TECHNICAL PAPERS ON
HEALTH AND BEHAVIOR MEASUREMENT

TECHNICAL PAPER 25

Implementing Audio-CASI on Windows' Platforms

Philip C. Cooley and Charles F. Turner

Reference Citation
Cooley, P.C., and C.F. Turner (1998). Implementing Audio-CASI on Windows' Platforms. *Computers in Human Behavior* 14(2):195-207, 1998.

Pergamon

Computers in Human Behavior, Vol. 14, No. 2, pp. 195–207, 1998

Printed in Great Britain
0747-5632/98 $19.00 + 0.00

PII: S0747-5632(98)00001-6

Implementing Audio-CASI on Windows Platforms

Philip C. Cooley and Charles F. Turner

Research Triangle Institute,
Research Triangle Park

Abstract — *Audio computer-assisted self-interviewing (Audio-CASI) technologies have recently been shown to provide important and sometimes dramatic improvements in the quality of survey measurements. This is particularly true for measurements requiring respondents to divulge highly sensitive information such as their sexual, drug use, or other sensitive behaviors. However, DOS-based Audio-CASI systems that were designed and adopted in the early 1990s have important limitations. Most salient is the poor control they provide for manipulating the video presentation of survey questions. This article reports our experiences adapting Audio-CASI to Microsoft Windows 3.1 and Windows 95 platforms. Overall, our Windows-based system provided the desired control over video presentation and afforded other advantages including compatibility with a much wider array of audio devices than our DOS-based Audio-CASI technologies. These advantages came at the cost of increased system requirements—including the need for both more RAM and larger hard disks. While these costs will be an issue for organizations converting large inventories of PCs to Windows Audio-CASI today, this will not be a serious constraint for organizations and individuals with small inventories of machines to upgrade or those purchasing new machines today.* © *1998 Elsevier Science Ltd. All rights reserved*

Keywords — Audio-CASI, Windows, surveys

Requests for reprints should be addressed to Philip C. Cooley, Research Triangle Institute, P.O. Box 12194, Research Triangle Park, NC 27709-2194, USA. E-mail: pcc@rti.org

A shortcoming of most DOS-based audio computer-assisted self-interviewing (Audio-CASI) technologies is their limited ability to manipulate video representations of the questions presented to respondents. Important screen environment features that are largely unexploited in DOS-based Audio-CASI technologies include selection of the size and typefaces of screen fonts, the choice of colors for text and background, and the display of digitized images. Technologies that permit manipulation of size, typeface, and color would allow a survey designer to adjust the visibility of screen display and to highlight important sections of text. The ability to embed digitized images would serve numerous purposes including the inclusion in the Audio-CASI instrument of "show cards" that are usually handed to respondent during an ordinary (non-Audio-CASI) interview. (Such show cards might be used, for example, to assist a respondent in identifying the type of oral contraceptives she was using by displaying the distinctive packages used to dispense the pills.)

The major obstacle to implementing these features in past Audio-CASI systems has been the constraints imposed by the DOS operating system. To provide these video features and to exploit the Windows' graphic user interface, we attempted to migrate one of our DOS-based Audio-CASI technologies to Windows in both 16-bit and 32-bit implementations. Our Windows Audio-CASI application is a true 32-bit application unlike the DOS version which uses a 32-bit CASI manager RTI Forms System (Cooley, 1994) calling a 16-bit playfile routine that plays previously recorded audio wave files. The invocation of the playfile executable is achieved via a DOS systems call; the 16-bit version of our Audio-CASI application operated flawlessly on both Windows 95 and Windows 3.1 platforms. However, the 32-bit implementation produced exceptions in Windows 3.1 but operated without problem in Windows 95. This paper describes our approach for implementing a survey of sexual and HIV-related behaviors (Rogers, Forsyth, Miller, Smith, & Turner, 1996) as a Windows application. The benefits and weaknesses of the resulting Windows-based Audio-CASI system are described.

