S.M., H.G. Miller, B.H. Forsyth, T.K. Smith, and C.F. Turner. (1996) Audio-CASI: The impact of operational characteristics on data quality. In *1996 Proceedings of the Section on Survey Research Methods* (ASA/AAPOR). Alexandria, Va.: American Statistical Association.

AUDIO-CASI: THE IMPACT OF OPERATIONAL CHARACTERISTICS ON DATA QUALITY

Susan M. Rogers, Heather G. Miller, Barbara H. Forsyth, Timothy K. Smith, and Charles F. Turner
Susan M. Rogers, Research Triangle Institute, 6101 Executive Blvd., Rockville, MD 20852

Key Words: Survey methods, audio-CASI, data quality, measuring sensitive behaviors

INTRODUCTION

The past five years have seen the development of several automated self-interviewing systems that portend an important shift in the way surveys of sensitive topics will be conducted in the 21st century. Those systems are

- video computer-assisted self-interviewing (video-CASI);
- audio computer-assisted self-interviewing (audio-CASI), a modification of video-CASI that adds audio features; and
- telephone audio-CASI, a modification of audio-CASI that permits interviewing by telephone.

During 1995, three large-scale surveys were fielded using audio-CASI technology: the 1995 National Survey of Family Growth, the 1995 National Survey of Adolescent Males, and the Adolescent Health Survey. In addition, the Substance Abuse and Mental Health Services Administration recently announced plans to convert the National Household Survey on Drug Abuse to an audio-CASI format. The number of such conversions of other surveys of sensitive behaviors is likely to accelerate during the remainder of the decade.

The fast-paced adoption of audio-CASI technology stems from its benefits: it affords a fully private, standardized interview even for respondents who are not literate in any language. Although the technology's adoption is gratifying to those who have been its early advocates, many implementation issues are poorly understood. In particular, the ways in which its implementation may affect data quality have not been adequately identified and assessed.

This paper reports the results of a field study that examined how key details of audio-CASI implementation influence the quality of audio-CASI measurements and the types of sensitive measurements that are most likely to be affected. Results were derived from a methodological survey ($N = 194$) of a sample of households in Baltimore County, Maryland. The survey varied two elements of audio-CASI implementation: the presentation of questions in video as well as audio format (versus presentation in audio-only mode) and the gender of the spoken voice. We assessed the impact of those variations and of the rated sensitivity of the questions on the quality of data obtained during the field interviews. Indicators of data quality were (1) the timing of responses relative to the reading of the survey questions and response categories and (2) rates of item nonresponse.

METHODS

For this study, we drew a sample of households with listed telephones in Baltimore County, Maryland (a suburb of the city of Baltimore). A quota sampling design was employed that targeted a population of 100 adult men and 100 adult women between 18 and 45 years of age. Eight interviewers made personal contact with 818 households; 453 households did not meet the sample requirements (primarily because of age ineligibility); 105 eligible respondents refused to participate; and in 56 instances, the eligible household member was unavailable or not at home. Before beginning the survey, each participant gave his or her informed consent. Completed interviews with timing information for the survey questions were collected from 194 respondents.

The audio-CASI systems allowed respondents to complete their questionnaires in conditions that maximized privacy. After a respondent was recruited and instructed in how to use the audio-CASI system, the interviewer retreated to a position from which the computer screen was not visible. In training all the interviewers, we stressed the importance of complete privacy for maintaining confidentiality of the data.

The questionnaire consisted of items selected from national surveys on important public health issues, including AIDS and drug use. It was divided into two parts. The first part comprised 57 questions on drug and alcohol use, contraceptive behaviors and condom use, number and characteristics of sexual partners, knowledge and history of sexually transmitted diseases (STDS), communication with partner about sex and contraceptive use, and sociodemographic characteristics of the respondent. The second section contained an additional 46 questions that collected information on the respondent's sexual practices and orientation, prior experiences with violence, history of arrest, psychological problems, and attitudes about gender and race.

The field experiment tested two variations in audio-CASI implementation. First, the survey instrument was recorded twice--once in a male voice and once in a female voice. Eligible subjects were randomly assigned to one of the voices based on their identification number. A

programming error in the randomization process resulted in a 30-70 split in the allocation of subjects to the voice variable rather than an even 50-50 split. Consequently, 60 subjects heard the female voice and 134 heard the male voice.

