

# **Overview of Blaise at Iowa State University Statistical Laboratory**

**L. L. Anderson, Iowa State University Statistical Laboratory**

**J. M. Larson, Iowa State University Statistical Laboratory**

**D. G. Anderson, Iowa State University Statistical Laboratory**

**S. M. Nusser, Iowa State University Statistical Laboratory**

## **Introduction**

In 1998, Iowa State University Statistical Laboratory became one of the first university users of Blaise in North America. The Statistical Laboratory (Stat Lab) began collecting data using Computer Assisted Telephone Interviewing (CATI) in 1995, but was interested in finding a different product, one that could meet a small shop's need for simplicity while providing the ability to conduct large and complex studies.

After previewing a number of software options, the Stat Lab selected Blaise for several reasons: its ability to handle large complex data sets; its ease of use for both interviewers and supervisors; its strong data editing and case management capabilities; it would soon be in a Windows version; and the prospect of strong user support from Westat.

In this paper, we outline our initial experiences with Blaise and identify features that we hope to explore in the future.

## **Description of Statistical Laboratory**

The Stat Lab's survey research data collection unit is small (20 CATI workstations and five professional staff, including one programmer). CATI projects typically require shifts of one to eight interviewers. Interviews are conducted primarily during evening hours, with a lesser amount of day and weekend contacts made as required by the project. Smaller projects can include some day and weekend shifts that are not staffed at all, while larger international projects might require virtually round-the-clock interviewing.

The nature and scope of projects varies tremendously. The Stat Lab has conducted numerous survey research projects with complex respondent selection criteria, extensive rostering, and intricate questionnaire routing with sub-interviews. Clients include university administration, faculty with research funded by external organizations, other universities, government agencies, and other non-profit organizations. Research topics cover issues such as welfare reform, business development, agriculture, economics, education, and health. Mail and web surveys, personal interviewing, and focus groups are all conducted by the Stat Lab.

## **Case management in a small shop**

We first used Blaise on relatively simple projects, a marketing study for the university student newspaper and an administrative study of university retirees. The questionnaires were straightforward and the target populations (students and retirees) were easily contacted list samples. Although the questionnaire itself performed well, our staff received the first indications that adaptations were needed for efficient case management. Specifically, we needed more extensive contact and background information available to

interviewers, and we needed to readjust the parameters in the CATI Specification Program since various cases sometimes appeared to interviewers at inappropriate times.

Our highest priority subsequently was to make a complete call history available to interviewers. Each dial result is now written to the database. These results, along with appointment and other information, are displayed on an infopane before the start of the interview (Figure 1). The call history screen is adapted for individual project needs. It provides interviewers with sufficient information about the case to increase their comfort level and to facilitate wise decisions.

To ensure that interviewers view the history screen before dialing, only the Questionnaire option is available on the Make Dial screen. All dial results are available within the questionnaire as parallel blocks in the tab format. After the introduction screen is passed, the dial results are no longer available.

