THREE ASPECTS OF CLASIC

John S. Gardenier, D.B.A., Centers for Disease Control and Prevention, National Center for Health Statistics (NCHS), Room 1120, 6525 Belcrest Rd, Hyattsville, MD 20782

KEY WORDS: CLASIC, CASIC, survey software

THE CLASIC SOFTWARE CONCEPT

Size and Complexity in NCHS Health Surveys

Health reform initiatives demand better and faster health data collection and reporting than can be met by traditional paper and pencil survey data collection methods. Computer- Assisted Survey Information Collection (CASIC) promises improvements in quality, timeliness, and flexibility.[1] Anticipating such benefits, NCHS began investigations in 1987 to automate its most complex personal interview surveys.[2] The results were encouraging, but CAPI then fell short of the full capability required.

We learned that the following requirements were exceptionally difficult to meet with existing software architectures:

Ability to adjust interview size dynamically during the interview itself. National Health Interview Survey (NHIS) interviews run from 1½ to 3½ hours, collecting extensive health data on up to thirty persons.

Ability for the interviewer to control the interview sequence, based on each respondent's availability.

Ability for the interviewer to track what/whose survey parts are: not required, required, or completed, dynamically throughout the interview.

Ability to back up, change or add responses or people, and go forward on the appropriate paths dictated by the new response set - without losing previous still-valid answers.

Ability to re-start an interrupted interview at the point of interrupt, even if a child or pet accidently dislodges the computer power plug.

In 1990, we decided the best way to meet these requirements was to develop survey software from scratch. Accordingly, we prepared full software requirements specifications and the preliminary design.[3] We adapted the CASIC acronym to focus on small computers, hence, the Computer/Laptop Assisted Survey Information Collection (CLASIC) system. The CLASIC system has three components: the Interview Software, the File Translation Postprocessor, and the Training System.

The Interview Software

CLASIC interview software is insensitive both to interview size and complexity. It is (largely) generic; one set of computer programs can run any survey. Survey-specific text AND LOGIC are placed in "survey definition" files.

The interview itself is run by a program we call the CLASIC "interview engine." That engine accesses the survey definition files as required to complete an interview. The interviewer can, at any point, interrupt the automatic flow of the interview for purposes of editing or redirecting the interview flow to meet changing conditions in the household, such as one respondent leaving or another one becoming available.

The most unusual feature of CLASIC - and its reason for being - is its navigational capability. Like all survey software systems, CLASIC supports automatic control of fills and skip patterns, warning and error messages for out-of-range or inconsistent data, and straightforward interviewing. Like some

other systems, CLASIC offers multiple-level rostering and the ability to back up, change answers, and continue with the appropriate subsequent questions.

Unlike most other survey software, CLASIC also:

> Allows direct editing of rosters, including capabilities to add, change, delete, or undelete rostered elements (including families and individuals within a household).

> Allows designation of individuals or families as refusals, skipping them while providing full coverage of other eligible subjects.

> Allows the interviewer to override the automatic navigation system when required by respondents' availability.

> Keeps track of survey component requirements and completions by person and by any other rostered element.

> Runs a "stopwatch" on the whole survey and on each component of a complex survey, accumulating timing data regardless of how often a section or supplement may be revisited.

CLASIC uses open-ended data structures during the interviews themselves. We never know how many persons may be involved – with how many health conditions, medicines, insurance plans, or other topics. Therefore, the interview engine dynamically re-sizes and revises the survey as it progresses.

The File Translation Postprocessor

CLASIC's open-ended data and file structures cannot be read directly by mainframe computers. To accommodate post-interview processing, editing, and analysis, we must translate the data into flat files. That requires

a variant of the "interview engine" - the file translation postprocessor.

NCHS does not collect its own data; rather, the Bureau of the Census and various contract survey organizations do the actual field work. That means CLASIC does not have to provide case management, but it does have to interface with any existing or future case management system.

Interfacing with external case management systems requires that we be able to interface software and files without violating memory or disk space constraints.

It also means we must be able to select any subset of the data gathered in a survey and reproduce it in any format needed by the data collectors. That requirement is in addition to the need to produce mainframe-readable files of the total data product.

CLASIC meets both the survey and the case management data needs by:

> 1. Incorporating each formal output specification.

> 2. Preparing a "script" which links each required output to the data source in the working files. This script also embeds any needed simple recoding or reformatting instructions.

> 3. Preparing logical rules for conditional data translation, for creating new data values from two or more "raw" data values in the working files, or other processing which is too complex for just a "script."

