TECHNICAL PAPERS ON
HEALTH AND BEHAVIOR MEASUREMENT

TECHNICAL PAPER 31

Automated Self-Interviewing and the Survey Measurement of Sensitive Behaviors

Charles F. Turner, Barbara H. Forsyth, James O'Reilly, Phillip C. Cooley, Timothy K. Smith, Susan M. Rogers, and Heather G. Miller

Reference Citation

Turner, C.F., B.H. Forsyth, J. O'Reilly, P.C. Cooley, T.K. Smith, S.M. Rogers, and H.G. Miller (1998). Automated self-interviewing and the survey measurement of sensitive behaviors. In *Computer-Assisted Survey Information Collection*. Edited by Mick P. Couper, Reginald P. Baler, Jelke Bethlehem, Cynthia Z, Clark, Jean Martin, William Nicholls II, and James M. O'Reilly, ISBN 0-471-17848-9, New York: Wiley and Sons, Inc.

CHAPTER 23

Automated Self-Interviewing and the Survey Measurement of Sensitive Behaviors

Charles F. Turner
Research Triangle Institute

Barbara H. Forsyth
Westat

James M. O'Reilly, Phillip C. Cooley, Timothy K. Smith, Susan M. Rogers, and Heather G. Miller
Research Triangle Institute

23.1 NOMENCLATURE: THREE FLAVORS OF CASI

The early 1990s witnessed the first ripples of what is likely to become a tidal wave of surveys using automated self-interviewing systems. Those systems currently fall into three main groups: video computer assisted self-interviewing (CASI), audio-CASI, and telephone audio-CASI (T-ACASI).

23.1.1 Video Computer Assisted Self-Interviewing

In video-CASI, respondents view the questions on a computer screen and enter their answers using the computer keyboard. This technique has precursors from the age of large mainframes and minicomputers (e.g., Evan and Miller, 1969), and it provided the earliest demonstration that automated self-interviewing might reduce underreporting of sensitive or stigmatized behaviors (Waterton and Duffy, 1984). Video-CASI, using laptops, has been explicitly tested in several methodological studies (e.g., O'Reilly et al., 1994; Tourangeau

Computer Assisted Survey Information Collection, Edited by Mick P. Couper, Reginald P. Baker, Jelke Bethlehem, Cynthia Z. F. Clark, Jean Martin, William L. Nicholls II, and James M. O'Reilly. ISBN 0-471-17848-9

and Smith, 1996); video-CASI also is available as an alternative interviewing mode in most audio-CASI systems.

23.1.2 Audio Computer Assisted Self-Interviewing

This technique adds audio features to those provided by video-CASI. Using laptop computers, respondents listen to questions through headphones and enter answers by pressing labeled keys. The recorded audio component has voice quality sound; it does not rely on synthesized voices and presents no significant delays in playing back audio-delivered questions. Early researchers and developers saw this technology as a way of providing complete privacy to *all* respondents (including those with poor reading skills) when reporting on sensitive topics. Although audio-CASI bears a superficial resemblance to early attempts to use a Sony Walkman to read survey questions (Camburn, Cynamon, and Harel, 1991), it is in fact fundamentally different. Because audio-CASI is computer controlled, it is capable of executing skip patterns, checking for out of range responses and inconsistencies across similar questions, and generating electronic data.

Today the use of this technology is experiencing explosive growth. A number of large-scale probability surveys using audio-CASI have been tested in the field (e.g., Mosher and Duffer, 1995; Udry, 1995; Turner et al., 1996a). In addition the federal government recently announced that its National Household Survey on Drug Abuse (NHSDA) would be converted to audio-CASI.

23.1.3 Telephone Audio Computer Assisted Self-Interviewing

In 1995 audio-CASI was adapted for use in telephone surveys, thus providing a promising new mode of survey administration. The technique was developed by adding telephone capabilities to a standard audio-CASI system. A successful T-ACASI pilot study has been completed, replicating the National AIDS Behavioral Survey (Catania et al., 1994) and comparing standard telephone survey administration with T-ACASI administration (Turner et al., 1996c). In addition implementation of T-ACASI in a large-scale methodological experiment (Turner, Miller, and Catania, 1995) began in November 1996 with the survey of a large sample of gay men (target $N > 6000$) residing in major cities in the United States. A parallel experiment is under way in the winter of 1997–1998 using a probability sample ($N = 3000$) drawn from the telephone accessible population ages 18 to 49 in the United States.

23.2 THEORETICAL ADVANTAGES OF CASI TECHNOLOGIES

Automated self-interviewing differs fundamentally from other technologies that have automated aspects of survey procedures, such as CAPI and CATI. In those applications, the respondent-interviewer interaction was not fundamen-

tally altered by the introduction of the new technology. In CAPI and CATI surveys the measurement process still relies on an interviewer asking questions and a respondent providing answers. Thus, not surprisingly, there is only weak and inconsistent evidence that these technologies affect the willingness of respondents to provide accurate responses to survey questions. Of course, the new technologies may have altered the error structure of *some* measurements.

All CASI methods remove or minimize the role of the human interviewer in the measurement process, a change that can have numerous effects. The key role of the interviewer in survey research has been well-recognized in the social sciences for more than half a century (e.g., Hyman et al., 1954). Indeed it has been argued that all survey data arise from a complex interpersonal exchange, that thus they embody the subjectivities of both interviewer and interviewee, and their interpretation requires a multitude of assumptions concerning, among other things, how respondents experience the reality of the interview situation, decode the "meaning" of survey questions, and respond to the social presence of the interviewer and the demand characteristics of the interview (Turner, 1981). By removing the requirement that respondents divulge sensitive, stigmatized, or counternormative behaviors to another human, CASI procedures may substantially reduce the extent to which response accuracy for such measurements is compromised by the social presence of the human interviewer.