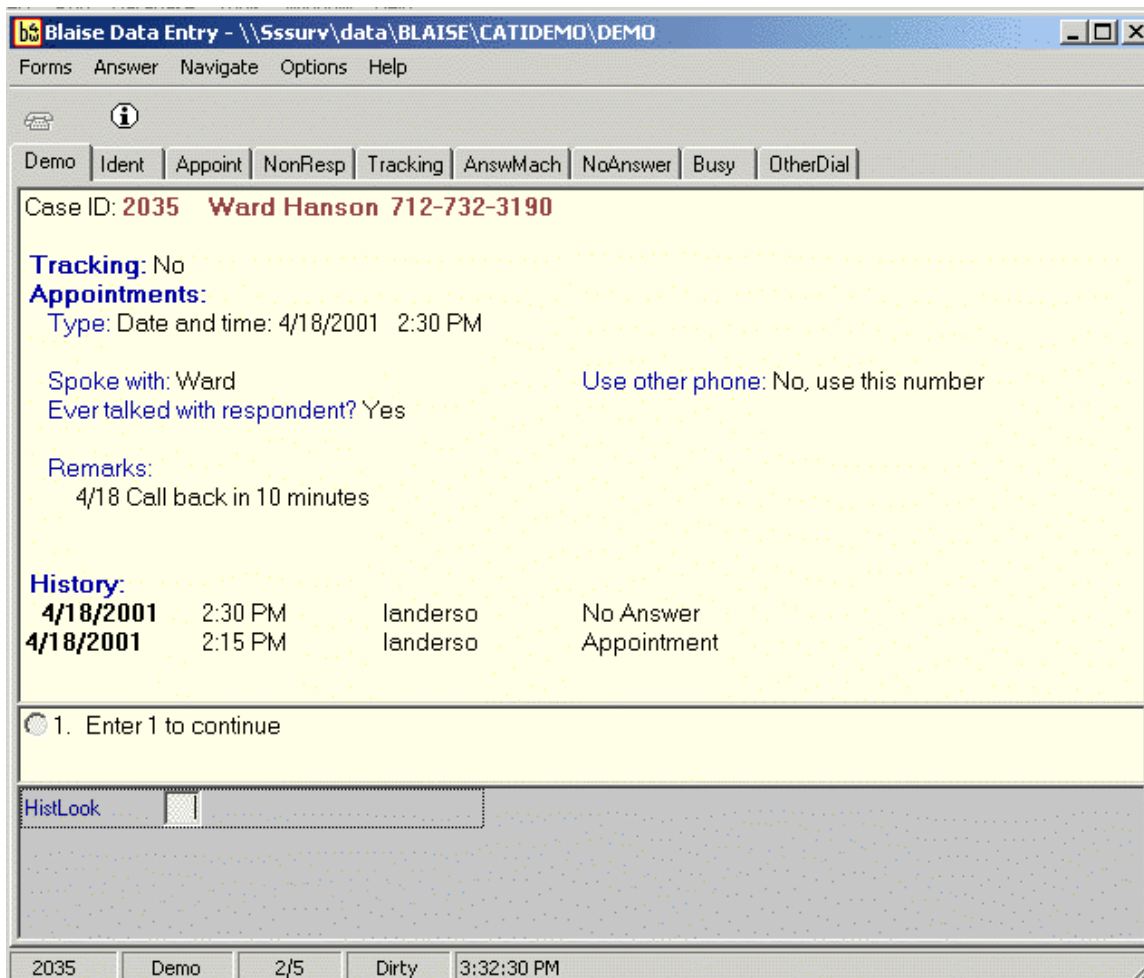


Figure 1. Call history screen.

Our difficulties in controlling the autoscheduler increased with a business development study, which indicated again that we needed to learn more about how to set parameters, how to handle various types of callbacks, and how to add another level of control when we have prior knowledge of when to call. For

example, most types of businesses should be called during daytime business hours, but a restaurant should not be called near lunchtime.

Callbacks on this study became a problem when we did not talk to the respondent (the manager of the business), but to an employee gatekeeper who would inform us that “He’s out of the office this week”, or “You can probably get her late in the afternoon”. We ended up with many soft priority cases that repeatedly came up before untried cases. Also, if the soft priority cases did not get a response within the period set, they would not come up again for a long time. We wanted a more even distribution between these cases, for which we had no real appointment information, and cases that had not been tried at all. Additionally, hard appointments that were not met sometimes surfaced inappropriately the following morning.

Increasing the number of days between no answer calls and reducing the maximum number of dials helped bring up more untried cases. We used interviewer groups to assign one interviewer only fresh numbers. We also trained the interviewers to make appointments more specific or more general, as appropriate. We need to look into methods to specify times to call for specific forms, before calling. Assigning different time slice sets to different types of businesses might help us call at appropriate times, but we need to make sure we always have someone working during all time slices. Our initial attempt to use time slices was unsuccessful because we did not always have someone scheduled to work during each time slice.