Background

During the past 5 years, researchers at the Research Triangle Institute (RTI) have developed a suite of Audio-CASI technologies to administer complex questionnaires in personal interview surveys (Cooley et al., 1996; O'Reilly, Hubbard, Lessler, Biemer, & Turner, 1994; O'Reilly & Turner, 1992; Turner, Lessler, & Gfroerer, 1992; Turner, Miller, Smith, Cooley & Rogers, 1996). Using portable laptop computers, these technologies are designed to afford maximum privacy to respondents without requiring that they be literate in any language. Both large-scale field studies and smaller-scale pilot studies have

found that it is possible not only to interview "typical" respondents, but also to interview those who pose special challenges. Turner, Rogers, Hendershot, Miller, and Thornberry (1996) and Hendershot, Rogers, Thornberry, Miller, and Turner (1996), for example, have demonstrated that it is possible for monolingual English-speaking field interviewers to use Audio-CASI to successfully interview subjects who speak only Korean or Spanish.

All of RTI's Audio-CASI technologies use digitized voice files rather than synthesized speech. Our in-person Audio-CASI interviewing systems appear to be as reliable as standard computer- assisted personal interviewing systems under field conditions. Field staff have generally reported few problems with the operation of this technology in a broad range of environments using a range of laptop PCs and sound devices.

RTI's in-person Audio-CASI technology has been used in several major surveys including the 1995 round of the National Survey of Family Growth (NSFG; $N = 10{,}000$) conducted for the National Center for Health Statistics and the National Survey of Adolescent Males (NSAM new cohort; $N = 1741$) funded by the National Institutes of Health ($N = 1741$). The 1995 NSAM included a randomized experiment comparing the reporting of sensitive behaviors using Audio-CASI with the reporting obtained using paper-and-pencil self-administered questionnaires (SAQs). Despite reasons to believe that both modes should be perceived as affording equivalent privacy (Turner, Danella, & Rogers, 1995), preliminary results suggest that Audio-CASI substantially increased respondents' willingness to report sensitive behaviors. Turner, Ku, Sonenstein, and Pleck (1996) found that the reporting of male–male sex increased four-fold under Audio-CASI (from 1.1 to 4.7% for first phase of experiment). Similarly, preliminary findings from a pilot test conducted in preparation for the NSFG indicate that Audio-CASI significantly increased reporting of abortions (Research Triangle Institute, 1994). Other pilot studies (Hendershot et al., 1996; O'Reilly et al., 1994; Rogers et al., 1996; Turner, Rogers, et al., 1996) indicate that:

1. Audio-CASI technology is stable and can be used with a minimum of disruption to typical survey and research routines.
2. Virtually without exception, respondents had no trouble using this new technology. This was true for both educated persons and those with substantial literacy problems, for the young and the old, for English-speaking respondents and for those who spoke only Spanish and Korean.
3. Even literate respondents reported preferring the new technology to paper-and-pencil SAQs.

The Audio Component of Audio-CASI

Two DOS-based systems that support Audio-CASI applications have been developed by RTI researchers. The first system uses Blaise—a computer-assisted interviewing package developed by Statistics Netherlands (1989a,b). The audio hardware component of this system uses an external digital audio processor—the Antex Audioport developed by ANTEX Electronics (O'Reilly et al., 1994). The second system (Cooley et al., 1996) uses the RTIFS in conjunction with Soundblaster-compatible audio devices. Elsewhere, we have described Soundblaster-compatible devices that support Audio-CASI applications on DOS platforms (Cooley et al., 1996). Two of these devices, a Texas Instrument TM4000M laptop computer with an built-in sound chip,[1] and a Toshiba 386 laptop with a Portable Sound Plus device connected via the PC parallel port, were used in field trials of our second Audio-CASI technology conducted in Baltimore County, MD, during 1995.

In the following section, we describe our experiences converting the DOS-based technology used in this Audio-CASI field test to operate in a Windows environment.

METHODS

Software

RTIFS (Cooley, 1994), a forms management software package originally developed in DOS, was rewritten for Windows platforms (both 16-bit and 32-bit applications were tested). The original system was developed in the computer language C (Turbo C 2.0 and Watcom C, The Watcom Group, 1994). Applications that use this system are described in Hoffman et al. (1995), Rogers et al. (1996), and Turner, Miller, et al. (1996). This system was redesigned with Windows-type controls using the Borland C++ 4.52 language (Borland International, 1994a) in conjunction with Borland's Object Window Library (OWL) 2.5 system (Borland International, 1994b).