Assignment to a survey mode was based on a crossover design in which the subject completed the first half of the questionnaire in one mode--either the audio-plus-video or audio-only mode--and then switched to the other mode for the second half of the questionnaire. In the audio-only mode, respondents could simply hear the recorded questions, whereas in the audio-plus-video mode, respondents read the questions on the computer screen *and* heard them. Fifty-eight respondents were assigned to the audio-plus-video mode for the first part of the interview and the audio-only mode for the second half. The remaining 136 respondents received the first section in audio-only mode and the second half in the audio-plus-video mode.

Immediately after completing each mode of presentation, the interviewers gathered data on the respondent's reactions to that mode. At the conclusion of the interview, the interviewer administered a paper-and-pencil questionnaire to collect information on subjects' general reactions to the survey questions and implementation procedures.

For this experiment, we analyzed the timing data for each question separately for the female and male recorded voices. In addition, the audio-CASI program provided timing data on the respondent. That is, for each respondent, we noted the time interval between the start of the audio and the time the respondent pressed a computer key in answer to the question being asked. The program also captured the number of times a respondent listened to each question and was able to distinguish between attempts to listen to the question again (playbacks) and attempts to change an answer (backups).

Using the question timing data, we determined how much of each question and each set of response categories the respondent heard before the keystroke indicating that a response was entered. We believe that this measure is a reasonable index of data quality, particularly for the audio-only mode, because respondents who enter keystrokes before hearing the full question may invite reporting error. We also investigated differences in the timing of responses between the male and female voices and the effects of the audio-only and audio-plus-video modes on data quality.

RESULTS

The results presented here come from 194 completed interviews. Respondents were equally distributed by gender (49 percent were male and 51 percent were female). Three-fourths of respondents were white, and 18 percent were black. One-third of the respondents had never married; more than half were currently married or living together. The respondents were relatively well off: nearly three-quarters (74 percent) had attended college, and almost one-third reported household income of more than $50,000.

Mode Preference Respondents who indicated a mode preference selected the audio-plus-video mode by a 9:1 ratio (see Table 1). In comparison with audio alone, respondents indicated that the audio-plus-video mode was

- easier to use and understand,
- more interesting,
- better for asking sensitive questions, and
- better for eliciting honest answers.

However, the video display was not without its drawbacks. The presentation of questions on the computer screen increased subjects' concerns about the privacy of their responses. Only 18 percent of respondents preferred the audio-plus-video mode for keeping answers private, although 39 percent were indifferent.

Mode preferences were related to some demographic

Table 1. Respondent mode preferences

	PREFERENCE		
DIMENSION	Audio+ Video (%)	Audio Only (%)	Indif-ferent (%)
Easier to understand	78	3	19
Easier to use	74	7	19
Most inter-esting to use	67	8	24
Better for asking sex questions	62	20	19
Better for asking about illegal activ.	57	18	26
Better for honest answers	45	16	39
Better for privacy	18	44	39
Method liked best	71	8	21

characteristics of the respondent. Whereas the majority of white subjects preferred to both read and listen to the survey questions, most of the black subjects did not state a preference. Preference for the audio-plus-video mode increased with education and income. Subjects who had attended college preferred the audio-plus-video mode to the audio mode alone, as did respondents living in households with higher income. Gender, age, and marital status had no significant effects on mode preference.

Aspects of Data Quality As shown in Table 2, it took respondents significantly longer to answer questions in the audio-only mode than in the audio-plus-video mode. Respondents in the audio-only mode took approximately four more minutes to complete the first half of the questionnaire, and we noted similar time differences for the second half of the questionnaire. Timing differences by gender of the recorded voice were minimal. However, the male voice tended to read the questions more slowly than did the female voice. Thus, questionnaires administered in the male voice took slightly longer to complete than those administered in the female voice.

Overall, a small proportion of survey questions were repeated; that is, respondents listened to or read the question or some portion of the question again before entering a response. Not unexpectedly, respondents repeated questions more often in the audio-only mode. Respondents seldom backed up to re-hear or re-read an earlier question; less than 1 percent of questions in either mode were backed up one or more times.

Table 2. Performance measures by mode of audio-CASI administration

Performance Measure		Audio + Video	Audio Only
Minutes to Complete Questionnaire			
Part 1:	Male voice	7.7	11.8*
	Female voice	7.1	11.3*
Part 2:	Male voice	8.4	13.6*
	Female voice	8.4	12.2*
Percentage of questions played 1+ times		2.3	3.8**
Percentage of questions backed up 1+ times		0.6	0.9*

Note: p values based on chi-square test, * $p < 0.01$; ** $p < 0.001$

We expected that the proportion of respondents who listened to the full question before entering a response would be an indicator of data quality, particularly for those in the audio-only mode. Data on timing of responses were compared by voice gender and mode of questionnaire administration. Based on preliminary analyses comparing data on the timing of responses by voice gender, gender does not appear to have a differential effect on responses. The analyses that follow focus on comparing data on timing of responses by mode.

