

Computer-administered versus paper-and-pencil surveys and the effect of sample selection

NICOLAOS E. SYNODINOS and C. S. PAPACOSTAS
University of Hawaii, Honolulu, Hawaii

and

GLENN M. OKIMOTO
State of Hawaii, Department of Transportation, Honolulu, Hawaii

Airport patrons answered a self-administered questionnaire regarding their satisfaction with various airport facilities and operations. The questionnaire was administered either by a computer touch-sensitive screen or by a contextually identical paper-and-pencil version. For the latter method, respondents were selected randomly, and for the former, they were either randomly selected or self-selected. The effect of the method of questionnaire administration on the rating scales was very small when the samples were selected at random. However, there were substantial differences in ratings between self-selected and randomly selected respondents: The former gave consistently more negative ratings. These differences are probably due to psychological factors such as motivation to participate. Also, it was found that self-selected persons using the computer were more likely to make comments. The findings of this study are discussed with emphasis on their implications for computer interactive surveys.

Computer administration of surveys started in the early 1970s (Fink, 1983; Nicholls & Groves, 1986; Shanks, 1983) was used primarily for telephone interviews. Since then, the use of computer-assisted telephone interviewing (CATI) has increased rapidly and expanded to other survey research modes (Bemelmans-Spork, Keller, & Sikkell, 1985; Lyberg, 1985; Shanks & Tortora, 1985).

Paper-and-pencil instruments—interviewer-administered questionnaires and self-administered questionnaires—are still the primary method of collecting data in opinion surveys and marketing research. However, they are being supplanted by interviewer-administered or self-administered computer questionnaires. Several authors (Frey, 1983; Groves, 1983; Saris, 1991; Shanks, 1983; Synodinos & Brennan, 1988; Tanur, 1985; Tortora, 1985; Vasu & Garson, 1990) have discussed the general advantages and disadvantages of computer interactive interviewing. Others have discussed the strengths and limitations of computer interviewing as it has been applied in particular situations such as counseling (Erdman, Klein, & Greist, 1985; Greist & Klein, 1980), psychological test-

ing (Bartram & Bayliss, 1984; Burke & Normand, 1987), personnel selection (Lautenschlager & Flaherty, 1990; Martin & Nagao, 1989), and other organizational settings (Rosenfeld, Doherty, Vicino, Kantor, & Greaves, 1989).

Many of the suggested advantages of computer-administered surveys have yet to be confirmed in controlled studies. The existing literature deals mainly with CATI's potential advantages in "survey costs, timeliness, data quality, and methodological enhancements" (Nicholls & Groves, 1986, p. 100). The limited research that is available "has produced little evidence that those potentials have been achieved" (Groves & Nicholls, 1986, p. 132). Also, the findings of studies comparing computer-administered and paper-and-pencil instruments have not been unequivocal (Booth-Kewley, Edwards, & Rosenfeld, 1992). For example, in a within-subjects design measuring reported drug use/abuse among high school students, Erdman, Klein, and Greist (1983) found that results from computer interviews closely resembled findings from paper-and-pencil surveys. Similarly, Kiesler and Sproull (1986) reported considerable similarities in the responses obtained by self-administered questionnaires using either a computer or a paper-and-pencil format, "but not so much that the two may be considered interchangeable without further research" (p. 411). Skinner and Allen (1983) compared three assessment formats (computer-administered, face-to-face, and self-report clinical interviews) and found no significant differences in the respondents' reported use of tobacco, alcohol, and drugs. However, Liefeld (1987, 1988), in intercept surveys of

This research was conducted under Contract 23181 from the State of Hawaii, Department of Transportation. Opinions expressed herein are those of the authors and do not reflect an official position of the State of Hawaii Department of Transportation or the University of Hawaii. The authors wish to express their appreciation to Chuck Fisher of Magnus Communications Design, Inc., and to Myra Tamanaha of the State of Hawaii Department of Transportation. Correspondence should be addressed to N. E. Synodinos, Department of Marketing, University of Hawaii at Manoa, Honolulu, HI 96822 (e-mail: nick@uhunix.uhcc.hawaii.edu).

shopping mall patrons, reported a tendency for computer-administered questionnaires to result in more positive responses than those found in similar self-administered and interviewer-administered paper-and-pencil instruments.

