Chapter 7

CATI-CAPI Technical

Chair: Fred Wensing, Australian Bureau of Statistics

Kevin Dooley

Timothy Triplett ◆ Beth Webb

Stanley E. Legum

Chapter 7

Questionnaire Programming Language (QPL)

Kevin Dooley, U.S. General Accounting Office

Abstract

The Questionnaire Programming Language (QPL) consists of a set of IBM/PC programs that automate many of the activities involved in gathering and preparing survey data for analysis. Using this software, complex computer-aided telephone interview (CATI) or data-entry programs can be written that are easy to use and provide a high degree of control over what information may be entered. Interviewers can be trained to use the CATI software in only minutes, and completed interview records can be edited quickly and accurately. Once a questionnaire program has been created, other QPL system programs can be used to automatically generate formatted questionnaire documents, Awk, QBasic, SPSS or SAS analysis programs; Lotus, dBase, or comma or tab delimited data files; or askSam text-based data files. ■

Chapter 7

Using a Parallel "CASES" Instrument to Edit Call Record Information and Remove Incorrect Interview Data

Timothy Triplett and Beth Webb
University of Maryland at College Park

Abstract

The Survey Research Center's "fix-it" program is a CATI (Computer-Assisted Telephone Interviewing) instrument that reads the same data files that are read by the main CATI questionnaire instrument. The Survey Research Center uses the Berkeley "CASES" software, which allows the CATI instrument writer to design multiple instruments that read the same data files. The term parallel instrument is the term used in the CASES documentation. The fix-it program works by correcting the following three types of errors. First, the fix-it program can change the recorded status of any call attempt or the interviewer ID number associated with any call attempt. For example, often interviewers will record an incorrect call disposition on one of their call attempts. With the fix-it program the person with access rights to the program can change any call attempt's call disposition. They can also change the interviewer ID number if the incorrect ID number is recorded. Interviewers sometimes use call back codes when they should be recording the call result as a refusal and also sometimes incorrectly code eligible households as ineligible.

In addition, sometimes interviewers simply make a data entry mistake. The fix-it program easily corrects these problems without having to manually edit the data file and fix-it also updates' the current status of the case by re-evaluating the entire history of call attempts.

Second, fix-it easily removes the information from the last call attempt. This is necessary when interviewers record call record information in the wrong case. After removing the last call attempt, the fix-it program re-evaluates the

Abstract (Cont'd)

case history to determine the current status of a case. While the removal of the last call attempt occurs less often than changing call attempt status, this feature is often combined with the final feature of fix-it, the removal of data. When an incorrect respondent is interviewed not only does the data need to be removed, but the last call attempt disposition must also be removed.

Third the fix-it program allows the supervisor to remove invalid interview data without having to edit the data file. Using the fix-it program to remove data prevents accidental deletion of both valid call record and respondent selection data. In addition, using the fix-it program ensures that all the invalid data is removed from the case.

There are a number advantages of using the fix-it program to edit sampling information CATI data files. Perhaps the most important is that it is easy and thus can get done in a timely fashion. Though just as important is that the fix-it program require no manual updating of a data file, thus safely updates and edits only those fields where a change is requested. Other advantages include the automatic re-evaluation and update of a case's call status. This automatic update is important for both keeping sample reports accurate and helping autoschedule programs accurately choose a sample to call.

Two other advantages of the fix-it program are first, it is easy to undo any changes, since the CASES software records key information in a history file. Second, the fix-it program is a generic program custom designed to work with the Survey Research Center's front end. The front end is the part of the CATI questionnaire instrument that stays the same from study to study. Thus, fix-it easily works for most SRC projects.

Using a Parallel "CASES" Instrument to Edit Call Record Information and Remove Incorrect Interview Data

Timothy Triplett and Beth Webb
University of Maryland at College Park

Introduction

The Survey Research Center's "fix-it" program was written using the Berkeley CASES (Computer-Assisted Survey Execution System) software. The fix-it instrument is used to edit data for CATI (Computer-Assisted Telephone Interviewing) studies. The three main functions of the fix-it program are editing incorrect sample disposition information, editing incorrect interviewer information, and removing survey data from interviews conducted with ineligible respondents.

The CASES software allows the instrument writer to design multiple instruments that read the same data files. These instruments are referred to as "parallel instruments" in the CASES documentation. The main questionnaire instrument is written and executed in the e-inst directory of a study. The fix-it instrument is usually written and executed in the e-inst2 directory. If there are multiple parallel versions of an instrument (e.g., different language versions), the fix-it instrument can be located in a different e-inst directory such as e-inst3 or e-inst4).

One advantage of writing the editing features into a parallel directory rather than in the main instrument, is that there is more control as to who is allowed to edit sample dispositions. Access can be limited, through the use of a password in the instrument, to those who are given responsibility for sampling issues. This reduces the likelihood of improper usage. When the fix-it program is executed the first screen asks for the user's identification and the second screen asks for that person's password. If the person does not have access rights to the fix-it instrument or an incorrect password is entered, execution is terminated.