The original DOS version of RTIFS consisted of a Form Builder/Editor and a forms-based Data Entry component. This builder/editor converted an electronic description of the form (developed in a word processor) into three data structures that define video layout, edit logic, and skip logic actions. The data entry component uses the three data structures to compose the screen, edit the keyed information, and determine the order of items presented to the respondent. During the keying process a fourth data structure, an output stream of formatted keystrokes, is created.

Survey Instrument

The survey instrument used in this test consisted of 95 questions on a wide range of sensitive topics including respondents' drug and alcohol use, sexual practices and orientation, number and characteristics of sexual partners, contraceptive behaviors and condom use, communication with partners about sex and contraceptive use, knowledge of and past history of sexually transmitted diseases, prior experiences with violence including history of arrest, psychological problems, and attitudes about gender and race, as well as the sociodemographic characteristics of the respondent. This survey instrument was successfully administered to 205 residents of Baltimore County, MD, in a 1995 methodological pilot experiment. We are currently conducting a full-scale methodological experiment using this survey instrument to compare the results obtained using Audio-CASI with those obtained using interviewer administration and paper-and-pencil SAQs.

Conversion to Windows

To implement this survey instrument in Windows, we modified our DOS user interface to fit the Windows environment. Four generic screen types were created as basic components for our redesigned user interface. These components correspond to the following question or activity types:

1. select one characteristic from a list of possible candidates (Select One Screen);
2. select all of the characteristics that apply from a list of possible characteristics (Select All Screen);
3. enter a single numeric variable (Numeric Entry Screen);
4. review information no entry (Review Screen).

We mapped the screens defined by the DOS version of RTIFS into these four generic screen types. Thus, the DOS and the Windows versions of this survey use the same data structures. In order to map the DOS screens into these four generic screen types in Windows, we made use of standard Windows-based controls. (By "standard" we mean that these controls are available in all Visual-type systems such as Visual Basic, Visual Fox-Pro, and all C++ systems.) The controls include:

1. *Static controls* which provide a Windows' dialog box with static text. Static text is text that the user can see but cannot change. Our Windows implementation of Audio-CASI uses this control to display such static

information to the user for the communication of instructions, or the reading of the question, etc.

2. *Push-button controls* that perform tasks each time they are pressed. The standard tasks implemented in our Windows Audio-CASI system are to (a) backup one question, (b) suspend the interview, (c) advance one question, (d) replay the audio, (e) respond "don't know" to the question, and (f) refuse to answer the question.
3. *Edit controls* that enable the user to type in and edit the text they enter into an input dialog box. Our Windows Audio-CASI system uses this type of control to enter numeric data or open ended responses.
4. *Check-box controls* that toggle a check mark. The control appears as a small rectangular button and a title that is displayed to the right of the rectangle. When the rectangle is clicked the check mark displayed in the rectangle is toggled, but the on–off state of the button is independent of any other check boxes on the screen. In our Windows Audio-CASI system, these controls are used to implement the "Check All That Apply" generic screen.
5. *Radio button controls* are check box controls with two states: "checked" or "unchecked". Note, however, that when one radio button is checked all other radio buttons on the page are turned off. Our Windows Audio-CASI system uses these controls to implement the "Select One" generic screen.

Examples of three of these generic screens are shown in Figures 1–3. The Select All Screen is nearly identical to the Select One Screen and is not shown. The Select One Screen is implemented as a static control with a varying number of radio button controls (e.g., a true–false list would have two). The Select All Screen is implemented as a static control with a varying number of check box controls. The Numeric Entry Screen is implemented as a static control with a single edit control and the Review Screen has only a static control.

A separate edit control is also included (if needed) for each data item required beyond the items comprising the generic form types. The parser that builds the generic screens automatically adds the necessary number of edit controls to the generic screens. As an example, note the radio button labeled "other" in Figure 1. This figure is an example of a Select One generic screen with an additional edit control added. If the "other" radio button is selected then the "specify text" field is highlighted indicating that additional information that defines the "other" category should be embedded. The extra edit control provides this vehicle. If the other category is not selected the edit control is automatically skipped.