Several authors (Couper, Groves, & Jacobs, 1990; de Bie, Stoop, & de Vries, 1989; Groves & Nicholls, 1986; Nicholls & Groves, 1986) have lamented the relative lack of empirical evidence and have called for further research. At this point, an all-encompassing framework that delineates the boundaries of similarities and differences between paper-and-pencil and computer-administered instruments does not exist. Individual studies in various contexts are slowly appearing in the literature, and there is a need to increase their numbers, especially in real-world applications. In the future, a detailed analysis of such studies (e.g., meta-analysis) may provide "the boundary contextual conditions that produce differences in computer versus paper-and-pencil responses" (Booth-Kewley et al., 1992, pp. 565–566). The present article contributes to the need for additional studies in real-world applications by describing a field study in this area.

Background and Purpose

The Honolulu International Airport (HIA) is operated by the State of Hawaii Department of Transportation (HDOT) and is one of the busiest airports in the country. It processes more than 20 million enplaning and deplaning passengers annually (Department of Business, Economic Development & Tourism, 1993).

Being responsive to user expectations, the HDOT has been conducting opinion surveys of airport patrons. In the past, airport patrons were selected by facilitators and given a paper-and-pencil self-administered questionnaire about their satisfaction with airport facilities and operations. Subsequently, a similar instrument was used as an unattended computer self-administered questionnaire (presented on a touch-sensitive screen connected to a stand-alone computer); respondents were self-selected.

The Questionnaire

The touchscreen instrument asks all respondents to rate 12 particular airport facilities (appearance, airport roads, airport parking, shuttle bus, loading zones, directional signs, baggage areas, restaurant/snackbars, gift shops, public conveniences, visitor information, cleanliness) and give an overall rating for the HIA. The choices for each of these questions range from 0 (*very poor*) to 10 (*very good*) plus the option *don't know*. In addition to the rating scales, the survey includes questions about basic demographic characteristics and airport usage. The last question consists of the presentation of a screen containing a simulated typewriter keyboard. This allows the respondent to enter comments by "typing" letters on the touchscreen.

For the present study, a paper-and-pencil instrument was constructed so that it would be identical to the touchscreen survey in terms of question phrasing and sequencing. The effect of intrinsic differences between the methods on the obtained responses is the subject of this study.

METHOD

Sample Selection

Three conditions were compared: (1) the computer self-administered questionnaire completed by randomly selected respondents (C/RS); (2) the paper-and-pencil self-administered questionnaire completed by randomly selected respondents (PP/RS); and (3) the computer self-administered questionnaire completed by self-selected respondents (C/SS).

The C/RS and PP/RS samples consisted of randomly selected persons who agreed to participate and completed either the attended computer self-administered questionnaire or the attended paper-and-pencil self-administered questionnaire. The C/SS sample consisted of self-selected persons who completed the touchscreen questionnaire unattended. This sample was included for purposes of comparison because the HDOT had used the touchscreen survey as an unattended questionnaire.

Randomly Selected Groups

Randomly selected adults were approached by facilitators near the touchscreen system, which was located in the central concourse of the HIA. The facilitators approached potential respondents between 8:00 a.m. and midnight during 36 data collection sessions spanning a period of 79 consecutive days.

The potential respondents were told that their answers would remain confidential, and they were either directed to the computer or given a clipboard with the paper-and-pencil questionnaire. The facilitators completed a contact record sheet (CRS) for each person approached. The *observed* age and gender were recorded along with information that allowed the concatenation of data of a particular questionnaire and its corresponding CRS.

Altogether, 1,645 persons were asked to participate in the study. The persons *approached*, for the two randomly selected groups (PP/RS and C/RS), were similar in terms of observed age [$\chi^2(6, N = 1,639) = 7.08, p = .31$] and gender [$\chi^2(1, N = 1,645) = 0.01, p = .93$].

