

SELF-ADMINISTERED QUESTIONS BY TELEPHONE

EVALUATING INTERACTIVE VOICE RESPONSE

ROGER TOURANGEAU
DARBY MILLER STEIGER
DAVID WILSON
The Gallup Organization

Over the past 25 years, computerization has swept over survey research, making computer-assisted data collection the de facto standard in the United States and Western Europe (Couper and Nicholls 1998). The move to computerization may now be ushering in a golden age for self-administered questions; the newest methods of survey data collection to emerge have reduced the role of the interviewer or eliminated it entirely, allowing the respondents to interact directly with the computer. The new modes of self-administered data collection include Web surveys and a technology variously referred to as interactive voice response (IVR), touchtone data entry (TDE), and telephone audio computer-assisted self-interviewing (T-ACASI). These different labels refer to the same data collection technology in which the computer plays a recording of the questions to the respondents over the telephone, who indicate their answers by pressing keys on the handsets (Appel, Tortora, and Sigman 1992; Blyth 1997; Frankovic 1994; Gribble et al. 2000; Harrell and Clayton 1991; Phipps and Tupek 1990; Turner et al. 1996b, 1998). We will refer to this method of data collection as IVR, the term used at Gallup and at most market research firms.

Automated telephone systems for gathering nonsurvey information are now widespread (e.g., for catalog sales, airline reservations, banking, and so on)

The work reported here would not have been possible without the aid of a number of our colleagues. We are particularly grateful to Wendy Moody, Rajesh Srinivasan, and Steve Hanway for sharing data and documentation with us. We also thank Kelly Green, who oversaw Study 4 and made invaluable contributions to that study. Many of the concepts presented here emerged in conversations with Mick Couper, who also commented on an earlier draft of this article. We happily acknowledge our debt to him. A longer version of this article with an extended discussion of nonresponse is available through the SMP Working Paper Series. Roger Tourangeau is now Senior Research Scientist at the University of Michigan's Survey Research Center and Director of the Joint Program in Survey Methodology at the University of Maryland.

and similar systems are being used to collect survey data as well. Within the federal statistical agencies, IVR and voice recognition entry (VRE)—in which the respondents speak their answers aloud—have until recently been limited to ongoing establishment surveys, but outside the government IVR has won wider acceptance, especially for brief interviews. Just as ACASI has begun to supplant earlier methods for gathering information in face-to-face settings, IVR may gradually displace other forms of data collection by telephone.

Proponents of IVR (e.g., Gribble et al. 2000; Turner et al. 1996a, 1998) cite a wide range of potential gains—particularly, reduced social desirability bias and decreased cost—but doubts remain about whether the technology is suitable for long or complex interviews. In part, the cost savings from IVR depend on whether the respondents (recruited through some other means) dial directly into the IVR system or are initially contacted by telephone interviewers, who switch them into the IVR system. Because it involves a live interviewer, the recruit-and-switch version of IVR does not yield such dramatic cost savings as the inbound version. Both forms of IVR can be subject to high levels of nonresponse, including high rates of breakoffs. In addition, with recruit-and-switch, a new form of nonresponse can occur in which potential respondents hang up during the switch to IVR.

The Gallup Organization uses IVR to collect data for a broad range of clients. In 1999 alone, Gallup completed more than 1 million IVR interviews, generally brief ones assessing customer satisfaction. In addition, it has carried out several experiments comparing data collected by IVR with data from other modes of data collection. This note describes the results of the four experiments done at Gallup in 1998 and 1999. Two of them compare IVR with CATI, one compares IVR with mail data collection, and the final study compares IVR with both CATI and mail. Three of the four studies focus on the impact of IVR on reporting; the fourth focuses on the effects of IVR on nonresponse.

Data Sets

All of the Gallup experiments involved the recruit-and-switch variant, in which live interviewers contact respondents initially and then switch them to IVR. The switch was immediate in the first three studies; in the fourth, sample members who agreed to take part were asked to schedule their interview for a later date. Each of the studies used a single female voice to record the questions for the IVR interviews. Although we have not made an extensive comparison of the different software packages for IVR, we believe the Gallup system is typical of those used throughout the industry.