Data-Editing Features

The first screen after the password screen offers the following:

- Option #1.--Change the status of a call attempt
- Option #2.--Remove the last call attempt
- Option #3.--Change the interviewer ID# for a particular call attempt
- Option #4.--Exit program, no change made.

Changing Disposition Codes

The first option allows the fixit program to change the disposition code for any call attempt. After a call attempt is completed the interviewer records the outcome status of that attempt using one of the outcome codes listed in Table 1. Supervisors review the interviewer's recording of call attempt disposition codes and sometimes feel the need to change the disposition code chosen by the interviewer. A common example is interviewers trying to hide refusals under the code of call back. Correctly coding a refusal call attempt is extremely important, since the call disposition code is the most important factor in determining how future call attempts are handled. In addition, one of the criteria of interviewer evaluations is their ratio of completed interviews to refusals (see Table 2); thus, it is important that the supervisor has the ability to assign the initial refusal to the correct interviewer.

Table 1.--Call Status Report

Code	Call Status	N	Percent
1	Completed Interviews	1,000	39.9
2	Partially completed interviews (Terms)	30	1.2
3	**Refusals**	**74**	**3.0**
4	**Call Backs (appointments)**	**65**	**2.6**
5	Home Recorders	10	.4
6	No Answers	21	.8
20	Language/Age/Health Problems	84	3.4
21	Finalized Partial completes	3	.1
22	2nd Refusal	174	6.9
23	Call Backs (finalized after 25 attempts)	35	1.4
24	Household no longer available	22	.9
25	Home Recorders (finalized after 25 attempts)	22	.9
26	Refusal (finalized after 25 attempts)	11	.4
30	No answer (after 20 attempts)	141	5.6
31	Non Working phone number	516	20.6
32	Non Household	297	11.9
98	**Possible ineligible respondents**	**1**	**.0**
		2,506	100.0

There are many other less common reasons why supervisors decide to change the call disposition code assigned by an interviewer. Some of these occur because of an interviewer keying error. Other sources of mistakes stem from interviewers misunderstanding how to properly code a call disposition. For example, an interviewer may code a call attempt as a respondent problem because the respondent is away for a week, whereas the proper code would have been to code the attempt as a call back for next week. In any event all mistakes by the interviewer in coding call dispositions need to be corrected so that future call attempts are handled correctly. In addition, corrections to the data file are needed so that the sample progress reports, interviewer progress reports and the final sample status report are accurate.

Table 2.--Cooperation Rate by Interviewer			
ID#	Completes	Refusals	Cooperation Rate Completes / Completes + Initial Refusals
1,009	33	8	80.5%
1,120	46	10	82.1%
1,121	27	10	73.0%
1,156	86	12	87.8%
1,160	45	22	67.2%
Column total	237	62	79.3%

The fix-it program not only allows easy editing of call disposition codes, it also updates all the sample status variables and call attempt counters affected by the change.
To change the call disposition code:

- ❒ Select Option #1, "change status of a call attempt."
- ❒ Enter the call attempt that is in error.
  (You will be told total number of attempts.)
- ❒ Enter new call attempt disposition.
  (The old disposition is shown at the top of the screen,
  and a list of possible disposition codes is displayed.)
- ❒ Verify the change.
  (The caseid #, the call attempt #, the old disposition,
  and the new disposition will appear in a confirmation screen).

There are safeguards built into this editing feature. For example, the status of a case can not be changed to a complete or partial complete when there are no interview data in the case. When there are data in the case and the status of the case is being changed to something other than a complete or a partial, a screen appears with the warning that the data will be deleted. The change must be confirmed at this point, in order for it to be made.

Removing Last Call Attempts

The second option, "remove the last call attempt", is a special case of editing an incorrect call attempt. This option is used when an interviewer either misdials a number or the interviewer dials correctly but then records the information in the wrong case-id. Often the interviewer will realize the error and report it immediately to the supervisor; other times it is caught when a supervisor attempts to verify the status of the case.

When the last call attempt is removed, the disposition of the case is changed to what it was on the call attempt previous to the last call attempt. To make this change:

- ❒ Select Option #2, "remove the last call attempt."

- ❒ Verify the change.
  (The caseid #, the call attempt #, and the disposition of the call attempt being removed will be displayed in a confirmation screen.)

Editing Interviewer IDs

The third option allows the editing of the interviewer identification number associated with any call attempt. The interviewer must record his or her identification number for each call attempt. Occasionally the identification number is miskeyed. Either the interviewer realizes the error immediately and reports it to the supervisor, or the error is discovered when the "Cooperation Rate by Interviewer" table (see Table 2) is reviewed.

The following steps are taken to change the interviewer ID number:

- ❒ Choose option #3, "Change the Interviewer ID# for a particular call attempt"

- ❒ Enter the call attempt which is in error
  (The total number of attempts will be displayed)

- ❒ Enter the correct interviewer ID number.
  (The current interviewer ID # will be displayed.)