Thirty-nine percent of the persons approached agreed to participate and completed the survey. Specifically, 38% (330 out of 878) participated in the PP/RS group and 41% (312 out of 767) in the C/RS group. These proportions were statistically similar [$\chi^2(1, N = 1,645) = 1.52, p = .22$], indicating that there was no difference in the refusal rates for the two conditions. Furthermore, those who agreed to complete the paper-and-pencil version were statistically similar with the persons who agreed to do the computer version in terms of observed age [$\chi^2(6, N = 636) = 3.54, p = .74$] and gender [$\chi^2(1, N = 642) = 0.05, p = .83$]. Therefore, in terms of these important variables, the sample selection process was successful in creating equivalent groups.

Self-Selected Group

The C/SS sample was drawn from computer records created by self-selected persons. These records were created at the same location and at approximately during the same time period as were the records created by the randomly selected persons. Specifically, the C/SS sample was selected from computer records that were created immediately before each of the 36 data collection sessions. In some cases, in which the computer-generated file did not contain a sufficient number of records before a session (as in the case

of file start-up), the cases were selected from records immediately following that data collection session.

Three hundred and fifty-seven computer records¹ were selected in this manner. One hundred and ten of these records consisted entirely of the code indicating that the touchscreen questionnaire timed out after someone started it without answering *any* of the questionnaire screens. Consequently, the usable sample for the C/SS condition was 247 surveys.

RESULTS

The questionnaire can be divided into three broad categories of questions: respondent characteristics, facility rating scales, and propensity to comment. This section compares the responses between the three samples in terms of these three categories of questions. Comparisons of the two randomly selected groups (C/RS vs. PP/RS) provide information regarding effects of the survey administration method. Comparisons between the groups completing the touchscreen questionnaire (C/RS vs. C/SS) provide information regarding the effect of the sample selection process. Comparisons between the PP/RS and C/SS provide information about the joint effect of different questionnaire administration methods and sample selection processes.

An observational study (Synodinos, Papacostas, & Okimoto, 1989) demonstrated that records created by self-selected persons should be filtered. Therefore, for the self-selected group, only complete (i.e., an answer was recorded in the last closed-ended question) records from persons reporting ages 20–69 were retained for analyses. For the sake of consistency, these criteria were followed also for retaining surveys for the two randomly selected groups. The resulting samples for each condition were 279 for the PP/RS, 266 for the C/RS, and 129 for the C/SS.

Comparisons of Respondent Characteristics

Table 1 summarizes the respondents in terms of several important characteristics. The respondents in the three groups were statistically similar in their purpose at the airport [$\chi^2(4, N = 672) = 3.43, p = .49$], education [$\chi^2(12, N = 673) = 17.35, p = .14$], and income [$\chi^2(14, N = 655) = 21.82, p = .08$]. The three groups differed significantly in proportions of residents and non-residents of the State of Hawaii [$\chi^2(2, N = 674) = 6.76, p = .03$], and in reported age [$\chi^2(8, N = 674) = 21.68, p = .006$].

The percentage of residents (see Table 1) in the randomly selected group for the touchscreen version was significantly higher than that of those selected randomly for the paper-and-pencil instrument [$\chi^2(1, N = 545) = 4.37, p = .04$] and those in the self-selected group [$\chi^2(1, N = 395) = 3.86, p = .05$]. There was no statistical difference [$\chi^2(1, N = 408) = 0.07, p = .79$] in the proportion of residents vis-à-vis nonresidents between the paper-and-pencil and the self-selected groups.

The self-selected group attracted a higher percentage of younger persons (see Table 1) than did the randomly selected group completing the touchscreen survey [$\chi^2(4,$

$N = 395) = 18.83, p = .001$] or the paper-and-pencil instrument [$\chi^2(4, N = 408) = 18.22, p = .01$]. The two randomly selected groups were similar in terms of reported age [$\chi^2(4, N = 545) = 0.55, p = .97$]. The similarity of the two randomly selected groups in terms of *reported* age agrees with the findings of the *observed* age recorded on the CRS by the facilitators. This indicates that respondents, at least on this question, answered the questionnaire truthfully.