The first three experiments examine customer satisfaction, comparing IVR data with data collected in CATI interviews (Studies 1 and 2) and mail questionnaires (Study 3). Although the samples assigned to the different methods

of data collection in these studies were comparable initially, those who completed the questionnaires under the different modes may differ systematically because of the effects of nonresponse. The potential confound between reporting differences and differential nonresponse is particularly serious for the studies we report because the response rates were low (see table 1). The analyses attempt to control for differences in the composition of the different mode groups, but we have limited background information on the sample members and cannot rule out the possibility that such differences contribute to the differences in reporting that we observe.

The final experiment we report examined what happened with a much longer questionnaire that used the items in the Census 2000 Long Form; this final experiment included all three modes—IVR, CATI, and mail. The response rates reported here follow the AAPOR Standard Definitions (using the RR3 formula) and take into account nonresponse to the screening/recruitment portion of the survey as well as nonresponse to the main interview. In Study 4, we examine breakoff rates in more detail, analyzing the proportion of cases who completed a questionnaire once they had agreed to take part.

Although the response rates in all four experiments are low by academic standards, they are typical of what is found in market research settings, where the premium is on the timeliness of the results. In these studies, there was limited refusal conversion during the recruitment phase and no attempt was made to recontact cases that broke off during the IVR portion of the interview. Gallup telephone interviewers work on both IVR and CATI studies; thus, the interviewers assigned to recruit the IVR cases in the four studies reported here and those assigned to interview the CATI cases were drawn from a single pool of interviewers and received the same training.

STUDY 1: CUSTOMERS AT A BANK

The first study compared IVR and CATI interviews with customers of a bank serving customers in 15 states and the District of Columbia. Two samples were randomly selected to represent current account holders. Telephone interviewers contacted members of both samples and administered screening questions via CATI to identify bank customers who had been to a branch in the prior month. Once they completed the screening questions, members of one sample were transferred to an IVR interview (“The rest of the survey is completed using our automated system, where you press your answers using the numbers on your phone”), and members of the other sample continued answering questions posed by live interviewers. A total of 502 respondents completed the IVR interview; another 503 completed the CATI version. Another 230 members of the sample were switched to the IVR interview but broke off before completing it. The data were collected during an 8-day period in August of 1998. The overall response rate for the IVR sample was 24

Table 1. Key Features of Four Mode Experiments

Study/Population	Mode/Response Rate	Breakoffs	Questionnaire Length	Key Items
Study 1: recent bank customers	IVR (<i>n</i> = 502/24%) CATI (<i>n</i> = 503/ 49%)	230 IVR breakoffs	26 questions (on average); 50 possible	6 questions on teller performance (10-point scales)
Study 2: customers of fast food chain	IVR (<i>n</i> = 410/45%) CATI (<i>n</i> = 405/ 34%)	5 IVR breakoffs	7 questions	4 items on satisfaction with last visit (5-point scales)
Study 3: recent bank customers	IVR (<i>n</i> = 24,095/ 33%) Mail (<i>n</i> = 638/17%)	170 IVR breakoffs	17 questions	11 questions on employee contact (5-point scales)
Study 4: Maryland residents	IVR (<i>n</i> = 75/29%) CATI (<i>n</i> = 81/38%) Mail (<i>n</i> = 72/36%)	53 IVR breakoffs	Varied depending on size and composition of household (see text)	7 demographic items and 8 items on income (in varied formats)

NOTE.—Response rates count IVR breakoffs as nonrespondents. (There were no breakoffs in the CATI conditions.) In Studies 1–3, cases were switched immediately to the IVR questionnaire; in Study 4, both IVR and CATI cases were encouraged to complete the main questionnaire at a later date.

percent; the response rate for the CATI group was 49 percent, roughly twice as high.