Audio-CASI and T-ACASI technologies also offer other important methodological advantages over standard survey methods, such as interviewer-administered questionnaires and paper self-administered questionnaires (SAQs). Most notably, audio-CASI and T-ACASI:

- can be used with any respondent who can hear and speak—without the requirement of literacy in any language;
- may permit efficient multilingual administration of surveys without requiring multilingual survey interviewers;
- offer the traditional advantages of computer assisted survey technologies (i.e., computer controlled branching through complex questionnaires; automated consistency and range checking; automatic production of data files; etc.); and
- provide a completely standardized measurement system—every respondent (in a given language) hears the same question asked in exactly the same way.

23.3 IN-PERSON AUDIO-CASI

The earliest video-CASI systems were designed for mainframes and minicomputers (Kiesler and Sproul, 1986; Griest, Klein, and Erdman, 1976; Lucas et al., 1977; Waterton and Duffy, 1984; see also Saris, Chapter 21). With the advent of personal computers (PCS) in the early 1980s came proposals for the

development of PC-based video-CASI (e.g., Dubnoff, Kiesler, and Turner, 1986). At about the same time the growing epidemic of the human immunodeficiency virus (HIV) generated a widely recognized argument for better measurements of sexual and other sensitive behaviors (see Turner, Miller, and Moses, 1989; Miller, Turner, and Moses, 1990).

CASI measurements have emerged as a methodological response to the need for improved data on such behaviors. The addition of audio to the CASI interviewing mode was accomplished in 1991 by groups at the University of Michigan and Research Triangle Institute (RTI) (see Johnston and Walton, 1995; O'Reilly et al., 1994). Development of RTI's audio-CASI interviewing system was stimulated by the suggestion of David Celentano that a "voice-administered CAPI" interview might reduce measurement bias in the collection of data on AIDS-related behaviors from samples of intravenous drug users, many of whom possess limited literacy skills (Project Light, 1991; Turner, 1991).

Early tests established the feasibility and advantages of using audio-CASI to administer complex questionnaires in personal interview surveys. The RTI audio-CASI pilot test conducted in 1991–1992 indicated that the technology was stable and could be used with a minimum of disruption to typical survey and research routines (O'Reilly et al., 1994). Soon thereafter Turner, Lessler, and Gfroerer (1992a) recommended testing the use of audio-CASI in the NHSDA. The first small-scale methodological tests embedded within the context of that survey were not completed until the winter of 1996, and the results are not yet publicly available. Nevertheless, a variety of other evidence accumulated since 1992 demonstrates the soundness of the above recommendation. In the next section we present some examples derived from our research program. Further data are presented by Tourangeau and Smith (Chapter 22).

In January 1995, RTI's in-person audio-CASI technology was field tested in two major national surveys: the National Survey of Family Growth, or NSFG (Cycle V: $N = 10{,}000$ females, ages 15 to 44) and the National Survey of Adolescent Males, or NSAM (new cohort: $N = 1741$ males, ages 15 to 19). Preliminary data from the surveys indicate that field interviewers and survey respondents had few problems accepting the audio-CASI technology and that audio-CASI produced substantial increases in the reporting of sensitive behaviors.

23.3.1 National Survey of Family Growth

The 1995 NSFG collected repeated measurements using interviewer-administered questioning (IAQ) and audio-CASI. Women in the 1995 NSFG first reported on their history of abortions, number of sexual partners, and a wide variety of other topics (e.g., history of pregnancies and contraceptive practices) to a female interviewer, who used CAPI technology to record their answers. In a second phase of the interview, a small number of the same questions were repeated, but this time the respondent used the audio-CASI technology. Thus all of the women who completed the NSFG's face-to-face IAQ (10,847

respondents) had a second opportunity to provide responses in the audio-CASI mode, which we hypothesized would encourage more complete reporting of sensitive events such as abortion.

Preliminary analyses of the data from the 1995 wave (Miller et al., 1997) confirm that among sexually active women, more women reported a history of abortion in the more private audio-CASI mode than in the interviewer-administered mode (see Table 23.1). Of the women who did not report an abortion in the original IAQ mode, 4.5 percent reported one or more abortions in the audio-CASI mode. (All percentages are weighted to account for variation in the probability of selection and nonresponse.) The preponderance of discrepant reports of abortion across modes occurred in the direction of increased reporting with Audio-CASI. Thus, among women who reported one abortion in the IAQ mode, only 1.8 percent decreased the number of abortions they reported while 5.8 percent increased the number in audio-CASI. It should also be noted that the effect of audio-CASI was greater among black women than among whites: 7.3 percent of black females who reported no abortions in

Table 23.1. Estimates of Percent of U.S. Women Who Have Had One or More Abortions by Mode of Questioning and Race (Estimates for Sexually Active Women Ages 15 to 44 Interviewed in the 1995 NSFG)

Original Report in IAQ	Subsequent Report in Audio-CASI			Unweighted *N*
	No Abortions	1 Abortion	2+ Abortions	
		All women		
No abortions	95.4	3.5	1.0	7,827
1 abortion	1.8	92.5	5.8	1,265
2 abortions	0.4	2.2	97.4	582
		White women		
No abortions	95.8	3.3	0.9	5,675
1 abortion	1.1	93.9	5.0	850
2 abortions	0.6	2.5	96.9	343
		Black women		
No abortions	92.8	5.3	2.0	1,742
1 abortion	4.4	85.3	10.3	355
2 abortions	0.0	1.8	98.2	217

Source: Miller et al. (in press).