> 4. Re-running the survey basically as it was executed in the household. This allows the postprocessor to pick up all still-valid original data even from survey segments which were later revised. It also results in overwriting any incorrect earlier data with the later corrections.

174

This separation of output formatting from the interview offers several advantages. The postprocessor recognizes that some collected data may become irrelevant due to later entries. It reads indications of subsidiary records (such as health conditions, doctor visits, or hospital stays) having been "marked for deletion" and ignores them.

That allows such records to be deleted and un-deleted as necessary in the interview itself. No data is actually "lost"; we simply ignore data still marked for deletion at the end of the interview.

Even the data which the postprocessor ignores is never lost. It remains in the archived working files, available for re-postprocessing with alternative parameters if later required by analysts. For example, all answers bear date and time stamps, but only aggregate elapsed time by section is reported initially.

During survey design, this system facilitates testing and revision by deferring detailed description of output formats until after the final form of the instrument has been decided. This feature supports quick turnaround in questionnaire research and development laboratories (or "cognitive labs".)

The CLASIC Training System (TS)

Another variant of the interview engine is not so much "required" as "opportunistic." The interviewers who have conducted major health surveys on paper are not necessarily computer literate. Even those who are intermediate or more advanced computer users in general have to learn CLASIC. It just so happens that minor adjustments to the interview engine turn it into a useful training system for computer-based instruction (CBI). The great advantage of using a form of the interview system to create its own CBI is the similarity of the resulting screens and key functions.

The CLASIC-TS system is designed to be able to provide training in the interview mechanics

at least, with minimal demands on the survey designers. It is a separate executable program, designed to be "embedded" (built into)[4] the actual interview program. This will allow survey designers to provide training material on any of their specific surveys with minimal adaptation.

Initially, NCHS developed the training system as a separate (not embedded) version of CLASIC. This was necessary to allow the interviewing system to progress while the training system was "frozen" for the incorporation of courseware oriented specifically to the NHIS.

During the next software revision, the training system will be merged with the interviewing system. At that point, it will be feasible to expand the training to deal with the concepts and definitions of each individual survey.

That will, however, require the survey designers (or their training consultants) to create specific training materials above and beyond what they would normally provide for on-screen instructions and context-sensitive help. They would have to identify specific questions to illustrate training points and provide the training text for those. They would also have to provide practice exercise (quiz) items and messages of positive or negative feedback on trainee actions.

Each training screen is based on a screen in the CLASIC survey software itself. A "header line" identifies the current survey topic and the current state of the computer: "READY" (for input) or busy, causing a "WAIT." The training screens add specific data about the training itself in the header line.

The training material is contained in a "TRAINING WINDOW" which overlays about one-third of the actual survey software screen - either at the top or the bottom. In cases where the training window necessarily obscures important material in the underlying screen, the trainee can reposition the training window (to the top or bottom) with a toggle

command. In cases where the training information or exercises take up too much space, material can be scrolled within the training window.

The system allows the interviewers to select the topics they most want to study. If they are not sure whether they need the full tutorial on a topic, they can page through the "quick review" to get a feel for the key points covered. Either before or after taking a lesson tutorial, they can take a quiz on the material covered - usually about 10 multiple choice questions.

Finally, for the NHIS, the interviewers also have available scripted practice interviews prepared by the Bureau of the Census for hands-on experience with the actual survey software - apart from the embedded training.

Acknowledgement

References

1. Subcommittee on Computer Assisted Survey Information Collection of the Federal Committee on Statistical Methodology, "Computer-Assisted Survey Information Collection," Working Paper 19. Washington, DC: Statistical Policy Office, Office of Management and Budget, May, 1990.

2. National Center for Health Statistics and Bureau of the Census, REPORT OF THE 1987 AUTOMATED NATIONAL HEALTH INTERVIEW SURVEY FEASIBILITY STUDY: An Investigation of Computer Assisted Personal Interviewing. NCHS Working Paper Series, Number 32. Hyattsville, MD: NCHS; November, 1988.

3. Gardenier, J. S., Biggar, R. S., Adams, C., and Shumate, P., SOFTWARE REQUIREMENTS SPECIFICATIONS AND DESIGN OVERVIEW: COMPUTER/LAPTOP ASSISTED SURVEY INFORMATION COLLECTION (CLASIC) INTERVIEWERS' SUBSYSTEM, Working Paper 36, Hyattsville, MD: National Center for Health Statistics, January, 1991.

4. Leslie J. Briggs, "Introduction," Chapter I of L. J. Briggs (Ed.) Instructional Design: Principles and Applications. Englewood Cliffs, NJ: Educational Technology Publications, 1977.