Journal of Official Statistics, Vol 11, No. 2, 1995, pp. 141–146

Data Quality in a CAPI Survey: Keying Errors

Lynn Dielman and Mick P. Couper[1]

This paper examines interviewer keying errors in a CAPI survey. Recorded responses to 16,778 closed-ended questions in 116 interviews were compared to audio tapes of the answers provided by respondents. A total of only 16 keying errors were detected, for an overall error rate of 0.095 %. This finding helps to alleviate concerns about keying errors to fixed-response questions in a CAPI survey.

Key words: Computer assisted interviewing; interviewer errors.

1. Introduction

The use of computer assisted interviewing (CAI) for telephone surveys has been well-established for several years, and the method is rapidly gaining widespread acceptance for face to face surveys (see Weeks (1992) for a review). Among the many advantages cited for both CATI (computer assisted telephone interviewing) and CAPI (computer assisted personal interviewing) are improvements in data quality resulting from increased control over the interviewing process. This is achieved primarily through automated skip patterns, consistency and range checks, and "fills" to adapt question wording to particular situations or respondents (see, for example, Martin, O'Muircheartaigh, and Curtice 1993).

In spite of these advantages of computer assisted interviewing, there are still concerns about different types of errors that may be introduced by CAI. One of these concerns relates to keying errors. It is argued that because interviewers are not necessarily expert typists or because they may be nervous using a computer, they may make a number of errors of miskeying that go undetected by the instrument's error-trapping facilities (see Baker 1992, p.151). This may be especially true of CAPI interviewers because of the compactness of the laptop computer keyboard and the possible awkward positioning of the computer for in-home interviewing.

Two general types of keying errors can occur in CAI surveys. The first is in the keying of open-ended responses or marginal notes. Although there is some evidence that

[1] Lynn Dielman is a Senior Research Associate and Mick Couper is a Study Director, both in the Survey Research Center at the University of Michigan, P.O. Box 1248, Ann Arbor, MI 48106, U.S.A. This research was supported by grant R01 AG08279-03 from the National Institute on Aging to principal investigators Regula Herzog and Hazel Markus. We are grateful to Regula Herzog, William Nicholls and two anonymous reviewers for their helpful comments.

the *amount* of open-ended text entered in CAI may be less than in paper and pencil interviewing (PAPI) (Groves and Nicholls 1986), there is little evidence that the *quality* of such responses is diminished (Bernard 1989; Catlin and Ingram 1988). The second type of keying error is that due to the miskeying of closed-ended responses. This paper addresses the second source of error.

Despite concerns about data quality in CAI surveys, virtually no research has focussed on keying errors in closed-ended responses. In a study of CATI interviewers, Kennedy, Lengacher and Demerath (1990) used reports from monitors in a centralized telephone facility to detect keying errors. They found keying error rates to average about 0.62% over four different surveys. To our knowledge, no study has examined keying error rates in a CAPI survey. One possible reason for this may be the expense associated with obtaining independent measures of what was actually said by respondents, against which the recorded response could be compared.

The paucity of research on keying or entry errors by interviewers extends to paper and pencil surveys. Few schemes for coding interviewer behavior, whether from tape recorded personal visit interviews (e.g., Cannell, Lawson, and Hausser 1975) or from monitored telephone interviews (e.g., Cannell and Oksenberg 1988), have included codes for recording errors. In an evaluation of mock interviews, Rustemeyer (1977) compared interviewer entries with recorded verbal responses. Rates of inconsistency between recorded entries and verbal responses ranged from 3.9% of entries for experienced interviewers to 6.2% of entries for newly trained interviewers. These figures include both incorrect entries and failures to make an entry where appropriate. Rustemeyer also examined clerical errors by interviewers in transcribing data from a continuing household roster to the interview form for that month. Experienced interviewers made transcription errors for 0.8% of items and failed to make required entries for an additional 0.5% of items. For newly trained interviewers the corresponding figures were 0.9% and 1.3%. While it is unclear how well the mock interview situation represents actual field experiences, these figures suggest that clerical errors are not uncommon in paper and pencil interviewing.

2. Keying Error Study Design

The present study aims to redress the lack of research on keying errors in a CAI survey by examining errors made by interviewers to closed-ended questions. This was made possible through the tape recording of a systematic subsample of interviews in a survey about self-concepts over the life span ("Self Portraits") in the Detroit metropolitan area in 1992 (Herzog 1994). This was a clustered probability sample of a three-county area in Southeastern Michigan fielded by the Survey Research Center at the University of Michigan from February–August 1992. Forty-eight interviewers completed 1,471 face to face interviews with respondents aged 30–95 using T2000SX Toshiba (386) laptop computers with the survey instrument programmed in AutoQuest.