Note: Of 9674 sexually active respondents who were interviewed in the IAQ mode, 2121 reported no history of pregnancy, and consequently interviewers did not administer the questions concerning abortion. In this analysis those women are considered as reporting no abortions. All women, regardless of pregnancy history, received questions on abortion in the audio-CASI mode. From this analysis 230 cases were excluded due to missing data.

the IAQ mode reported one or more abortions when using audio-CASI; for white participants the comparable percentage was 4.2 percent. In addition 10.3 percent of black females reporting one abortion in the IAQ mode reported two or more abortions in the audio-CASI mode, compared with 5.0 percent of white women. Overall, the odds of reporting an abortion were approximately 1.3 times greater when information was collected using the audio-CASI technology rather than an IAQ.

We suspect that these NSFG results understate audio-CASI's actual effect. Because of the design of the NSFG experiment, all women completed the IAQ before the audio-CASI reinterview; thus the order of presentation was not randomized. If respondents were motivated to provide consistent responses—or to avoid admitting that they had deceived the human interviewer—we would expect the audio-CASI mode to have less effect on the results than might have been obtained in a simple randomized experiment. In that kind of design, respondents would experience no pressure to provide consistent responses.

The second major audio-CASI survey conducted in 1995, the National Survey of Adolescent Males, used such a simple randomized design—but with a quite different population.

23.3.2 National Survey of Adolescent Males

Since 1988, NSAM has tracked the sexual, contraceptive, and HIV risk-related behaviors of a national probability sample of U.S. men ages 15 to 19 in 1988. A follow-up of the original cohort was conducted in 1991. In 1995 a third round of interviews with the original cohort, as well as interviews with a new cohort of young men ages 15 to 19 in 1995, was carried out.

In the 1988 and 1991 rounds of NSAM data collection, behavioral and demographic data were collected in face-to-face interviews. Questions about sensitive behaviors, such as use of illicit drugs, male-male sexual contacts, and violence were presented in a self-administered questionnaire that followed the interview and that the respondent sealed in an envelope before returning it to the interviewer (Sonenstein, Pleck, and Ku, 1991). The data contain some perplexing puzzles, the most troubling of which involve measurements of male-male sexual contacts, the most common mode of HIV transmission during the 1980s. The estimates are considerably lower than one might predict, based on the retrospective reports of national samples of adult men (see Turner, Danella, and Rogers, 1995).

Concerns about potential underreporting bias in the 1988 and 1991 rounds of NSAM and the desire to increase the actual and perceived privacy of the interview prompted the addition of a methodological experiment to the 1995 round (Turner et al., 1996a). Respondents were randomly assigned to receive the most sensitive NSAM questions in either the format employed in 1988 (personal interview followed by self-administered questionnaire) or the audio-CASI mode. Preliminary results are available from the first 928 respondents, or approximately one-half of the total sample. These are reproduced in Table 23.2.

Table 23.2. Percentages Reporting Different Types of Male–Male Sexual Contact, and Odds Ratios, in a National Sample of Males Ages 15 to 19, by Mode of Data Collection (Self-Administered Questionnaire or Audio-CASI)

	Estimated Percentage			
Measurement	Paper SAQ	Audio-CASI	Odds Ratio	*p*
Ever masturbated another male	1.1	2.3	2.07	0.29
Ever been masturbated by another male	0.6	3.0	5.44	0.03
Ever had insertive oral sex with another male	0.6	2.5	4.42	0.07
Ever had receptive oral sex with another male	0.6	2.1	3.67	0.13
Ever had receptive anal sex with another male	0.0	1.2	—	0.05
Ever had insertive anal sex with another male	0.6	1.6	2.93	0.23
Any male–male sex	1.1	4.7	4.26	0.01
(*N*)	(176)	(728)		

Source: Turner et al. (1996a).

Notes: Preliminary (unweighted) results from the 1995 NSAM. The *p*-values are for likelihood ratio chi-square for fit of independence model to the two-way table of mode by reporting of behavior.

Table 23.2 shows the percentage of respondents who reported engaging in each of six types of male-male sexual contacts. The final line of the table shows the results for a composite measure that indicates whether the respondent reported at least one type of contact. As the table shows, there were substantial and statistically reliable differences in the reports given in the audio-CASI interview and the paper and pencil SAQ. Indeed respondents were more than four times more likely to report some male-male contact in the audio-CASI interview. Those preliminary results suggest that audio-CASI reduced the underreporting of male-male sex that occurred in the prior NSAM surveys.

23.4 TELEPHONE AUDIO-CASI

23.4.1 Motivation for Implementing Audio-CASI in Telephone Surveys

Because of the substantial cost of sending field interviewers to tens of thousands of households across the country, some of the largest surveys of HIV-related and other sensitive behaviors have used telephone survey techniques (e.g., Catania et al., 1992; ACSF Investigators, 1992). Large samples are

particularly important in such surveys because they permit analyses of subpopulations that are crucial to understanding the dynamics of the HIV epidemic (e.g., the population with multiple new sexual partners in the past year). These telephone surveys have interviewed substantially larger samples than typical in-person surveys of HIV-related behaviors (e.g., Tanfer, 1993; Laumann et al., 1994).