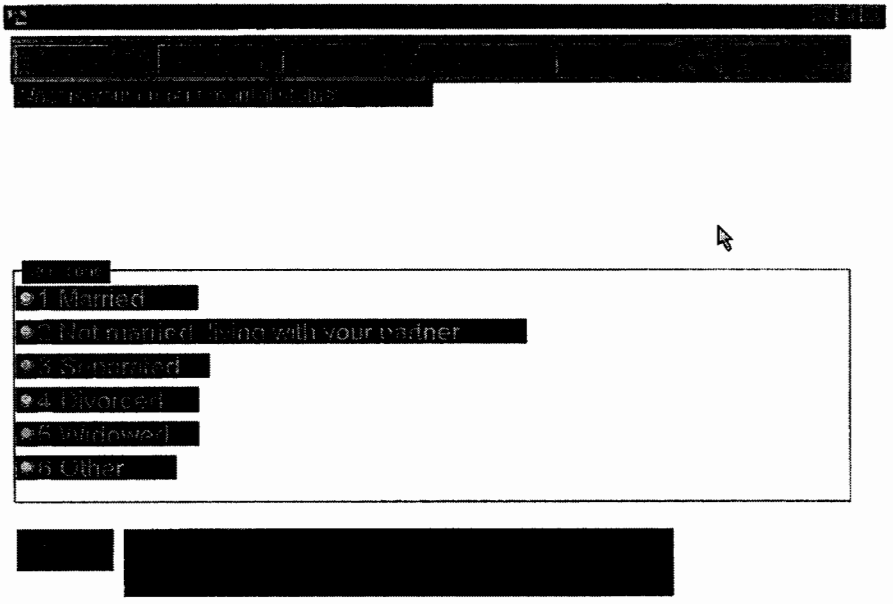


Figure 1. Example of Select One Screen.

Because we maintained the same basic set of screen conceptual elements in both the DOS and Windows versions of our Audio-CASI system, we were able to use the DOS versions of RTIFS' EDIT and SKIP logic processors in the Windows' implementation. For example, each screen has a definable address (for skip logic branching purposes), audio play back files are recognized by all generic screen types, and each edit and skip item is linked to a specific data entry variable. The logic processors parse the skip and edit logic

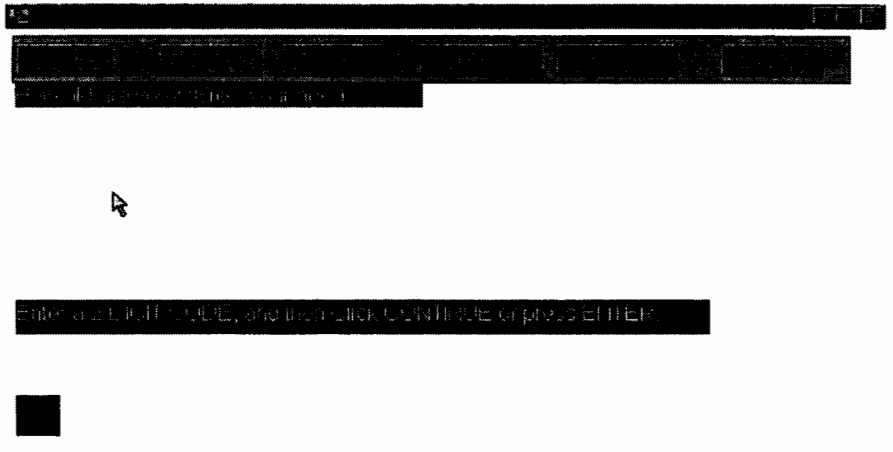


Figure 2. Example of Numeric Entry Screen.

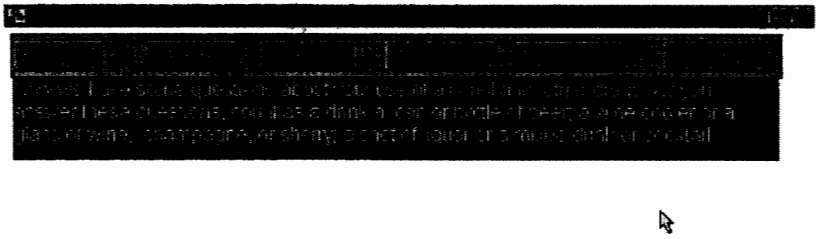


Figure 3. Example of Review Screen.

in the DOS and Windows applications to determine if the condition for branching to a new screen (skip logic) or if the conditions for accepting the entered data as valid (edit logic) apply. Also—as in the DOS version of the software—the Windows version creates a formatted data structure of ASCII keystrokes. This constitutes the keyed data record.