The main interview included questions about the teller service at the branch, about other staff there, about the customer's experiences in opening an account, and about the resolution of any problems encountered. Our analysis focuses on the questions concerning the teller service, which applied to virtually all the respondents. All six of these items used a 10-point scale, with respondents pressing the one and zero keys to indicate they were "extremely satisfied" and pressing the one key to indicate they were "not at all satisfied." Across both mode groups, respondents were administered an average of 26 questions and were skipped out of an average of 24 items (which did not apply to them); the six teller items and five demographic questions were administered to all the respondents. In the IVR version of the questionnaire, respondents could replay a question by pressing the star key or skip a question by pressing the pound key.

STUDY 2: CUSTOMERS AT A FAST FOOD CHAIN

The second study compared IVR and CATI interviews with customers of a fast food chain. Two list-assisted, random-digit-dial (RDD) samples were selected to represent households in the continental United States. The CATI interviewers administered screening questions to identify eligible members of both samples (persons over 12 years of age who had visited one of the chain's restaurants during the 4 weeks prior to the interview). When more than one household member was eligible for the survey, the interviewer asked to speak with the one with the most recent birthday. After the interviewer identified a respondent, members of one sample continued with the CATI interview, completing six additional items about their most recent visit to the chain's restaurants and a question on their age; members of the other sample were switched to an IVR version of the same seven questions. A total of 815 respondents took part in the experiment, 410 completing the IVR interview and 405 the CATI interview. In addition, five members of the IVR sample broke off without completing the questionnaire.

STUDY 3: CUSTOMERS OF A SECOND BANK

The respondents in the third study were also bank customers (at a different bank from the one in Study 1). In August of 1998 and again in April of 1999, samples of customers were contacted by telephone and switched to an IVR interview. These customers had recently visited a branch of the bank and were selected from customer lists provided by the bank. At the same time, Gallup selected smaller comparison samples of bank customers from the same customer lists and mailed them a paper version of the questionnaire. Our analysis combines the data from both waves of data collection. In total, 24,095

customers completed IVR interviews and another 170 broke off without completing the questionnaire. A total of 638 returned mail questionnaires with usable data. The response rates for the IVR survey were approximately 32 percent in both August and April, reflecting losses during both the initial recruitment by a live interviewer and the IVR interview. The mail surveys had response rates of 18 percent in August and 13 percent in April.

Our analysis focuses on 11 items about the respondent's contact with the bank employee during their visit to the bank. These items consisted of a statement about the bank employee (" . . . communicated clearly with you"), with respondents indicating their agreement with the statement on a 5-point scale. Respondents in both groups were administered six additional items.

STUDY 4: MARYLAND RESIDENTS

The final study compared IVR data collection with both CATI and mail. Gallup interviewers contacted an RDD sample in Maryland and asked residents to take part in the study by completing a demographic questionnaire. Once they agreed to take part, respondents were randomly assigned to receive one of the three versions. A major purpose of the study was to determine IVR's suitability for long interviews, and we used the same questions that appeared on the Long Form questionnaire for Census 2000. The Long Form includes from 16 to 67 questions about each member of the household (the exact number depending on each person's age and employment status) plus an additional 17 to 33 questions about the housing unit (the exact number depending mainly on whether the household rents or owns its home). A typical family of four with two working parents and two children who live in a single-family house that it owns would receive close to 200 questions. A single person who had not worked in the past 6 years and who rents an apartment would still receive a minimum of 54 items. (A copy of the Long Form questionnaire is available at the Bureau of the Census web site—www.census.gov/dmd/www/pdf/d-61b.pdf.) Gallup telephone interviewers contacted a total of 1,882 households during September and October of 1998. Of these, 658 agreed to take part in the study; 200 were mailed a paper questionnaire, 257 were assigned to an IVR interview, and 211 were assigned to a CATI interview. Ultimately, 36 percent of the mail households returned questionnaires, 29 percent of the IVR households completed the questionnaire (although 49 percent completed one or more questions), and 38 percent of the CATI households completed a CATI interview.