Nevertheless, the cost advantage that argues for use of telephone survey methodology to collect sensitive data ignores the effect and importance of other factors. Researchers must also take into account that traditional telephone surveys of sensitive behaviors are vulnerable to serious reporting biases (Aquilino, 1994; Gfroerer and Hughes, 1992; Tourangeau and Smith, Chapter 22; Turner, Miller, and Rogers, in press). In an attempt to provide a technology that would reduce such biases, we extended our audio-CASI system for use in telephone interviewing.

23.4.2 Early Development Efforts

Our goal in modifying the audio-CASI technology was to enable the conduct of complex call-in or call-out telephone ACASI surveys. In a call-in survey the respondent initiates the interview by calling a number that is answered by the T-ACASI system. In a call-out survey a human telephone interviewer calls the respondent and subsequently transfers the call to the T-ACASI system. We developed a software platform that fully integrated audio-CASI and T-ACASI capabilities and that could be implemented on a wide array of hardware. In our T-ACASI system, PCs equipped with a hardware interface handle incoming and outgoing calls; as a result the administration of a survey can be painlessly transferred from one environment to another by cloning the relevant software and digitized voice files.

We use the term *T-ACASI* for these applications both for consistency with the nomenclature for other CASI modalities and to distinguish them from the more circumscribed data collections that have been carried out with earlier touchtone data entry (TDE) systems. Early experimentation with TDE was begun at the Bureau of Labor Statistics (BLS) during the late 1980s (Werking, Tupek, and Clayton, 1988; Phipps and Tupek, 1990). These early TDE applications have been limited to quite simple data collection tasks, typically involving only 5 to 10 questions asked without complex skip patterns or other tailoring of the survey instrument (see Weeks, 1992).

23.4.3 Major Effects

Our T-ACASI system became fully operational in February 1995. Testing began with a pilot study using the questionnaire from the National AIDS Behavioral Survey (Catania et al., 1992). The study used a cross-over experimental design to test respondents reactions to the new technology (see Turner et al., 1996 for a full description). Half of the sensitive NABS questions, which

dealt with heterosexual and same-sex encounters, HIV serostatus, and drug use, were asked by using standard telephone interview methods; the other half were administered with the T-ACASI technology. The order of presentation was balanced across the experiment. At the end of the interview, a human telephone interviewer queried respondents about their experience with each mode of interviewing.

The pilot test not only evaluated the feasibility of the new T-ACASI technology, but it also tested two hypotheses: (1) whether respondents would feel more comfortable reporting sensitive sexual behaviors to a computer compared with a human interviewer and (2) whether they were more likely in the T-ACASI mode to report engaging in stigmatized or sensitive behaviors and less likely to report normative behaviors than in the standard telephone interview mode. The quota sample for the experiment comprised people 18 to 49 years of age recruited either from households with listed telephones in Cook County, IL, or from the Wake County Sexually Transmitted Disease (STD) Clinic in Raleigh, NC. Quotas were set to produce approximately equal numbers of males and females in each stratum. Preliminary data from the first 142 cases in the study are encouraging and suggest that T-ACASI surveys are not only feasible but that they may improve the quality of data on sensitive topics (Turner et al., 1996c).

Touchtone phone ownership was quite common among respondents, a finding that supports the feasibility of T-ACASI. Data on touchtone ownership from the entire sample indicated that among the 306 screened households that reported having an adult ages 18 to 49 and that also reported on the types of telephones in their households, 302 (99%) said they had a touchtone phone. Touchtone telephone ownership was considerably less common among elderly households (i.e., those without an 18- to 49-year-old). Only 153 of 203 households (75%) in that group reported having a touchtone phone.

Secondary analysis of data from a national telephone survey conducted in 1993 (Frankovic, Ramnath, and Arnedt, 1994) indicate a similar pattern for the telephone accessible portion of the U.S. population. Approximately 90 percent of respondents ages 18 to 49 reported that they were speaking from a touchtone phone. For 85 percent of respondents in this age group, the interviewer verified that the phone was *currently switched* to generate the tones required for T-ACASI. In 5 percent of cases, the phone appeared to be switched to generate pulses; in those instances the telephone mode selector would have to be reset for a T-ACASI interview. Touchtone telephones, however, are less widely available among people age 50 and older, and in both age groups they are less widely available among people with lower levels of income and education.

Also supporting the finding of T-ACASI's feasibility was the stability of the system and the relative ease with which interviewers used it. In general, respondents were enthusiastic about the system as well. Preliminary results from questions to respondents about their experiences with T-ACASI and standard telephone interviewing indicated that respondents thought T-ACASI

was better at protecting privacy, provided a more comfortable environment for answering sensitive questions, and was more likely to elicit honest reporting of sexual and drug use behaviors.

Even with the small sample sizes available for analysis (maximum *n*'s for each condition were 79 and 67), the preliminary data from this study indicated significant differences, or differences bordering on significant, in the responses given to many of the most sensitive questions asked in the survey (see Table 23.3). For example, among respondents who reported engaging in anal sex during their lifetime, preliminary data indicated a 17 percentage point difference between the two modes of administration (25.4% for interviewer administered questioning vs. 42.0% for T-ACASI). That result appears to be attenuated in the full study; however, Table 23.3 shows other results indicating that, compared with standard telephone interviewing, T-ACASI substantially increased the likelihood that respondents would report sensitive behaviors or experiences.