- ❒ Verify the change
  (The caseid #, the current interviewer ID number, and the correct Interviewer ID number will appear in a confirmation screen.)

There is a disposition code for complete or partial interviews with possible ineligible respondents(see Table 1). This code is used when during or after an interview, the interviewer is given information that leads him or her to believe that the respondent is not eligible for the study (e.g,. the population is households and we've reached a respondent in a nursing home; or adults are being interviewed and the demographics section of the questionnaire indicates that the respondent is 17). The interviewer then codes the case as an interview (or partial interview) with a possible ineligible respondent. A case is automatically assigned this code if the respondent does not verify the phone number called at the end of the interview.

When the fix-it program is executed for a case with this code, the following options appear:

- ❒ ***No, Don't Clear Data -- But Return This Case to the Status It was Before It Became a Problem.***

- ❒ ***Exit the Program, No Change Made.***

- ❒ ***Yes, Clear the Data.***

If it is decided that the interviewer was incorrect and the respondent **is** eligible to be interviewed for the particular study, Option #1 is chosen and the case becomes a completed interview.

If the interviewer was correct and the respondent is **not** eligible to be interviewed for the study, Option #3 is chosen. This clears the data and returns the case to the disposition it had on the previous call attempt. Before the data are cleared, a confirmation screen appears displaying the caseid number and giving the same options as above.

Summary

The old method of editing sample disposition information data involved manually editing the physical data files, which was a very time-consuming and error-prone process. This was especially true when an early call attempt had to be changed, necessitating the editing of the data for each subsequent call attempt. The fix-it program always writes changes to the correct variables and all status and counter variables are automatically updated. This greatly reduces the chance of keying error and the amount of time needed to fix sample problems. When the status and counter variables are accurately and frequently updated, this improves the accuracy of programs which use these variables such as sample reports and auto-schedulers.

Another important feature is the portability of the program from study to study. The shell of the fix-it program can be written in as little as a week by someone who is familiar with the front-end. As long as the front end stores the sampling variables in the same location, it takes less than 10 minutes to update and prepare the fix-it program to be used for individual studies. ■

Chapter 7

A Computer-Assisted Coding and Editing System for Non-Numeric Educational Transcript Data

Stanley E. Legum, Westat, Inc.

Abstract

Westat has completed a number of large studies in which the basic data have come from thousands of high school transcripts produced by hundreds of schools. These data needed to be coded and combined into a common database. Some of the challenges presented by these studies have been:

- Transcripts from different schools have different formats and present different information;
- Courses with the same titles in different schools often have different content;
- Courses with similar content may have different names in different schools;
- Within a school, remedial and honors courses may be distinguished from regular courses of the same name by codes on the transcript;
- The number of student contact hours represented by one credit differs from school to school;
- Schools use different grading systems.

We will demonstrate two software tools developed for the transcript studies:

- The Computer Assisted Data Entry (CADE) system; and
- The Computer Assisted Coding and Editing (CACE) system.

Abstract (Cont'd)

The Computer-Assisted Data Entry System

The Computer Assisted Data Entry system was designed for use by clerks who enter data directly from the transcript onto forms appearing on their computer screens. The CADE system has the advantage of naturally guiding the clerk to look for the needed information wherever it may appear on the transcript and of providing a consistent means for entering it. The CADE system, which is written in Clipper, includes range and logic checks which are active during data entry. It also includes a provision for double keying by a second clerk. When the second keyer enters information different from the first keyer, the system gives the second keyer a message and provides an opportunity to change the entry or confirm that the second entry is correct.

The Computer-Assisted Coding and Editing System

The Computer-Assisted Coding and Editing system is designed for use by subject matter experts (in this case, curriculum specialists) in the process of mapping objects (high school courses) to a predefined classification system (in this case, the Classification of Secondary School Courses). Since subject matter experts are selected for their knowledge rather that their data entry skills, the CACE system minimizes the amount of keying that needs to be performed.

Clerks pre-key all the titles from a high school catalog into a file which is read by the CACE system. The system, which is written in Paradox, presents the coder with one course title at a time from a school, a suggestion list of classification codes which might match the course, and a window displaying the full text of the classification manual that applies to the currently highlighted suggestion. Coders are free to ignore the suggestions and browse the entire classification manual and select any applicable code. A system of "flags" lets the coder record specialized information about a course such as it being the first course in a sequence or a later course in a sequence or that it is taught off campus. Only valid codes and flags are accepted by the system.

A subsystem of the CACE system is used by coders to match course titles in course catalogs to the corresponding course titles on the transcripts. This process cannot be fully automated because of the sometimes idiosyncratic ways in which course titles are abbreviated on transcripts.

Summary

In developing solutions to the challenges of entering and coding complex educational transcript data, Westat has developed generalizable software that may be applicable to other content domains such as medical record abstracting or document cataloging. ■