Finally, both versions were designed to accept data without using the mouse and without adhering to standard Windows keyboard protocol. The function keys are used to backup to the previous question, replay a question, and suspend an interview. However, the Windows application will accept (if desired), mouse input (i.e., radio button and check box items can be clicked instead of numerically entering data). For example, the "next screen," the "previous screen" (backup), and the "sound replay" functions can be invoked by clicking the appropriate button located on the upper part of the screen.

Testing

To test the Windows version of the data entry software, the existing questionnaire developed in the DOS version of RTIFS was implemented using our new Windows-based Audio-CASI system. This test provided an assessment of the versatility of the generic screen types described already and permitted an assessment of the utility, from the user's perspective, of the Windows version. In addition, all of the audio hardware devices examined by Cooley et al. (1996) on the DOS platform were retested for compatibility on the Windows' platform.

RESULTS AND DISCUSSION

Table 1 presents the results of our mapping of the survey instrument from the DOS based Audio-CASI system to our Windows-based system using the generic screens already defined. As Table 1 shows, most of the DOS-based screens "fit" into one of the four generic screens without modification. Those

Table 1. Frequency of Mapping Survey Questions to Generic Screen Types in Windows-Based Audio-CASI Systems

Generic screen types	Edit control	
	Without	With
Select One Screen	119	0
Select All Screen	0	0
Numeric Entry Screen	9	1
Review Screen	16	–
Total	144	1

Note. Questions are from survey instrument described in Rogers et al. (1996). The number of screens (145) exceeds the number of survey questions because alternative formats were used for questions directed at particular subpopulations (e.g., male and female versions of questions on reproductive history).

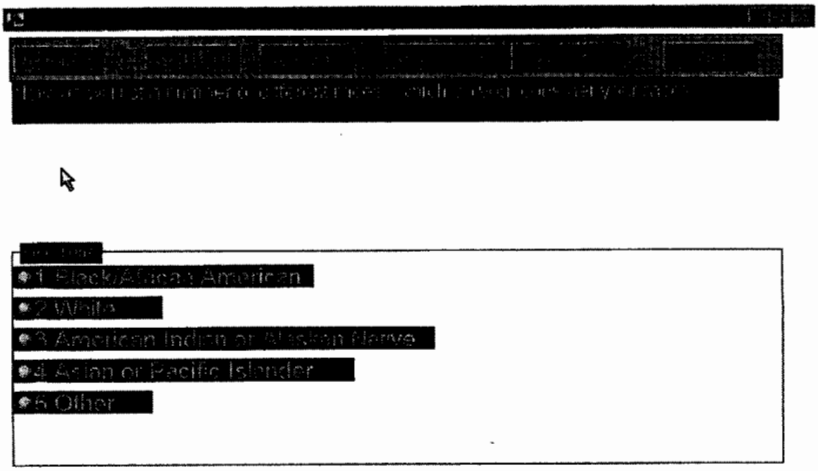
that did not fit required only the addition of a single edit control to augment the generic screen type.

To provide readers with an appreciation (albeit in black and white) of the video features available in the Windows implementation of our Audio-CASI system, Figure 4 compares one of the screens from the DOS-based implementation of this survey instrument (bottom of Figure 4) with the Windows-based version (top of Figure 4). It should be noted that the Windows screen controls vary the font size and can be written over a bitmap image with a different bitmap assigned to each screen.

In addition to providing a more flexible user interface, the Windows version provides significantly improved audio quality. Of the devices listed in Table 2, only 3 of the 13 devices provided Soundblaster compatibility in a DOS-based Audio-CASI application (Cooley et al., 1996). However, under Windows, all of the devices supported our Audio-CASI application, and thus they could be considered candidates for Audio-CASI sound devices on a Windows platform.

Our testing demonstrates that migrating DOS-based Audio-CASI technologies to Windows is both feasible and practical. There are not substantial impediments to this migration, and our testing suggests that most survey instruments can be supported with a small number of generic screen types. The benefits of implementing Audio-CASI technology on a Windows platform are significant and include:

1. improved (easier-to-read) fonts;
2. improved color handling (for highlighting and emphasis);



```
Now we will list a number of different races.  Which do you consider
your race?
     1   Black/African American
     2   White
     3   American Indian or Alaskan Native
     4   Asian or Pacific Islander
     5   Other
     _
```

Figure 4. Comparison of screens for same question in Windows-based (top) and DOS-based (bottom) systems.