Comparisons of Rating Scales

The 13 facility rating questions represented the respondents' perceptions of various aspects of the airport. The relationship between the ratings as a function of treatment condition was examined in a series of *t* tests. The possibility of using a multivariate approach was ruled out because of the existence of missing values (either *unanswered* or *don't know*). Procedures for estimating missing values are controversial (Anderson, Basilevsky, & Hum, 1983), especially in cases such as the present one, where these values were not necessarily random.²

Although the ratings of the PP/RS group were slightly higher than those of the C/RS in all but one of the 13 scales, they reached statistical significance in only two of these scales [restaurant/snackbars, $t(380) = 2.29, p = .03$; gift shops, $t(413) = 2.91, p = .004$]. Therefore, the preponderance of the findings support the conclusion that the responses obtained via these two methods of survey administration are similar when the samples are selected randomly.

All the mean ratings of the C/SS respondents were consistently lower than those of the C/RS group (see Table 2). That is, self-selected persons were more critical about the airport. The differences reached significance in the scales of directional signs [$t(315) = 2.62, p = .009$], gift shops [$t(322) = 2.77, p = .006$], public conveniences [$t(365) = 2.58, p = .01$], visitor information [$t(264) = 2.01, p = .05$], and overall airport rating [$t(392) = 2.01, p = .05$], but they did not reach significance in the other scales. It can be concluded that there are several differences in the ratings when the questionnaire administration method is held constant and the respondents are selected differently.

All the mean ratings by respondents in the C/SS condition were consistently lower than those of the PP/RS sample (see Table 2). These differences reached significance (at the .05 level and beyond) in 9 of the 13 rating scales. The differences did not reach significance for appearance [$t(402) = 1.37, p = .17$], airport parking [$t(192) = 1.91, p = .06$], shuttle bus [$t(250) = 1.74, p = .08$], and cleanliness [$t(402) = 1.27, p = .21$]. Therefore, these comparisons indicated that there are substantial differences in the ratings obtained from groups that use different methods of questionnaire administration and are selected differently.

As previously noted, the three groups differed in their age and in residency distributions of respondents. The

Table 1
Respondents' Characteristics by Condition

Characteristics	Condition					
	PP/RS		C/RS		C/SS	
	<i>N</i>	%	<i>N</i>	%	<i>N</i>	%
Purpose						
Leaving	214	77.3	192	72.2	91	70.5
Arriving	20	7.2	22	8.3	14	10.9
Other	43	15.5	52	19.5	24	18.6
Residency Status						
Nonresident of Hawaii	218	78.1	186	69.9	103	79.8
Resident of Hawaii	61	21.9	80	30.1	26	20.2
Education						
< High school diploma	2	0.7	5	1.9	2	1.6
High school diploma	18	6.5	34	12.8	15	11.6
Trade/vocational diploma	12	4.3	20	7.5	14	10.9
Some college	78	28.1	65	24.4	27	20.9
College degree	87	31.3	73	27.4	37	28.7
Some postgraduate	23	8.3	23	8.6	8	6.2
Postgraduate degree	58	20.9	46	17.3	26	20.2
Age						
20–29	76	27.2	76	28.6	53	41.1
30–39	75	26.9	65	24.4	43	33.3
40–49	65	23.3	66	24.8	18	14.0
50–59	41	14.7	38	14.3	6	4.7
60–69	22	7.9	21	7.9	9	7.0
Income						
Under \$ 15,000	8	3.1	20	7.5	8	6.2
\$15,000 to \$19,999	14	5.4	21	7.9	3	2.3
\$20,000 to \$24,999	16	6.2	20	7.5	9	7.0
\$25,000 to \$34,999	39	15.0	32	12.0	12	9.3
\$35,000 to \$49,999	49	18.8	52	19.5	19	14.7
\$50,000 to \$64,999	46	17.7	35	13.2	17	13.2
\$65,000 to \$74,999	23	8.8	20	7.5	16	12.4
\$75,000 and over	65	25.0	66	24.8	45	34.9

Note—PP/RS, paper-and-pencil self-administered questionnaire, randomly selected; C/RS, computer self-administered questionnaire, randomly selected; C/SS, computer self-administered questionnaire, self-selected; *N*, number of respondents.

self-selected group attracted a significantly larger percentage of younger respondents than did those selected for the other groups. The PP/RS and the C/SS had approximately equal proportions of residents and nonresidents, and the C/RS had a significantly higher proportion of residents than did the other two groups. Two-way analyses of variance were conducted to determine the effects, if any, of these differences on the obtained ratings.