Although such preliminary results from a small sample must be approached with caution, they suggest that subjects prefer T-ACASI for answering sensitive questions. They also suggest that T-ACASI increases the likelihood that subjects will report sensitive behaviors and decreases the likelihood that they will overreport normative ones. Of course much work remains to be done to secure a broader range of evidence on the effects of T-ACASI on respondents' reporting. Part of that effort is likely to take account of the cognitive demands of such new technologies present.

23.5 COGNITIVE ASPECTS OF AUTOMATED SELF-INTERVIEWING

With the rapid development of enhanced CASI technologies, including audio-CASI and T-ACASI, early methodological work has generally focused on preliminary questions of feasibility and general data quality (O'Reilly et al., 1994; Turner et al., 1996b, 1996c; Hendershot et al., 1996; Tourangeau and Smith, 1996, Chapter 22). In these early stages less attention was given to the new cognitive demands that accompany the technologies. In addition CASI technologies are frequently used for surveys about highly sensitive or stigmatized behaviors in research that focuses on the cognitive demands of the question answering tasks rather than the cognitive demands that are unique to the data collection technologies.

Because of the limited research and knowledge base and the tentativeness of the conclusions that can be drawn, we present below—in a very abbreviated fashion—some of the emerging findings from the limited cognitive research that has been done on CASI interviewing.

Offering both audio and video questioning modes improves respondent acceptance. O'Reilly et al. (1994) found that respondents rated audio-plus-video CASI as easier to use and more interesting than the video-only mode. Rogers

Table 23.3. Estimates of the Prevalence of Sensitive Behaviors Obtained from Telephone Interviews Using Human Interviewers and Telephone Audio-CASI (Preliminary Results)

Measurement	Estimated Percentage: Human Interviewer	Estimated Percentage: T-ACASI	Odds Ratio	*p*
	Anal intercourse			
Ever had anal intercourse	25.4	42.0	2.13	0.03*
Had anal intercourse in past 6 months	3.0	12.0	4.43	0.03*
	Oral sex			
Given oral sex (since age 18)	79.7	79.5	0.99	ns
Received oral sex (since age 18)	89.8	89.0	0.92	ns
	Limited sexual experience			
Had no sex partners since age 18	1.6	7.6	4.93	0.09
Had no sex in last 5 years	4.8	11.4	2.53	0.15
Did not have sex in past 6 months	1.5	8.0	5.74	0.01*
Had sex fewer than 10 times in past 6 months	22.7	41.3	2.51	0.01*
	Condom use			
Never used a condom in lifetime	8.1	18.4	2.57	0.07
Used condom every time had sex in past 6 months	14.8	6.8	0.42	0.14*
Used condom almost every time had sex in past 6 months	27.8	15.9	0.49	0.14*
	Stability and quality of relationships			
Most recent sexual relationship lasted less than 6 months	5.8	21.3	4.42	0.01
Never discussed sex life with most recent partner	1.9	14.8	8.83	0.03*
Discussed sex life less than once a month	28.8	49.2	2.39	0.03*
Ever had a one-night stand since age 18	59.0	64.4	1.26	ns
(Approximate *N*)	(67)	(79)		

Source: Turner et al. (1996c).

Note: *p*-value marked with asterisk do not apply to test of individual odds ratios but are derived from testing of multi-category response distribution for a mode effect of trend (see source for details).

et al. (1996) found similar results for audio-plus video CASI compared with audio only. These results suggest that offering multiple CASI modes enhances respondents' favorable reactions to CASI methods. In cognitive testing with audio-plus-video CASI administration for NSAM and the NHSDA, respondents had the option of turning off either the audio or the video presentation. In both studies all cognitive interview participants completed interviews with both the audio and the video presentations turned on. During debriefing interviews several cognitive interview participants indicated that they thought the audio presentation was unnecessary for them. However, none of them exercised the audio-off option, which suggests that the redundant mode was not sufficiently annoying to cause respondents to turn it off.

Even in sex surveys, the gender of the voice is unimportant. Most of our methodological studies have involved administering potentially sensitive items about sexual behavior. Thus we hypothesized that the gender of the audio-CASI voice might affect how respondents reacted to the audio-CASI interview and how they answered audio-CASI items. In cognitive testing for NSAM, we used three voices to present the audio-CASI interview: one male voice and two female voices. Follow-up debriefing items asked respondents about their reaction to the audio-CASI voice they had heard, and most respondents were unable to report whether their interview was administered by a male or a female. Rogers et al. (1996) included voice gender as an experimental factor in their methodological study and found no effects of voice on either ratings of the audio-CASI methods or on survey response distributions, even when controlling for effects of the respondent's gender and race.

Respondents will back up to change answers during CASI interviews, but backing up is likely to occur less frequently than in interviewer questioning. Rogers et al. (1996) calculated how often respondents repeated questions or backed up to re-hear earlier questions. Overall, a small proportion (3.2%) of survey questions were repeated; that is, very few respondents listened to the question or a portion of the question again before entering a response. Not unexpectedly, respondents repeated questions more often in the audio-only than in the audio-plus-video mode. In either mode, respondents backed up over less than 1 percent of the questions one or more times. In T-ACASI, respondents rated interviewer questioning as overwhelmingly superior to T-ACASI for changing answers (Turner et al., 1996c).