3. ability to provide descriptive images (bitmaps) electronically during the data entry process;
4. access to a wider selection of audio devices that support the playback of WAV files;
5. true 32-bit applications.

On the negative side, the Windows platform requires more extensive computer resources. So, for example, a minimum of 4 MB of RAM are required for Windows 3.1 and 8 MB for Windows 95. Consequently, implementing Audio-CASI applications in Windows may require costly new hardware purchases for large-scale surveys. Somewhat surprisingly, the Windows version of our Audio-CASI technology did not require faster processors to achieve the same level of performance as our DOS-based implementation. We suspect that greater efficiency is achieved in our

Table 2. Sound Devices Tested for Use in DOS- and Windows-Based Audio-CASI Systems

Device	Manufacturer	Approximate price[a] ($)	Type	Weight
Antex AudioPort	Antex Electronics	389	Parallel	1 lb 7 oz
Port-Able Sound Plus	DSP Solutions, Inc.	189	Parallel	1 lb 13 oz
Sound Exchange Model B	Interactive	229	Parallel	2 lb 6 oz
Audioman	Logitech, Inc.	179	Parallel	13 oz
Audioport	Video Associates Lab., Inc.	295	Parallel	13 oz
The Cat	Vocal Tec.	179	Parallel	10 oz
IBM PCMCIA 16 Bit Audio	IMB Corporation	265	PCMCIA	4 oz
Tempe Audio Card	IO Magic Corporation	299	PCMCIA	4 oz
WAVJammer	New Media Corporation	399	PCMCIA	3 oz
Thinkpad 750[b]	IBM Corporation	2439*	Built-In	NA
Thinkpad 760C	IBM Corporation	5369*	Built-In	NA
TM4000	Texas Instrument	1650*	Built-In	NA
TM5000	Texas Instrument	2775*	Built-In	NA

[a] Prices shown are suggested list prices, except for those prices with an asterisk.
[b] Model has been discontinued and replaced by new model described as offering similar features.
* Approximate mail order prices in 1996.

Windows implementation as a result of the better integration of multimedia features into the Windows operating system. Both the DOS-based and Windows-based versions of our Audio-CASI system ran well on laptops equipped with Intel 486SX/25 processors.

Future Development Work

Our development work to date has provided us with a working Windows-based Audio-CASI system that provides equivalent functions and improved features to its DOS-based Audio-CASI parent. This new system has not yet been subjected to large-scale field testing, and there are some aspects of the user interface that merit greater attention. For example, our interface presently makes the assumption that survey respondents will be literate enough to use task-bar buttons labeled "replay," etc. Clearly substituting icons might have some value for subjects with reading impediments. Similarly, the visual layout of the screen is presently somewhat cluttered and could be improved.

We are also anxious to exploit the multilingual font capabilities afforded by our transition to the Windows platform. We have previously demonstrated that monolingual English-speaking survey interviewers can use Audio-CASI to obtain interviews from non-English speaking respondents (Hendershot et al., 1996; Turner, Rogers, et al., 1996). This demonstration was conducted using an Audio-CASI survey instrument that had been recorded in Korean and Spanish. Native language video presentation was not attempted because

of the difficulties of presenting multiple non-Roman fonts on video displays in the DOS environment. Since Windows loosens this constraint, we are anxious to develop and test a second-generation multilingual Audio-CASI application that will present questions in the respondents' native languages through both the audio and video channels.

Acknowledgments — The Audio-CASI system described in this paper was developed by Philip Cooley with support from the Research Triangle Institute. Methodological testing and preparation of this article were supported, in part, by the National Institute of Child Health and Human Development and the National Institute of Aging under NIH Grant R01-HD/AG31067 to Charles Turner.

NOTE

[1]The convenience of the laptop computer with the built-in sound device proved to be an attractive feature of the Audio-CASI technology implemented for this field trial. Interviewers expressed a strong preference for the laptop with the built-in sound device because of the simplicity of the set-up which eliminated the necessity for connecting auxiliary devices to the portable computer and removed a potential source of hardware unreliability. As a result of the interviewers' preferences, more than 70% of the 205 interviews in this field were obtained using this hardware type.