None of the 26 experienced or 22 newly-trained interviewers who worked on the Self Portraits Study had previous CAPI experience (indeed, most had no real computer experience at all). Besides any usual general interviewer training and study

specific substantive training, 3 days of supervised and home practice with computer assisted survey techniques helped the interviewers become comfortable with the use of the laptop computer. The visual programming of the instrument produced consistent spacing and placement of questions and response categories on each screen. Almost all answers were keyed by pressing the *number* of the correct category and then pressing ENTER to proceed to the next screen.

The interview lasted an average of 58 minutes and was divided into two main parts. The first half (26 minutes average) consisted of two components: requests for open-ended descriptions of the respondent's current, past and possible selves; and a computer assisted self-interview (CASI) where the respondents were taught to use a restricted laptop keyboard to rate themselves several times in 44 domains. This part of the interview was not used in the keying error study.

The latter half of the questionnaire (32 minutes average) consisted of varied categories for several topics of closed-ended questions, including frequency of participation in various activities; functional and cognitive health; occupation; income and assets; social support; implicit theories of aging; life events; and satisfaction. The software made it possible for interviewers to back up easily to correct a checkpoint answer and then automatically proceed down a correct path. Depending on skip patterns, there were 130 to 194 closed-ended questions checked in this half of the interview.

A quarter of the closed-ended questions examined in this study had only two possible responses (yes/no, true/false, etc.), while a third had five response options (a variety of ordinal scales), 10 % required continuous responses (e.g., age, years of education), and the remainder had a varying number of responses. All these questions required input using the numeric keys.

3. Assessment of Keying Errors

Interviewers were asked to tape 2 of their first 10 interviews (3rd and 7th depending on the permission of the respondent and the availability of tape recorders) and then about every 10th interview. The preliminary taping was proposed so the supervisors could use the tapes for general interviewer evaluation, but was continued through 1,353 cases for this study. All taped interviews were evaluated by field team leaders for interviewing techniques and then turned over to the study staff for further processing. Although the interviewers knew that their team leader would be listening to the tapes to evaluate probing and pace, they were not aware that keying errors would be studied.

Of the 48 interviewers who worked on the Self Portraits study, 6 did less than 3 interviews, none of which were taped; 1 interviewer with 13 interviews had no usable tapes. The remaining 41 interviewers in the keying error study took between 5 and 75 interviews each and produced 116 usable tapes for this study. Four of the tapes received were unusable for a variety of reasons (e.g., the interviewer forgot to turn the cassette over). The 116 interviews yielded a total of 16,778 responses verified using the tapes, for an average of 145 responses per interview.

A hard copy of all variables in the data set was printed as a worksheet for each of

the 116 taped interviews. Study staff listened to and matched each closed-ended variable with the answer given by the respondent on tape. Senior staff "check listened" to 26 complete random interviews (no errors in coding were found) and every keying error or questionable answer was rechecked. There were four false-positive errors coded which the senior staff determined were unusual situations in which the correct judgement was made by the interviewer.

The number of closed-ended answers was counted for each case – the interview variable base differed in number of questions due to skip patterns, and also occasional non-recorded variables missed because of tape recording situations such as changing tape cassettes or unintelligible answers.

4. Results

Of the 16,778 answers that were checked for accuracy, there were only 16 final keying errors for an overall error rate of 0.095%, or less than 1 error for every 1,000 entries. A summary of the 16 errors is presented in Table 1. Only 9 of the 41 interviewers had

Table 1 Description of keying errors detected

Interviewer	Interviewer experience	Interview number	Errors	Description of errors
A	New hire	6th of 46	1	Keyed 2 (quite often) instead of 3 (sometimes)
B	New hire	4th of 60	1	Keyed 3 (sometimes) instead of 4 (not very often)
C	Experienced	4th of 6	1	Keyed 1 (excellent) instead of 2 (very good)
D	New hire	3rd of 6	3	Keyed 3 (somewhat disagree) instead of 4 (strongly disagree) on two adjacent questions. Keyed 1 (strongly agree) instead of 2 (somewhat agree)
E	Experienced	12th of 74	1	Keyed 5 (no) instead of 1 (yes)
F	Experienced	3rd of 47	2	Keyed 2 instead of 3 (7-point scale) on two adjacent questions
G	New hire	3rd of 19	1	Keyed 5 (no) instead of 1 (yes)
H	New hire	3rd of 41	1	Keyed 2 (quite often) instead of 3 (sometimes)
I	Experienced	4th of 73	1	Did not change entry when R corrected answer from 15 to 10 (no. of cigarettes smoked)
I		6th of 73	3	Keyed 1950 when R said "50 years old" to question on "what year did you end a job". Then keyed 1 (self employed) instead of 2 (someone else) and 1 (yes) instead of 5 (no) on next two questions
I		10th of 73	1	Keyed 2 instead of 12 (years of education)