For the majority of questions, respondents assigned to the audio-only mode listened to the entire question before responding. That pattern held true for both sensitive or personal questions as well as less sensitive questions and provided some indication of the quality of the data.

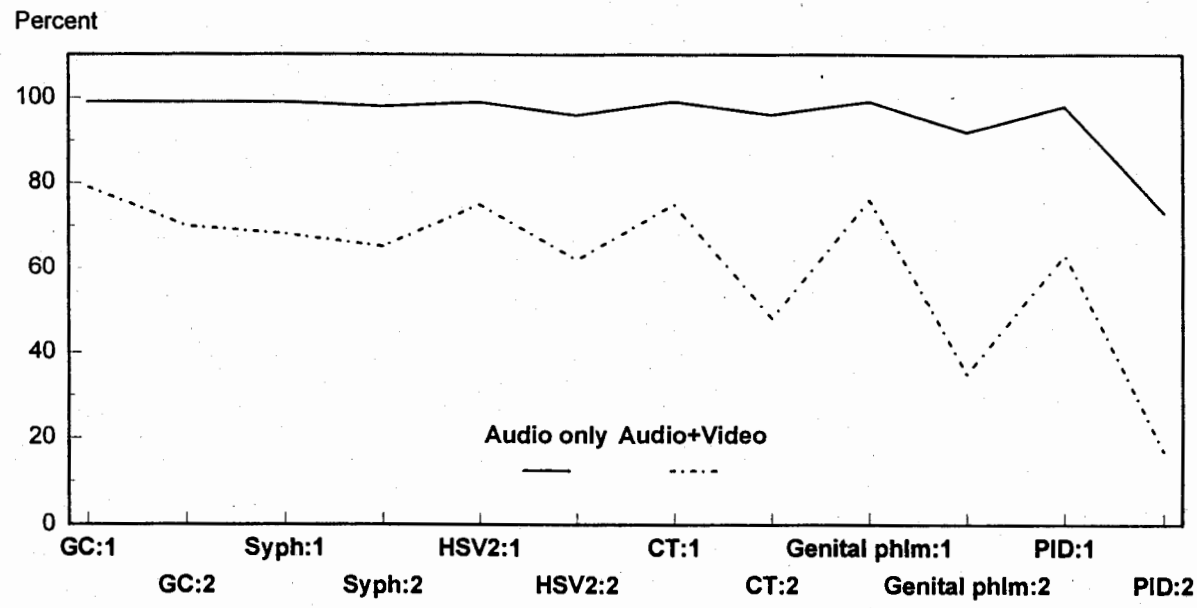
Several questions were open-ended including such items as respondent's age, number of children, number of abortions, and number of new sex partners. For example, the question "How many children do you have?" was followed by additional information to standardize the calculation across subjects. (After the question, respondents heard "Please include natural children, stepchildren, adopted children, and other children who live in your home and you are helping to raise. If the total is more than nine, please enter 9.")

The audio files show that respondents did not always listen to all of the qualifying statements before entering their answer. Unlike face-to-face interviews, respondents to an audio-CASI instrument may misinterpret pauses in the recorded voice and consider them completed questions; unlike respondents in the audio-plus-video mode, respondents in the audio-CASI mode cannot simultaneously view the questions and responses on the computer screen. Based on those observations, we determined that a more suitable format for these "extended" questions was needed. For example, qualifying statements could be placed before the question as an introduction or preface.

Most of the questions in the instrument were not open-ended, and many required a yes-or-no response. A series of 12 questions asked about knowledge and history of sexually transmitted diseases; the series comprised gonorrhea, syphilis, herpes, chlamydia, PID (pelvic inflammatory disease), and one disease-sounding entity, genital phlemoria. Each of the questions required a yes-or-no response. For example, respondents heard the question "Have you ever heard of a disease called chlamydia? If yes, press 1; if no, press 2." It was followed by "Has a doctor or nurse ever told you that you had chlamydia? If yes, press 1; if no, press 2."

Respondents quickly adapted to that pattern of questioning (see Figure 1). Those in the audio-only mode

Figure 1. Percentage of respondents who listened to the full survey question on a sexually transmitted disease before entering a response, by interview mode.