The Long Form questionnaire is designed primarily as a mail instrument, and it includes some items (e.g., questions on annual utility costs) that are easier to answer if the respondent checks relevant records (e.g., utility bills). In addition, some of the items offer numerous response options (e.g., an item about educational attainment lists 16 possible answers). To help respondents cope with such items, we attempted to mail paper questionnaires to everyone

who agreed to take part in the study. During the initial recruitment, interviewers asked respondents for an address where we could send the paper questionnaire. However, some of the IVR and CATI cases insisted on completing the questions on the spot, and these were switched immediately to the CATI or IVR interview. Forty-eight of the 137 cases who started the IVR interview were switched immediately; similarly, 20 of the 81 cases who began the CATI interview started the questionnaire immediately.¹ Table 1 summarizes the key features of the four studies.

Results

Our analysis of the results focuses on two questions: How does IVR affect reporting? How does it affect willingness to complete an interview, particularly a long one?

REPORTING DIFFERENCES

Aside from its impact on data collection costs, the major potential advantage of IVR is its reduction of socially desirable responding (e.g., Turner et al. 1996a, 1998). Although the items in customer satisfaction surveys are not especially sensitive, the responses may still exhibit a form of reporting bias—the tendency to give positive or at least lenient ratings (e.g., Landy and Farr 1980; Sears 1983; see also Tesser and Rosen 1975). Because respondents are reluctant to give negative ratings, satisfaction scores tend to pile up on the positive end of the scale. Since IVR does not involve a live interviewer, it may increase respondents' willingness to express complaints about products or services.

Overall satisfaction ratings. Studies 1 and 2 compare customer satisfaction ratings under IVR and CATI. The studies involve different types of business (banking vs. fast food), different satisfaction questions, and different response formats (10-point satisfaction ratings vs. 5-point agreement scales). In addition, in one study, the IVR group had a lower response rate than the CATI group, but, in the other, the IVR group had the higher response rate. Despite all these differences between the two studies, the reporting results are quite consistent: On all six satisfaction items in Study 1 and on all four in Study

1. We also modified some of the Long Form questions to make them easier to administer by telephone. Our analysis focuses on the items administered in identical formats in all three versions of the questionnaire.

2, the IVR group reported significantly lower satisfaction than did the CATI group. A key difference appears to be the presence of a live interviewer.²

The hypothesis that the presence of a live interviewer is a key variable is consistent with the results for Study 3, which compared two modes that eliminate interviewer contact—IVR and mail. In that study, the differences between modes are fewer in number, and the IVR respondents reported higher satisfaction on average than did the mail respondents on 10 of 11 questions. Only five of the individual items show significant mode differences in univariate tests (although the difference considering all 11 items in Study 3 was significant in a multivariate test, $F = 3.44$, $df = 11$, 20876, $p < .001$). This comparison demonstrates that, while different modes of self-administration may not yield identical results, the outcomes are closer than when self- and interviewer-administered methods are compared. Figure 1 plots the means of the satisfaction items from the three studies.³

Even when a live interviewer does not administer the questions, satisfaction ratings still pile up at the positive end of the scale. This bunching was particularly clear in Study 3, where the majority of responses on all 11 items were “top box” responses, with most respondents describing themselves as extremely satisfied.

Random responding? It is possible that some of the IVR respondents in Studies 1 and 2 were simply picking answers at random to get through the interview. Random responding would tend to reduce the pileup at the positive end of the scale and lower the overall mean level of satisfaction reported. To rule out the possibility that IVR respondents were pressing random keys, we examined the relationship between the satisfaction ratings obtained in the IVR portion of the interview and other indicators of customer satisfaction. In Study 1, we calculated the multiple correlation between the six teller satisfaction items and a rating of overall satisfaction with the target bank obtained during the screener (before respondents were switched to IVR). In Study 2, we computed the multiple correlation between the four satisfaction items and the number of reported visits to the fast food chain during the 4 weeks prior to