There were no main effects of age and no interactions involving age. Therefore, in this context, age does not contribute to the differences/similarities in the ratings obtained from respondents in the three conditions.

On the other hand, there were significant main effects of residency: Ratings of residents of Hawaii were lower than those of visitors. There were basically no significant interactions between method and residency. The main effect of residency would explain the small differences in the ratings between the randomly selected groups (PP/RS and C/RS), because there were more residents (who are more likely to give lower ratings) in the latter. Consequently, taking into account the effect of

residency brings the C/RS closer to the ratings of the PP/RS. This strengthens the previously stated conclusion that the method of questionnaire administration has no effect on the obtained ratings when the groups are selected randomly.

The significantly lower ratings given by the self-selected group in comparison with the two randomly selected groups are not explained by residency but must be attributed to other factors. These factors are elaborated further in the Discussion.

Comparisons of Propensity to Comment

The analyses showed that significantly [$\chi^2 (2, N = 665) = 55.42, p = .001$] different proportions of the three groups offered comments. Specifically, 49.6% of the self-selected group wrote comments with the touch-screen simulated keyboard. This exceeded the 36.6% of the randomly selected respondents who wrote comments on the paper-and-pencil questionnaire. Only 15.1% of the randomly selected respondents using the computer provided comments.

Table 2
Respondents' Ratings of the Airport by Condition

Rating Scales	Condition								
	PP/RS			C/RS			C/SS		
	<i>M</i>	<i>SD</i>	<i>N</i>	<i>M</i>	<i>SD</i>	<i>N</i>	<i>M</i>	<i>SD</i>	<i>N</i>
Appearance	7.16	2.55	277	6.86	2.67	264	6.77	2.76	127
Airport roads	7.00	2.29	228	6.77	2.40	235	6.30	2.92	115
Airport parking	6.29	2.38	117	6.09	2.80	153	5.56	2.92	77
Shuttle bus	6.35	2.68	159	6.19	3.10	162	5.68	3.41	93
Loading zones	6.25	2.61	195	6.12	2.99	219	5.51	3.00	103
Directional signs	6.65	2.54	211	6.94	2.70	216	5.79	3.15	101
Baggage areas	6.97	2.30	201	6.50	2.60	208	6.13	2.99	109
Restaurant/snackbars	6.25	2.48	195	5.63	2.81	187	5.18	3.23	103
Gift shops	6.72	2.18	204	6.05	2.48	211	5.19	2.97	113
Public conveniences	7.20	2.18	220	7.09	2.37	246	6.37	2.77	121
Visitor information	6.91	2.54	175	6.72	2.68	174	5.99	3.05	92
Cleanliness	7.79	2.09	275	7.80	2.11	266	7.49	2.40	129
Overall	7.60	2.01	274	7.43	1.97	265	6.97	2.49	129

Note—PP/RS, paper-and-pencil self-administered questionnaire, randomly selected; C/RS, computer self-administered questionnaire, randomly selected; C/SS, computer self-administered questionnaire, self-selected; *N*, number of respondents.

DISCUSSION

In this article, the touchscreen survey has been compared with a contextually similar paper-and-pencil version. The C/SS and PP/RS were selected because these were normally used at HIA. In addition, data were collected from a randomly selected group of respondents who used the touchscreen. Comparisons between these three groups captured the effects of method, the effects of sample selection, and their combined effects.

An important comparison was that between randomly selected respondents completing the survey via two methods of questionnaire administration. It was found that in samples selected at random (C/RS vs. PP/RS) the effect of the method of questionnaire administration was small. The difference was even smaller when residency (the only demographic variable that was found to have a significant effect on ratings) was taken into account. This finding, that similar results are obtained through the use of these two methods of questionnaire administration, agrees with previous findings (Erdman et al., 1983; Kiesler & Sproull, 1986; Skinner & Allen, 1983) and extends them to a different context.