Elderly people, individuals who are unfamiliar with computers, the less educated, and some other segments of the population may hesitate to participate in a CASI interview. However, most people in those segments appear able to complete interviews and to provide reasonably reliable data. Couper and Rowe (1996) have reported that interviewers read questions or entered answers for a substantial proportion of respondents (21%) in the video-CASI segment of a 1992 Detroit area survey. Even among respondents with "prior computer experience," interviewers provided assistance in 16 percent of cases. Interviewers in the study had discretion in providing such assistance, although it was intended that "these options were to be used as a last resort in the case of refusal or incapacity" (p. 92).

In contrast to that experience, the NSFG and NSAM did not permit such direct interviewer involvement, and there were only a handful of cases in which an audio-CASI interview could not (or would not) be completed by a respondent. Similarly Hendershot et al. (1996) reported relatively little difficulty in conducting multilingual audio-CASI interviewing with a sample of elderly, monolingual Koreans (mean age of 71.5 years)—even though the field interviewers who conducted the survey spoke no Korean themselves. The addition of the audio mode might explain the observed divergence from Couper and Rowe's results; however, Saris (personal communication, 1996) reports similar findings for his video-CASI surveys, which include interviewing of the elderly. These results lead us to suspect that allowing interviewers to use their discretion in providing assistance may transform signs of respondent "hesitancy" into interviewer perceptions that "help" is necessary.

The foregoing conclusions should probably be treated as interim speculations. Relatively little is known with certainty about the cognitive aspects of automated self-interviewing, and much work remains to be done.

We are currently conducting a broadly focused program of research on the applications and effect of audio-CASI and T-ACASI on survey and other measurements of sensitive behaviors. Among the activities under way are a number of experimental comparisons of these techniques in both general samples and clinic studies. Over the next several years we hope that this research program and those of other investigators will provide a firmer empirical foundation than now exists for drawing inference about the cognitive aspects of CASI and its effect on the quality of survey and other measurements of sensitive behaviors.

23.6 HARDWARE AND SOFTWARE CONSIDERATIONS

The audio-CASI systems currently used in survey research have been developed by a few large survey organizations to meet the needs of their own research programs and those of major clients. These organizations have drawn on internal programming staff with specialized skills to develop either an audio-CASI subsystem for their existing computer assisted interviewing (CAI) package or a custom audio-CASI interviewing system. In the following sections we describe how audio-CASI systems have been designed and implemented. We also discuss some of the technical factors that make DOS implementations of audio-CASI inferior to Windows-based implementations, and we consider how future Windows-based CAI packages will offer audio and other multimedia capabilities.

23.6.1 In-Person Audio-CASI

In 1996 virtually all audio-CASI interviewing in survey research was based on DOS platforms despite the difficulty of implementing audio and other multi-

media features in DOS. In the DOS environment the interfaces between the audio component and the central processing unit are achieved by using audio-specific device drivers. However, the design of the drivers that operate on DOS platforms is based on Intel 16-bit architecture assumptions, whereas the CAI systems currently available are based on Intel 32-bit architecture. As a result compatibility between the audio and CAI parts of the audio-CASI system can be directly achieved only by developing a CAI routine that does not utilize the full capabilities of the 32-bit architecture used by new personal computers.

In order to use audio routines from CAI systems, researchers have developed two distinct methods: terminate and stay resident processes (TSRs), which are invoked by a "hot key" or a software equivalent, and external DOS-executable routines, which are invoked by system calls from the CAI software. In both methods, either an external custom audio device connected to the PC's parallel port or an internal chip added to the PC's motherboard provides the audio capability.

This rather cumbersome and somewhat fragile architecture in DOS audio-CASI systems is unlikely to change. Audio services are most widely used on PCs in conjunction with multimedia applications. In the DOS environment audio and other multimedia are implemented only in a limited, specialized manner, that is, for computer games. However, the success of Windows in the early 1990s created what has become a growing consumer market for multimedia. Today most vendors of multimedia devices and software tools have shifted their business focus almost entirely to Windows, which means that audio device drivers are now based on Intel 32-bit architecture assumptions.

Although the mass market has moved strongly to Windows, survey research organizations continue to use DOS for the great bulk of their interviewing systems, in large part because the CAI systems they employ were, and continue to be, DOS-based. A second reason, we suspect, is that prior to the development of audio-CASI, there was no serious claim that a Windows CAI application would add significant value to the automated interviewing process.

23.6.2 Telephone Audio-CASI

In T-ACASI the sound card used for audio-CASI is replaced by a telephony card. That change in hardware and associated software makes it possible to ask questions using audio sent over the telephone line and to collect a respondent's answers to those prerecorded questions by decoding the tones generated when the respondent presses the keys on a touchtone telephone. The telephony card provides audio features for T-ACASI applications in a manner that is completely analogous to the functioning of the sound card in audio-CASI applications. Similarly, in T-ACASI, the telephone keypad replaces the PC's keyboard. There are, however, obvious differences between T-ACASI and audio-CASI applications that arise from the distinct technological features of the two media. For example, T-ACASI keypad responses are limited to 10

digits (0 to 9) and two special characters (# and *). In addition, T-ACASI systems have no special screen handling needs.