2. Multivariate tests that examined the mode effect across all of the satisfaction items were significant for both studies as well. For Study 1, $F = 6.17$, $df = 6$, 933, $p < .001$; for Study 2, $F = 5.83$, $df = 4$, 804, $p < .001$. These differences between the IVR and CATI groups persist when we control for differences in the background characteristics of the members of the two mode groups in each study. These include respondent age, race/ethnicity, sex, and income in Study 1 and age and number of visits to fast food restaurants in the last 4 weeks in Study 2. Unfortunately, these are the only background variables available. The results are also unchanged when we include the satisfaction data from the cases in Study 1 who broke off before completing the IVR interview. On the average, the IVR cases who broke off in Study 1 gave even lower satisfaction ratings than those who completed the IVR interview. (There were only five breakoffs in Study 2, and we did not analyze the partial data they provided.)

3. The lower overall ratings in the modes that do not involve live interviewers were also apparent in the lower proportion of “top box” responses (the highest possible satisfaction ratings). The IVR respondents gave significantly fewer top box responses than did CATI respondents on all six teller satisfaction items in Study 1 and on all four satisfaction items in Study 2.

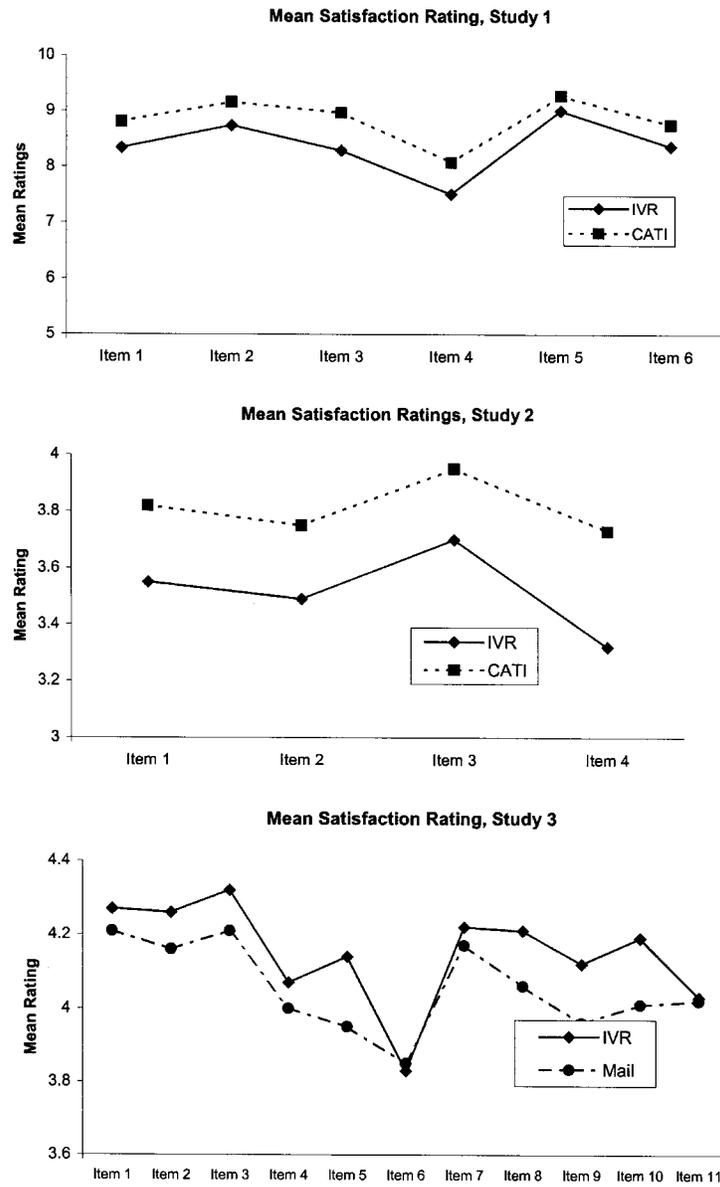


Figure 1. Mean satisfaction ratings, by study and mode of data collection

the interview. Random responding would lower these correlations, but, in both studies, the correlations are higher in the IVR than in the CATI samples (in Study 1, the R^2 is .390 for the IVR group but .301 for the CATI sample; in Study 2, the R^2 values are .104 and .087 for the IVR and CATI groups, respectively). Thus, there is little evidence that the IVR responses are any sloppier or less valid than the responses obtained through CATI.