However, the present results differ from those of Liefeld (1987, 1988), who reported a tendency for higher mean ratings in computer-administered surveys than those found in paper-and-pencil instruments. Liefeld's study had two things in common with this study's comparison of the two randomly selected groups. First, the respondents were selected by facilitators among patrons of a facility. Second, the present study and Liefeld's obtained ratings of object characteristics (not necessarily ego involving) such as those routinely measured in consumer psychology and marketing research studies. Liefeld attributed the obtained differences, in part, to response effects caused by the initial placement of the computer screen cursor. Variations among computer software and hardware in their han-

dling of various formats of questions is an important consideration. Therefore, application-specific assessments, such as the one described here, should be routinely conducted to determine the extent and direction of any effects that could be attributed to the survey mode. The advice of Kiesler and Sproull (1986) that "at a minimum, electronic survey results should be calibrated against those obtained through more conventional means" (p. 412) is appropriate. It is recommended that, in situations where different methods are employed, studies similar to the one described in this paper are necessary in order to determine the compatibility and proper interpretation of the results.

In its original use at the HIA, the touchscreen survey relied on self-selected respondents' completing the computer version of the questionnaire. Therefore, a sample of self-selected respondents that used the touchscreen survey was drawn. Their responses were compared with those of the randomly selected respondents, and these comparisons revealed substantial differences between the self-selected and the randomly selected groups.

The fact that there are differences between self-selected and randomly selected samples is not unexpected. However, the direction and degree of these differences is not predictable a priori because reasons of self-selection vary from situation to situation. For example, in a different context (a survey of subscribers to the oceanographic division of a computer network regarding their scientific endeavors and network experiences), self-selected respondents provided higher quality and more extensive answers (Walsh, Kiesler, Sproull, & Hesse, 1992). These researchers suggested that "planned comparisons of self-selected samples with random samples under a variety of conditions in more heterogeneous groups could contribute to our understanding of the nature of motivated response and participation bias" (p. 244). The present study has contributed another

block to the foundation that will support the future development of a general framework for understanding the complexity of this phenomenon.

In the present study at the HIA, surveys administered to self-selected respondents had lower ratings than did those obtained from randomly selected persons who completed the survey. Two possible reasons for this difference can be postulated.

First, it is possible that self-selected persons may *feel* more anonymous than respondents in the randomly selected conditions and are thus less inclined to give socially desirable answers in the form of high ratings. Regarding social desirability, Greist and Klein (1980) suggested that answers to sensitive or potentially embarrassing questions can be obtained more readily on computer self-administered questionnaires than in personal interviews conducted by trained professionals such as psychiatrists. However, the literature regarding the effects of social desirability in computer-administered surveys is inconclusive, as has been illustrated by recent studies (Booth-Kewley et al., 1992; Lautenschlager & Flaherty, 1990; Martin & Nagao, 1989). Given the type of rating questions asked in the present study (e.g., not particularly ego involving), the effect of social desirability is probably small.

Second, and most likely, self-selected respondents are motivated to participate for different reasons than randomly selected respondents are. It was shown that demographic characteristics (including residency) did not explain the difference in ratings between the self-selected and the randomly selected respondents. Consequently, the differences in the ratings are probably due to psychological factors such as motivation to participate. This explanation is reinforced by the fact that self-selected persons were more likely to make comments even though it required more effort to enter responses on a computer keyboard than to write them on a questionnaire. The least likely group to provide comments comprised the randomly selected persons using the touchscreen.

The study provides evidence that when respondents are self-selected, in situations such as the present one, lower ratings are to be expected. This should be taken into account when one is using data obtained from self-selected individuals for evaluating performance characteristics in applied behavioral studies.

Comments from self-selected respondents are probably negatively biased. However, from an applied perspective, the high volume of these comments, coupled with the fact that they can be collected continuously, makes the touchscreen system a useful source of information for monitoring the performance of a facility. Comments obtained in this manner can provide information about various issues that are not included in the structured questions, and user concerns can be considered in a timely manner.

Finally, it is important to repeat that there is a need for additional systematic comparisons of computer-administered

questionnaires and traditional methods in diverse contexts. Much of the prior evidence about the advantages of computer interviewing and related issues consists predominantly of "unpublished comparisons and impressionistic evidence" (Catlin & Ingram, 1988, p. 437). The evidence accumulated from systematic studies will eventually contribute to an understanding of the circumstances and contexts in which computer questionnaires are similar and those in which differences exist. Thus, the accumulation of empirical evidence about these issues will be valuable to researchers and practitioners of behavioral and social research.