Automated touchtone response systems, known as interactive voice response (IVR) systems, have proliferated in the past two years and are used in a large number of commercial applications. Many vendors market such systems. A major difference between IVR and T-ACASI systems is that the latter are designed to support complex survey operations, and they have at their core a CAI routine that manages the survey component of the application. In contrast, IVR systems use relatively simple data collection instruments to capture limited amounts of data from a large number of respondents. Consequently the design of T-ACASI systems focuses on maintaining the survey processing features of the CAI component of the package, whereas the design of IVR systems focuses on handling many respondent calls simultaneously (i.e., on supporting channel multiprocessing features).

For T-ACASI systems implemented under DOS, which is a single-user operating system, the multitasking features have to be built directly into the IVR software. Consequently event-based programming procedures must be used to develop channel multitasking capabilities. That constraint is a fundamental problem for T-ACASI designs. T-ACASI systems have a more complex CAI core than IVR systems; the core frequently supports a separate language (with its own grammar and syntax) to describe edit and skip logic. Such systems are more difficult to convert from a procedural style of programming, which is typical of DOS-based applications, to an event-based style typical of Windows-based applications. Hence, on the one hand, the migration of T-ACASI applications to Windows is not a trivial task. On the other hand, moving the T-ACASI application to Windows enables T-ACASI to make use of a channel multitasking environment. That means that different T-ACASI processes can use different channels on the same telephony card simultaneously—in other words, the same machine can handle multiple interviews. (Such simultaneity cannot be achieved by running multiple DOS-based T-ACASI applications on a Windows 95 machine.)

Recently we moved both IVR and T-ACASI applications to a Windows NT platform, and the performance of the two applications improved significantly compared with the DOS versions. In addition the multitasking capabilities of the NT operating system, though not totally free of bugs, have been successfully demonstrated in both IVR and T-ACASI applications. Improvements in the NT drivers are expected to resolve the few remaining problems.

23.6.3 Emerging Availability of Audio-CASI in Standard Systems

Currently three major CAI system vendors (CASES, Blaise, and Surveycraft) are developing Windows versions of their DOS-based CAI systems. Many market research software vendors have already made the move to Windows. Because multimedia features are available in the languages and systems that reside on Windows platforms, one can expect that audio-CASI capabilities will

using another external audio device. In the Adolescent Health Study, Grilley, Kean, and Nichols (1996) cited the following:

- Because of malfunctions, an estimated 10 percent of the field interviewers probably did not use the audio equipment.
- Twelve percent of documented support calls to the central office "involved problems with the programming and hardware associated with the Audio-CASI portion of the interview."
- The audio device had to be carefully screwed into the parallel port before raising the laptop's cover, or the device would not function properly in the audio-CASI section of the interview, which occurred after a length CAPI section.
- Interviewers had to be careful not to attach the audio device too tightly or they would be unable to remove it at the end of the interview.

Those experiences suggest that a number of factors determine whether an external audio device is robust and reliable, including the design of the device itself, the laptop, the software, organizational experience, and so forth. As organizations gain experience, serious difficulties are reduced to a low level. Overall, despite problems and shortcomings, external audio devices have established a strong, successful track record in demanding field situations. Therefore, although internal sound chips promise the highest level of reliability and ease of use, at least for the interim, external sound devices can be a capable alternative.

23.6.5 Prospects for General Audio-CASI Availability

The key factor in determining general audio-CASI availability will be the arrival of general-purpose CAI software for the Microsoft Windows and Windows 95 graphical environment. Our experience in transporting a CAI system from DOS to Windows 95 demonstrated the difficulties of moving a sophisticated software package from a linear procedural design to an event-based design (Cooley and Turner, in press). Nevertheless, the resulting benefits proved well worth the effort. They included true 32-bit performance without any of the side effects associated with accessing external audio processes, a single program that works with all Windows-compatible sound systems, and flexible screen handling characteristics such as font and bitmap manipulation. All of those capabilities are now available without any apparent loss of performance.

We can look to such Windows-based CAI systems for other features as well. They should be able to interface easily with virtually all types of audio devices, including built-in audio chips, PCMCIA cards, and external parallel port devices. Almost all PC laptops sold since 1995 are capable of running such software. The key remaining obstacle is the release of Windows versions of CAI

soon become standard. In the meantime some vendors are developing audio-CASI capabilities for existing DOS-based systems.

Our experience in developing audio-CASI systems in DOS and Windows suggests that the difficulties in implementing a general-purpose CAI system with audio features in the Windows environment should not be underestimated. A key design consideration is that the system's target user must be assumed to have little or no computer experience and only a few minutes of training before being let loose to complete a complex interview. That demanding usability requirement means the system must provide a high level of speed, simplicity, and robustness along with the desired graphical and audio features (Cooley et al., 1996; Cooley and Turner, 1996). The system must also be able to work well with people who may never have used a computer or even a keyboard. Thus the system must permit respondents to engage the instrument with full audio accompaniment without confusion, delays, or intrusions that disrupt their concentration. Even small extra demands on the respondent may undercut support, participation, and data quality.

23.6.4 External Sound Device versus Built-in Audio Capability

Cooley et al. (1996) present a thorough review of technical issues and document their experience in implementing audio-CASI for PC laptops. They conclude that for (non-Windows) DOS-based audio-CASI applications using laptop PCs, the only viable options are the use of one of a limited but increasing number of laptop PCs with built-in Sound Blaster-compatible audio, or the use of an external audio device connected to the laptop's parallel printer port. PCMCIA cards are not feasible for DOS platforms because the audio interface is typically designed for Windows.