Note: GC, gonorrhea; Syph, syphilis; HSV2, herpes; CT, chlamydia; genital phlm, genital phlemoria (a fictitious disease); PID, pelvic inflammatory disease; 1 refers to the question on whether the respondent had heard of the disease, and 2 refers to the question on whether a doctor or nurse had ever told them they had the disease.

listened to both questions on gonorrhea at the beginning of the series. For questions at the end of the series, on PID, for example, all respondents heard the first PID question, "Have you ever heard of PID?" but only 73 percent listened to the second, "Has a doctor or nurse ever told you that you had PID?" in its entirety. That pattern of results was even more dramatic when respondents could read as well as hear the questions. (Interestingly, 16 percent of respondents reported that they had heard of the fictitious disease *genital phlemoria*, and 1 percent indicated that they had been told they had the "disease." Of the 58 percent of respondents who had heard of PID, 3.1 percent reported that they had had PID, and one of them was male.)

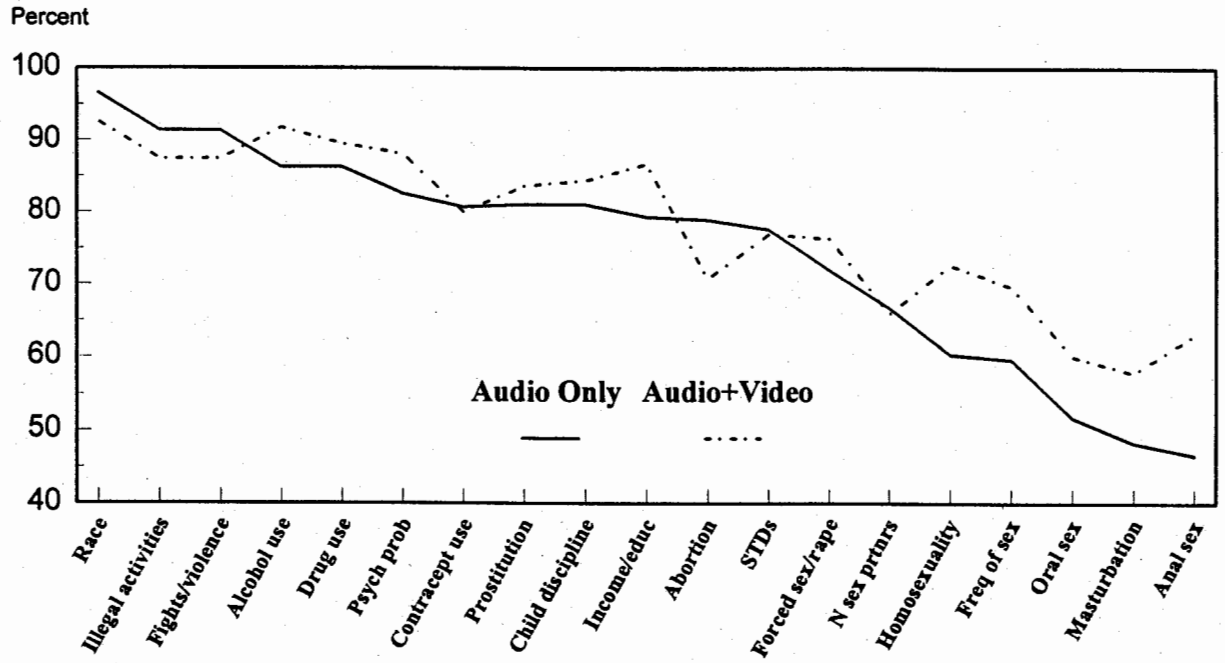
Item Sensitivity One of the goals of this study was to assess respondents' reactions to sensitive, personal questions administered in different presentation modes. One-third of respondents indicated that they had been reluctant to answer some questions truthfully, although only 2 percent reported that they had not given honest answers. Presumably, questions that respondents judged offensive or intrusive would be vulnerable to misreporting or nonresponse.

After completing the entire questionnaire, respondents were asked to rate certain categories of questions using the following scale: not at all offensive, slightly offensive, somewhat offensive, and very offensive. As illustrated in Figure 2, questions that did not deal with sexual behavior were judged least offensive. In other words, more respondents considered questions on race, violence, drugs, or alcohol use to be "not at all offensive" compared with questions on anal sex, oral sex, or masturbation.

The perceived offensiveness of some questions varied by mode of administration, and those differences increased with the sensitivity of the question being asked. For example, compared with respondents in the audio-plus-video mode, respondents in the audio-only mode were less likely to rate questions on oral sex, same-gender sex, and masturbation as "not at all offensive." For questions on anal sex, this difference was significant: 47 percent of respondents in the audio-only mode rated anal sex as not offensive versus 63 percent of respondents in the audio-plus-video mode.