REFERENCES

- ANDERSON, A. B., BASILEVSKY, A., & HUM, D. P. J. (1983). Missing data: A review of the literature. In P. H. Rossi, J. D. Wright, & A. B. Anderson (Eds.), *Handbook of survey research* (pp. 415-494). Orlando, FL: Academic Press.
- BARTRAM, D., & BAYLISS, R. (1984). Automated testing: Past, present and future. *Journal of Occupational Psychology*, *57*, 221-237.
- BEMELMANS-SPORK, M. E. J., KELLER, W., & SIKKEL, D. (1985). Use of handheld microcomputers in the Netherlands Central Bureau of Statistics. In *Bureau of the Census First Annual Research Conference: Proceedings* (pp. 47-52). Reston, VA: Bureau of the Census.
- BOOTH-KEWLEY, S., EDWARDS, J. E., & ROSENFELD, P. (1992). Impression management, social desirability, and computer administration of attitude questionnaires: Does the computer make a difference? *Journal of Applied Psychology*, *77*, 562-566.
- BURKE, M. J., & NORMAND, J. (1987). Computerized psychological testing: Overview and critique. *Professional Psychology: Research & Practice*, *18*, 42-51.
- CATLIN, G., & INGRAM, S. (1988). The effects of CATI on costs and data quality: A comparison of CATI and paper methods in centralized interviewing. In R. M. Groves, P. R. Biemer, L. E. Lyberg, J. T. Massey, W. L. Nicholls, & J. Waksberg (Eds.), *Telephone survey methodology* (pp. 437-450). New York: Wiley.
- COUPER, M., GROVES, R. M., & JACOBS, C. A. (1990). Building predictive models of CAPI acceptance in a field interviewing staff. In *Bureau of the Census 1990 Annual Research Conference: Proceedings* (pp. 685-702). Arlington, VA: Bureau of the Census.
- DE BIE, S. E., STOOP, I. A. L., & DE VRIES, K. L. M. (1989). *CAI software: An evaluation of software for computer assisted interviewing*. Amsterdam: VOI Association of Social Research Institutes.
- DEPARTMENT OF BUSINESS, ECONOMIC DEVELOPMENT & TOURISM (1993). *The State of Hawaii Data Book 1992*. Honolulu: Author.
- ERDMAN, H. P., KLEIN, M. H., & GREIST, J. H. (1983). The reliability of a computer interview for drug use/abuse information. *Behavior Research Methods & Instrumentation*, *15*, 66-68.
- ERDMAN, H. P., KLEIN, M. H., & GREIST, J. H. (1985). Direct patient computer interviewing. *Journal of Consulting & Clinical Psychology*, *53*, 760-773.
- FINK, J. C. (1983). CATI's first decade: The Chilton experience. *Sociological Methods & Research*, *12*, 153-168.
- FREY, J. H. (1983). *Survey research by telephone*. Beverly Hills, CA: Sage.
- GREIST, J. H., & KLEIN, M. H. (1980). Computer programs for patients, clinicians, and researchers in psychiatry. In J. P. Sidowski, J. H. Johnson, & T. A. Williams (Eds.), *Technology in mental health care delivery systems* (pp. 161-181). Norwood, NJ: Ablex.
- GROVES, R. M. (1983). Implications of CATI: Costs, errors, and organization of telephone survey research. *Sociological Methods & Research*, *12*, 199-215.
- GROVES, R. M., & NICHOLLS, W. L., II. (1986). The status of computer-assisted telephone interviewing: Part II—Data quality issues. *Journal of Official Statistics*, *2*, 117-134.
- KIESLER, S., & SPROULL, L. S. (1986). Response effects in the electronic survey. *Public Opinion Quarterly*, *50*, 402-413.