Both built-in audio and external devices provide equivalent functionality including acceptable voice quality audio and performance that is adequate for most low-end PCs. The main differences between these options are the comparative level of their reliability and their effects on the field interviewing process. External devices add complexity, increase weight, and take time to set up. Thus an interviewer in a respondent's home must not only set up the laptop but also attach the audio device to it, although extensive and demanding field experience of the 1995 NSFG and NSAM indicates that external audio devices are a minor inconvenience that field staff are usually able to handle relatively easily. Indeed, in the NSFG pretest and main study, RTI field staff reported infrequent malfunctioning of the audio devices and few difficulties in managing the equipment (Kinsey et al., 1995).

In NSAM the same equipment worked reliably after a problematic start-up phase with older model laptops. Interviewers did, however, complain that the external audio equipment was somewhat cumbersome to use. After the early spate of equipment problems, those complaints dwindled. Nevertheless, survey managers reported that they preferred laptops with built-in audio capability.

The National Opinion Research Center reported some significant problems

systems, such as CASES, Blaise, and Surveycraft, that offer the other features required for large-scale complex surveys. Given the actual or imminent availability of suitable software and hardware, plus the growing body of evidence that audio-CASI is a powerful, if not essential, survey research method for addressing sensitive topics, much wider adoption of audio-CASI can be anticipated in coming years.

23.7 FUTURE DEVELOPMENTS

The growing use of CASI—and of audio-CASI in particular—has sometimes led us to speculate that these technologies could well revolutionize the way all types of social measurements will be made in the future. Key elements of that speculation will, we believe, stand the test of time. Nevertheless, the more general claim is likely to be found wanting. The ultimate result for those concerned with social measurement may still be dramatic, but the drama will play itself out in scenes and with nuances that are only vaguely anticipated or appreciated today.

The great rush of enthusiasm that fueled the adoption of in-person audio-CASI interviewing in research (our own and others) has arisen in part from the fact that the technology elegantly solved a problem of great import for *some* areas of the social and statistical sciences. In particular, the idea of audio-CASI interviewing emerged to meet a pressing need in the field of behavioral research on the spread of HIV—that is, a measurement method that would reduce bias in the reporting of highly stigmatized behaviors (such as male-male sexual contacts) that risked HIV transmission. The need was particularly apparent among subpopulations, such as injection drug users, for whom literacy levels were inadequate to permit confident use of paper and pencil questionnaires. Indeed that specific problem was the initial stimulus for our current program of research and development of audio-CASI interviewing systems.

In considering the application of audio-CASI technology to social measurement, we believe nonetheless that the range of areas in which one might expect large improvements in response validity is not sufficient to sustain a true "revolution" in measurement methods, nor is the target population extensive. Literacy may be a problem for a notable fragment of the U.S. population, but that fraction is considerably smaller than a majority of potential respondents. Since audio-CASI is required only for that segment of the population that is incapable of responding to written questions (or of following written instructions), the theoretical advantages of in-person audio-CASI are delimited both in terms of the target population and in terms of the content areas in which major improvements in measurement validity might reasonably occur. Furthermore, for typical in-person surveys, the economic benefits of audio-CASI, if any, will not be large because interviewers typically must wait in the respondent's home while the respondent completes the audio-CASI interview.

So, one might ask, where is the "revolution" coming from? We offer two

replies. First, CASI and audio-CASI are fundamentally altering our understanding of the prevalence and patterns of certain hidden and highly sensitive characteristics of the U.S. population. Even preliminary data indicate that those alterations are not always occurring in ways that would have been expected. Thus the first revolutionary aspect of the adoption of these technologies is the substantial rethinking of the realities we and others have described in our past attempts to assess the prevalence and patterns of sexual, drug using, and other sensitive behaviors (e.g., see Catania et al., 1994; Laumann et al., 1994; Turner, 1989; Turner, Miller, and Moses, 1989; Miller, Turner, and Moses, 1990; Fay et al., 1989; Rogers and Turner, 1991; Turner, Danella, and Rogers, 1995). In that regard we would offer the additional observation that what is called the "mode effect" in survey measurements (i.e., the difference in responses obtained from in-person versus CASI questioning) is likely to become a useful instrument for judging the social sensitivity of survey questions and topics and the variation in that sensitivity over time.

The second revolutionary aspect of the use of CASI technology arises from a technological afterthought that will, we believe, induce fundamental change in the way social measurements of sensitive aspects of human behavior are carried out in the future. That afterthought was the extension of audio-CASI technology to the telephone, an almost serendipitous move that has fundamentally redefined what is possible when conducting telephone surveys. Before this merging of technologies, there was no way to offer respondents in a telephone survey a completely private mode of data collection. Although in-person surveys have long used paper and pencil questionnaires—and can now use CASI—telephone surveys have required respondents to report on even the most private aspects of their sexual and other behaviors to a human interviewer. Now the advent of T-ACASI has transformed the potential of telephone surveys by providing a fully private mode of data collection within the more economical telephone survey process. This development should allow researchers to feel more comfortable in using telephone methodologies for future surveys of sensitive behavior.

ACKNOWLEDGMENTS

Preparation of this chapter was supported by the U.S. National Institutes of Health through grants to the first author from the National Institute of Child Health and Human Development and the National Institute on Aging (grant R01-HD/AG31067-04) and from the National Institute of Mental Health (grant R01-MH56318-01). The opinions expressed in this chapter are the sole responsibility of the authors.