Item Sensitivity and Nonresponse Presumably, questions that a respondent judges as more sensitive may be more difficult to answer and hence subject to nonreporting. Furthermore, one can assume that interview modes in which a respondent feels a lack of privacy will also invite nonreporting or errors in reporting. In this

Figure 2. Percentage of respondents who reported selected categories of questions as "not at all offensive," by interview mode.



experiment, it appears that the mode and perceived sensitivity of a question had little effect on item nonresponse rates (see Table 3). For the majority of questions, nonresponse rates were trivial; only six questions had nonresponse rates that were greater than 3 percent. Two of those questions were about masturbation. The first asked, "On average, in the past 12 months, how often did you masturbate?" (to which 4 percent [N = 8] did not respond). The second asked, "When masturbating in the past 12 months, how often did you reach orgasm, that is, come or come to climax?" (Of those who responded positively to the previous question, 4 percent refused to answer.) An additional question on racial attitudes registered a nonresponse rate of 5 percent. It asked, "If you could find the housing that you would want and like, would you rather live in a neighborhood of all black; mostly black; half black, half white; mostly white; or all white?" Two questions concerned the respondent's performance in the survey: "If you were tempted to mislead a survey interviewer about how often you had sex, would you be inclined to say that you had sex more often or less often than you really did?" and "If you were tempted to mislead a survey interviewer about how many sexual partners you had in the past 12 months, would you be inclined to say that you had more sex partners or fewer sex partners than you really did?" Twelve and 22 percent of respondents, respectively, did not answer those questions. (We suspect that some of that nonresponse arose because the questions forced respondents to indicate either more often or less often, more partners or fewer

Table 3. Percentage of respondents who reported selected behaviors and attitudes (%) and number of "don't know" and "refusal" responses given (N), by mode of audio-CASI administration

	Audio Only		Audio+Video	
	%	N	%	N
Drug use				
Injected drugs	3.7	1	1.7	0
Sexual behavior				
Never had sex	4.4	1	6.9	0
Masturbated 1+ times last yr.	85.5	3	77.1	5
Ever had abortion	40.9	1	39.3	0
STDs				
Dripping or oozing	17.0	1	12.3	1
Anal sex (females)	26.7	0	30.4	0
Attitudes				
Prefer white neighborhood	62.5	2	69.5	8
Report more frequent sex	41.7	10	45.1	14
Report more sex partners	19.5	17	27.3	26
Income >$50K	28.3	5	32.3	10

partners; it provided no neutral response category.) A final question on income registered a nonresponse rate of 8 percent ($N = 15$).

SUMMARY

Adoption of audio-CASI technology is accelerating in response to its obvious benefits--the privacy it offers and the standardization of the interview process it affords, even for respondents who are not literate (Hendershot et al., 1996; Turner et al., 1996). But issues related to its implementation have yet to be resolved. One particularly troubling gap in knowledge is the ways in which implementation may affect data quality.

The purpose of this experiment was to test whether individuals felt more comfortable responding to questions about their personal experiences in an audio-only mode or in an audio-plus-video mode, and whether the gender of the recorded voice had any effect on their responses. Our preliminary analyses indicate that voice gender does not appear to have much of an effect on responses. The effect of presentation mode is somewhat more complex. Apparently, respondents (particularly more educated respondents) found the audio-plus-video mode quicker and easier to use and to understand; they completed the survey in less time in that mode than in the audio-only format. Respondents also preferred the audio-plus-video mode for answering sensitive questions. However, the longer time it took respondents to complete the questionnaire in the audio-only mode may be offset by their sense that the audio mode provided more privacy than the audio-plus-video mode.

Despite respondents' perceived sensitivity of the survey questions, only a very few items resulted in nonresponse rates greater than 3 percent. For the majority of sensitive behaviors, mode differences were negligible. It thus appears that audio-CASI, with or without the video option, provides an equally attractive alternative to traditional survey administration techniques for measuring sensitive behaviors.

REFERENCES

Hendershot, T.P., S.M. Rogers, J.P. Thornberry, H.G. Miller, and C.F. Turner. 1996. Multilingual A-CASI: Using English-speaking interviewers to survey elderly members of Korean-speaking households. In Richard Warnecke (ed.), *Health Survey Research Methods: Conference Proceedings.* Hyattsville, Md.: National Center for Health Statistics, pp. 165-169.

Turner, C.F., S.M. Rogers, T.P. Hendershot, H.G. Miller, and J.P. Thornberry. 1996. Improving representation of linguistic minorities in health surveys: A preliminary test of multilingual audio-CASI. *Public Health Reports* 3(3):276-279.