BREAKOFFS

Although it is possible for respondents to break off a CATI interview, it happens quite rarely in general and did not occur at all in the studies reported here. Respondents probably consider it rude to hang up on a live interviewer; in addition, interviewers often provide encouragement and positive feedback to respondents, spurring them on to finish the interview. With IVR, both the barriers to quitting and the inducements to finishing the interview are reduced. Not surprisingly, breakoffs are quite common in IVR. In Study 1, there were 230 IVR breakoffs (out of 732 cases switched to IVR). Even in Study 2, where the IVR section consisted of a mere six questions, there were five breakoffs (out of 415 switched to IVR).

In Study 4, we examined the data from the breakoffs in detail to determine what types of respondents were most likely to break off the interview and when they were likely to quit. A total of 137 cases were switched to IVR in Study 4, but nine of them hung up without answering a single question, and another 53 cases failed to complete the entire questionnaire. (Operationally, we defined a breakoff as someone who started the interview but who skipped the three final items in the questionnaire and the seven key demographic items for the last household member.) By contrast, none of the CATI cases broke off. We examined rates of breaking off the IVR interview by respondent education and marital status; neither variable was significantly related to breaking off the interview without completing it.

Breaking off was nonetheless quite predictable in Study 4. The key variable was household size. The proportion of respondents who completed the IVR version of the Long Form questions dropped from 89 percent (20 respondents out of 23) to 0 percent (0 out of 10) as the household size increased from one person to five or more. Though the sample sizes are small, the linear association between the proportion breaking off and household size is statistically significant ($\chi^2 = 6.97$, $df = 1$, $p < .01$; see fig. 2). In the CATI group, each additional household member added about five and a half minutes to the length of the interview. (The CATI times give a better picture of the time needed per person, since they are unaffected by breakoffs.) That amount of extra time was apparently just too much for many of the IVR participants.

This same pattern of breakoffs by household size was apparent when we examined the completeness of the data by person within the household. In the Long Form questionnaire, respondents provide the same information about

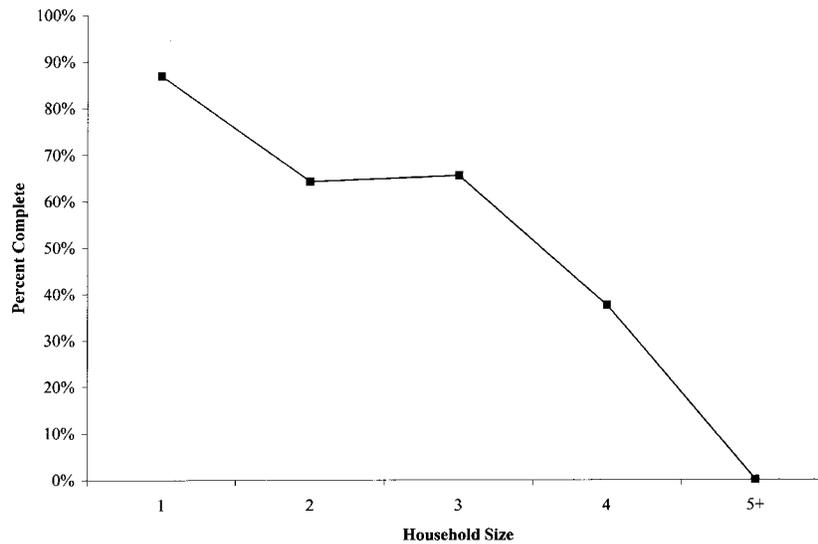


Figure 2. Percent of respondents in Study 4 providing complete data, by household size.

each household member. Thus, everyone who began the questionnaire was asked questions about “Person 1”; those in households with another member were also asked questions about Person 2; and those in larger households were asked questions about Persons 3, 4, and 5. (Our version of the questionnaire stopped at Person 5.) Participants made it to the end of Person 1 in 87.5 percent of all households; none made it to the end of Person 5. It is not that people in large households gave up immediately; some (four out of 10) got as far as Person 5 before they quit. In fact, on average, the CATI and IVR cases spent about the same amount of time—around 30 minutes—attempting to answer the questions. The difference is that the CATI cases managed to finish in that time.