- LAUTENSCHLAGER, G. J., & FLAHERTY, V. L. (1990). Computer administration of questions: More desirable or more social desirability? *Journal of Applied Psychology*, *75*, 310-314.
- LIEFELD, J. P. (1987). Computer interviewing: Cursor and response task effects. *Canadian Journal of Marketing Research*, *6*(Nov.), 20-25.
- LIEFELD, J. P. (1988). Response effects in computer-administered questioning. *Journal of Marketing Research*, *25*, 405-409.
- LYBERG, L. (1985). Plans for computer-assisted data collection at Statistics Sweden. In *Proceedings of the 45th Session, International Statistical Institute* (Book 3, Topic 18.2, pp. 1-11). Amsterdam: International Statistical Institute.
- MARTIN, C. L., & NAGAO, D. H. (1989). Some effects of computerized interviewing on job applicant responses. *Journal of Applied Psychology*, *74*, 72-80.
- NICHOLLS, W. L., II, & GROVES, R. M. (1986). The status of computer-assisted telephone interviewing: Part I—Introduction and impact on cost and timeliness of survey data. *Journal of Official Statistics*, *2*, 93-115.
- ROSENFELD, P., DOHERTY, L. M., VICINO, S. M., KANTOR, J., & GREAVES, J. (1989). Attitude assessment in organizations: Testing three microcomputer-based survey systems. *Journal of General Psychology*, *116*, 145-154.
- SARIS, W. E. (1991). *Computer-assisted interviewing*. Newbury Park, CA: Sage.
- SHANKS, J. M. (1983). The current status of computer-assisted telephone interviewing: Recent progress and future prospects. *Sociological Methods & Research*, *12*, 119-142.
- SHANKS, J. M., & TORTORA, R. D. (1985). Beyond CATI: Generalized and distributed systems for computer-assisted surveys. In *Bureau of the Census First Annual Research Conference: Proceedings* (pp. 358-371). Reston, VA: Bureau of the Census.
- SKINNER, H. A., & ALLEN, B. A. (1983). Does the computer make a difference? Computerized versus face-to-face self-report assessment of alcohol, drug, and tobacco use. *Journal of Consulting & Clinical Psychology*, *51*, 267-275.
- SYNODINOS, N. E., & BRENNAN, J. M. (1988). Computer interactive interviewing in survey research. *Psychology & Marketing*, *5*, 117-137.
- SYNODINOS, N. E., & BRENNAN, J. M. (1990). Evaluating microcomputer interactive survey software. *Journal of Business & Psychology*, *4*, 483-493.
- SYNODINOS, N. E., PAPACOSTAS, C. S., & OKIMOTO, G. M. (1989). *Evaluation of computer interactive data collection at the Honolulu International Airport*. Project report prepared for the State of Hawaii Department of Transportation, Honolulu.
- TANUR, J. M. (1985). Survey research methods in environmental psychology. In A. Baum & J. E. Singer (Eds.), *Advances in environmental psychology: Vol. 5. Methods and environmental psychology* (pp. 147-184). Hillsdale, NJ: Erlbaum.
- TORTORA, R. D. (1985). CATI in an agricultural statistical agency. *Journal of Official Statistics*, *1*, 301-314.
- VASU, M. L., & GARSON, G. D. (1990). Computer-assisted survey research and continuous audience response technology for the political and social sciences. *Social Science Computer Review*, *8*, 535-557.
- WALSH, J. P., KIESLER, S., SPROULL, L. S., & HESSE, B. W. (1992). Self-selected and randomly selected respondents in a computer network survey. *Public Opinion Quarterly*, *56*, 241-244.

NOTES

1. That is, the number of records selected for the touchscreen self-selected group was equal to the 357 records created by the randomly selected group. For the C/RS, typically each one of the created computer records represented a survey completed in its entirety by each of the 312 respondents. However, 45 additional computer records were created accidentally by respondents or by persons not selected by the facilitators. These invalid records were identified by the facilitators' notes on the CRS and were excluded.

2. In the paper-and-pencil questionnaire, it is possible to leave an item unanswered and continue. However, in the computer version, as originally programmed by the independent consultant, "refusing" to answer an item causes an abort from that point on (this does not apply to questions that provide the option *don't know*), resulting in missing values for that and subsequent items. It would have been more appropriate had the respondents been given the option to refuse to answer particular items and still continue (Synodinos & Brennan, 1990).

(Manuscript received January 11, 1993;
revision accepted for publication March 11, 1994.)