any final keying errors and no one question produced more than one keying error. Errors were made by both novice and experienced interviewers. For the few interviews with multiple keying errors, the errors tended to cluster together on adjacent questions. Many of the errors (9/16) were simple mispunches one number away from the respondent's answer. None of the final errors resulted in a changed skip pattern. The number of keying errors detected was too low to permit further analysis.

5. Discussion

The number of keying errors in closed-ended questions found in this study was negligible. CAPI surveys vary on a number of dimensions, including complexity (e.g., the variety of question types), the type of questions asked (e.g., fixed-response vs. open-ended vs. numeric entry such as dollar amounts), the software used and program quality (design constraints), the data entry method used (e.g., numeric vs. alpha vs. cursor keys), and the interviewers employed (e.g., typing skills and computer experience). It should be remembered that the range of options available in closed-ended questions is limited, with only a small number of numeric keys being used. In addition, out-of-range entries are not accepted by the instrument. Whether the results from this study are generalizable to other CAPI surveys or other types of questions is not known. We also do not know the relative recording error rates for comparable PAPI (paper and pencil) surveys. Finally, the number of initial entry errors that were detected (either by the interviewer or by checks embedded in the instrument) and corrected during the interview was not recorded. The focus of this study was on inaccuracies in the final data set attributable to interviewer keying errors rather than inefficiencies resulting from corrections of miskeyed entries during the interview. Nonetheless, the low keying error rate found for this CAPI survey is encouraging. The miskeying of answers to closed-ended questions with limited sets of allowable entries does not appear to be a major cause for concern on carefully designed CAPI surveys.

6. References

Baker, R.P. (1992). New Technology in Survey Research: Computer-Assisted Personal Interviewing (CAPI). Social Science Computer Review, 10, 145–157.

Bernard, C. (1989). Survey Data Collection Using Laptop Computers. Paris: INSEE (Report No. 01/C520).

Cannell, C.F., Lawson, S.A., and Hausser, D.L. (1975). A Technique for Evaluating Interviewer Performance. Ann Arbor MI : Institute for Social Research.

Cannell, C.F. and Oksenberg, L. (1988). Observation of Behavior in Telephone Interviews. Chapter 30 in Groves, R.M., Biemer, P.P., Lyberg, L.E., Massey, J.T., Nicholls, W.L., and Waksberg, J. (eds.), Telephone Survey Methodology. New York: John Wiley.

Catlin, G. and Ingram, S. (1988). The Effects of CATI on Costs and Data Quality: A Comparison of CATI and Paper Methods in Centralized Interviewing. Chapter 27 in Groves, R.M., Biemer, P.P., Lyberg, L.E., Massey, J.T., Nicholls, W.L., and Waksberg, J. (eds.), Telephone Survey Methodology. New York: John Wiley.

Groves, R.M. and Nicholls, W.L. II, (1986). The Status of Computer-Assisted Telephone Interviewing: Part II – Data Quality Issues. Journal of Official Statistics, 2, 117–134.

Herzog, A.R. (1994). The Self in All Its Glory. Ann Arbor MI: Institute for Social Research (unpublished manuscript).

Kennedy, J.M., Lengacher, J.E., and Demerath, L. (1990). Interviewer Entry Error in CATI Interviews. Paper presented at the International Conference on Measurement Errors in Surveys, November, Tucson, AZ.

Martin, J., O'Muircheartaigh, C., and Curtice, J. (1993). The Use of CAPI for Attitude Surveys: An Experimental Comparison with Traditional Methods. Journal of Official Statistics, 9, 641–662.

Rustemeyer, A. (1977). Measuring Interviewer Performance in Mock Interviews. Proceedings of the Social Statistics Section, American Statistical Association, 341–346.

Weeks, M.F. (1992). Computer-Assisted Survey Information Collection: A Review of CASIC Methods and Their Implications for Survey Operations. Journal of Official Statistics, 8, 445–465.

Received April 1994
Revised November 1994