Discussion

IVR has been widely adopted for market research, but few systematic evaluations have assessed its impact on the quality of the information collected. Studies done at the Bureau of Labor Statistics have looked at the feasibility of inbound IVR for establishment surveys, particularly for the collection of employment data (e.g., Clayton and Harrell 1989; Harrell and Clayton 1991; Phipps and Tupek 1990; Working, Tupek, and Clayton 1988). In addition,

Turner and his colleagues have reported results from two experimental comparisons of IVR and CATI (Gribble et al. 2000; Turner et al. 1996a). Mingay's (2000) review remains the most extensive discussion of the technology to date.

Despite their low response rates, the Gallup experiments add some resolution to our picture of this method of data collection. The results provide tentative support for two main conclusions about IVR. First, as supporters of IVR have argued, IVR appears to yield more honest answers than CATI—the average level of reported satisfaction was lower and answers to the customer satisfaction items showed less bunching at the high end of the scale under IVR than CATI. This interpretation of the results assumes that lower reported satisfaction indicates greater honesty. In addition, however, the satisfaction ratings obtained under IVR were more highly correlated with related questions on overall satisfaction (Study 1) and repeat business (Study 2). The Gallup findings are generally consistent with results reported by Turner and his colleagues, who found that respondents were more likely to report a number of sexual behaviors in an IVR interview than in a CATI interview (Turner et al. 1996a), and with the study by Gribble and his co-workers, who found that IVR respondents were more likely to report illicit drug use than were CATI respondents (Gribble et al. 2000).⁴

The second conclusion is less sanguine—IVR has a down side, providing additional opportunities for sample members to become nonrespondents. Cases who are reluctant to take part can opt out during the switch to IVR or they can quit partway through the questionnaire. Either way, they can get out of the interview without giving offense to a live interviewer. People seem quite willing to take advantage of these opportunities. In our fourth study, several participants who had just promised the interviewer that they would complete the IVR questionnaire nonetheless hung up without answering a single question. These losses during the switch to IVR can be substantial. For example, in the study by Gribble and his colleagues, about 18 percent of the sample dropped out during the switch to IVR.

IVR also sharply increases the proportion of respondents who start the interview without completing it. In Study 1, 230 of the 732 respondents switched to the IVR interview did not finish the questionnaire. The breakoff rate was even higher in Study 4, when more than 40 percent of the cases who began the IVR interview hung up without finishing it. (None of the CATI cases broke off the interview.) Cooley and his colleagues report a breakoff rate in this same range; with an IVR questionnaire that took about 30 minutes

4. Study 4 also provided some additional suggestive evidence that IVR promotes greater candor than CATI interviews. Although the items in Study 4 are, for the most part, straightforward demographic questions, some of the income questions touch on relatively stigmatized sources of income, such as welfare and Supplemental Security Income payments. The IVR respondents were more likely than CATI respondents to report both types of income, although the differences were not statistically significant.

to finish, about 24 percent of the sample broke off (Cooley et al. 2000; see also Gribble et al. 2000). In our Study 4, where the length of the questionnaire varied markedly with the number of household members, that variable correlated strongly with the rate of breakoffs. As the participants doubtless realized, the length of the questionnaire depended directly on the number of household members. We suspect that the prospect of another 5 minutes or more of dull questions disheartened many respondents, leading to the high rate of breakoffs. It seems likely that breakoffs will be common in IVR whenever respondents can foresee bad news about the amount of time required to finish the questionnaire. Keeping the questionnaire short may be the best strategy for avoiding breakoffs.