REFERENCES

Borland International. (1994a). Borland C++ (Version 4.52) [Computer programming language]. Scotts Valley, CA: Borland International, Inc.

Borland International. (1994b). Borland Object Windows Library (OWL) (Version 2.5.) [Computer software]. Scotts Valley, CA: Borland International, Inc.

Cooley P. C. (1994). RTIFS user's manual [Software manual]. Research Triangle Park, NC: Research Triangle Institute.

Cooley, P. C., Turner, C. F., O'Reilly, J. M., Allen, D. R., Hamill, D. N., & Paddock, R. E. (1996). Audio-CASI hardware and software considerations in adding sound to a computer-assisted interviewing system. *Social Science Computer Review, 14*(2), 197–204.

Hendershot T. P., Rogers, S. M., Thornberry, J. T., Miller, H. G., & Turner, C. F. (1996). Multilingual Audio-CASI: Using English-speaking interviewers to survey elderly Korean-speaking households. In R. Warnecke (Ed.), *Health survey research methods* (pp. 165–170). Hyattsville, MD: National Center for Health Statistics.

Hoffman, J. A., Schneider, S. J., Koman, J. J., Flynn, P. A., Luckey, J. W., Cooley, P. C., Wish, E. D., & Diesenhaus, H. I. (1995). The centralized intake model for drug abuse treatment: The role of computerized data management. *Computers in Human Behavior, 2*, 215–222.

O'Reilly, J. M., Hubbard, M. L., Lessler, J. T., Biemer, P. P., & Turner, C. F. (1994). Audio and video computer assisted self-interviewing: Preliminary tests of new technologies for data collection. *Journal of Official Statistics, 10*, 197–214.

O'Reilly, J.M. & Turner, C. F. (1992, March). *Survey interviewing using audio-format, computer-assisted technologies*. Paper presented at the Washington Statistical Society, Washington, DC.

Research Triangle Institute. (1994). *NSFG Cycle V Pretest Report*. Research Triangle Park, NC: Author.

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

Statistics Netherlands. (1989a). Blaise [Computer software]. Voorburg, The Netherlands: Netherlands Central Bureau of Statistics.

Statistics Netherlands. (1989b). Blaise 2.0 CAPI/CATI user manual [Software manual]. Voorburg, The Netherlands: Netherlands Central Bureau of Statistics.

Turner, C. F., Danella, R. D., & Rogers, S. M. (1995). Sexual behavior in the United States, 1930–1990: Trends and methodological problems. *Sexually Transmitted Diseases, 22*, 173–190.

Turner, C. F., Ku, L., Sonenstein, F. L., & Pleck, J. H. (1996). Impact of Audio-CASI on bias in reporting of male–male sexual contacts. In R. Warnecke (Ed.), *Health survey research methods* (pp. 171–176). Hyattsville, MD: National Center for Health Statistics.

Turner, C. F., Lessler, J. T., & Gfroerer, J. C. (1992). Improving measurements of drug use: Future directions for research and practice. In C. F. Turner, J. T. Lessler, & J. Gfroerer (Eds.), *Survey measurement of drug use: Methodological studies* (pp. 299–306). Washington, DC: Government Printing Office.

Turner, C. F., Miller, H. G., Smith, T. K., Cooley, P. C., & Rogers, S. M. (1996). Telephone audio computer-assisted self-interviewing (T-ACASI) and survey measurements of sensitive behaviors: Preliminary results. In R. Banks, J. Fairgrieve, L. Gerrard, et al. (Eds.), *Survey and Statistical Computing 1996* (pp. 121–130). Chesham, UK: Association for Survey Computing.

Turner, C. F., Rogers, S. M., Hendershot, T. P., Miller, H. G., & Thornberry, J. P. (1996). Improving representation of linguistic minorities in health surveys: A preliminary test of multilingual Audio-CASI. *Public Health Reports, 111*(3), 276–279.

The Watcom Group. (1994). WATCOM C/386 (Version 10.0) [Computer programming language]. Waterloo, ON, Canada: The Watcom Group, Inc.