References

- Appel, M. V., R. D. Tortora, and R. Sigman. 1992. "Direct Data Entry Using Touch-Tone and Voice Recognition Technology for the M3 Survey." Research Report Series no. RR-92/01. Washington, DC: Bureau of the Census, Statistical Research Division.
- Blyth, W. G. 1997. "Developing a Speech Recognition Application for Survey Research." In *Survey Measurement and Process Quality*, ed. L. Lyberg et al., pp. 249–66. New York: Wiley.
- Clayton, R., and L. Harrell. 1989. "Developing a Cost Model of Alternative Data Collection Methods: Mail, CATI, and TDE." In *Proceedings of the Survey Research Methods Section of the American Statistical Association*, pp. 264–69. Alexandria, VA: American Statistical Association.
- Cooley, P. C., H. G. Miller, J. N. Gribble, and C. F. Turner. 2000. "Automating Telephone Surveys: Using T-ACASI to Obtain Data on Sensitive Topics." *Computers in Human Behavior* 16:1–11.
- Couper, M. P., and W. Nicholls II. 1998. "The History and Development of Computer Assisted Survey Information Collection." In *Computer Assisted Survey Information Collection*, ed. M. P. Couper et al. New York: Wiley.
- Frankovic, K. A. 1994. "Interactive Polling and Americans' Comfort Level with Technology." Paper presented at the annual meeting of the American Association for Public Opinion Research, Danvers, MA.
- Gribble, J. N., H. G. Miller, P. C. Cooley, J. A. Catania, L. Pollack, and C. F. Turner. 2000. "The Impact of T-ACASI Interviewing on Reporting Drug Use among Men Who Have Sex with Men." *Substance Use and Misuse* 80:869–90.
- Harrell, L., and R. Clayton. 1991. "A Voice Recognition Technology in Survey Data Collection: Results of the First Field Tests." Paper presented at the National Field Technologies Conference, San Diego.
- Landy, F. J., and J. L. Farr. 1980. "Performance Rating." *Psychological Bulletin* 87:72–107.
- Mingay, D. M. 2000. "The Strengths and Limitations of Telephone Audio Computer-Assisted Self-Interviewing (T-ACASI): A Review." Paper presented at the annual meeting of the American Association of Public Opinion Research, Portland, OR.
- Phipps, P., and A. Tupek, A. 1990. "Assessing Measurement Errors in a Touchtone Recognition Survey." Paper presented at the International Conference on Measurement Errors in Surveys, Tucson, AZ.
- Sears, D. O. 1983. "The Person-Positivity Bias." *Journal of Personality and Social Psychology* 44:233–50.
- Tesser, A., and S. Rosen, S. 1975. "The Reluctance to Transmit Bad News." In *Advances in Experimental Social Psychology*, vol. 8, ed. L. Berkowitz, pp. 193–232. New York: Academic Press.
- Turner, C. F., B. H. Forsyth, J. M. 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*, ed. M. P. Couper et al. New York: Wiley.

- Turner, C. F., H. G. Miller, T. K. Smith, P. C. Cooley, and S. M. Rogers. 1996a. "Telephone Audio Computer-Assisted Self-Interviewing (T-ACASI) and Survey Measurement of Sensitive Behaviors: Preliminary Results." In *Survey and Statistical Computing 1996*, ed. R. Banks et al. Chesham Bucks, U.K.: Association for Statistical Computing.
- Turner, C. F., S. M. Rogers, T. P. Hendershot, H. G. Miller, and J. P. Thornberry. 1996b. "Improving Representation of Linguistic Minorities in Health Surveys: A Preliminary Test of Multilingual Audio-CASI." *Public Health Reports* 111:279-79
- Werking, G., A. Tupek, and R. Clayton. 1988. "CATI and Touchtone Self-Response Applications for Establishment Surveys." *Journal of Official Statistics* 4:349-62.