Our first welfare reform study used a list sample of welfare recipients. Because it was expected that a great deal of the contact information from the frame would be incomplete or incorrect and extensive tracking would be required, the autoscheduler was not used. All locating was done by telephone, using references such as National Change of Address (NCOA) and cross directories. Because supervisors did most of the tracking, and because most of the information they received from calling was vague and not easily classified, it was easier to have the information for each case on paper record-of-calls (ROCs).

In a later welfare reform study, we used the autoscheduler rather than the paper ROCs, making use of a five-row tracking table to store information from five tracking attempts (Figure 2). This information includes old telephone numbers, new telephone numbers, the source of the new telephone number, and an open field for any information available. If a new number is not known, any information that the interviewer has may be entered, and that information will be available in reports. If a new number is known, the number can be tried immediately. The new number replaces the old telephone number in the original telephone number field used in the Make Dial screen.

Blaise Data Entry - \\Sssurv\data\BLAISE\CATIDEMO\DEMO

Forms Answer Navigate Options Help

Demo Ident Appoint NonResp Tracking AnswMach NoAnswer Busy OtherDial

**What is the source of the new telephone number?**

1. Person at previous number  
 2. DHS  
 3. Internet  
 4. Directory assistance  
 5. Respondent called in  
 6. Respondent replied by postcard  
 7. Other

	TrackQ	OldPhone	NewPhone	Source	MoreInfo	CallNewPhone
Track[1]	2	319-386-2480	319-385-2211	1		
Track[2]						
Track[3]						
Track[4]						
Track[5]						

2145 Tracking 1/1 Dirty 2:27:08 PM

Figure 2. Tracking block

All cases in need of tracking were written to html reports, with each case linked to a page giving all information collected in the tracking table. These pages were printed and given to the two interviewers we had in the field to locate respondents. The field interviewers conducted the interview with pencil-and-paper if they were able to speak to the respondents, or they provided the office with a new telephone number when they were unable to conduct the interview in person.

In this study we used a dual frame sample, with a list of welfare recipients to target the welfare population and an RDD sample to target both the general population and low income respondents who were not on welfare. A different set of dispositions was required for the RDD and list portions of the study. We also made use of an external file to code car make and model in a roster of household vehicles.

Finally, our most complex Blaise study to date has been the Family Business Research Project. Interviews were conducted first in 1997. All 1997 respondents were contacted again in 2000, this time using Blaise. The study was conducted for a consortium of researchers from 16 universities in the U.S. and Canada. Data were collected both about the business and how the business affected the family. Accordingly, both the business owner/manager and the individual responsible for managing the household were interviewed. In some cases, this was the same person. In 2000, we were thus attempting to contact either one or two people.

Routing for the questionnaire depended on whether the business was still owned or operated by the business manager respondent, and whether the business manager and household manager were still in the same household. If the business was no longer in operation or managed by our respondent, a short series of questions was asked. If the business manager and household manager were no longer in the same household, the business manager would be asked all business and household questions (a combination questionnaire), and the household manager would have a short series of questions. If the household manager was now the business manager, the household manager would answer the combination questionnaire.

Because of the difficulty of making appointments for two respondents who may no longer live in the same household, we used paper ROCs rather than the autoscheduler. We needed to have one form for both respondents because the preliminary part of the questionnaire determined the routing for both respondents, and the next part of the questionnaire (the household roster) was answered by the first respondent interviewed.

### **Data Management**

The Blaise CATI Management reports and the edit mode are important to our case management. Because we may not have a crew working during the daytime on a small project, it is necessary for supervisors to be aware of upcoming appointments. The appointments graph and other reports in CATI Management make case management efficient by making it easy to keep track of appointments and the status of all cases.

The edit mode makes it possible to check the quality of the data and make changes immediately after an interview. Supervisors can access the production data model in edit mode with CATI disabled, using a shortcut icon.

For editing and coding, we use a new data model. This new data model includes additional codes for enumerated fields which have a value of “other” and are followed by an open field to specify the “other” answer. It also includes additional enumerated fields to code other open fields. The supervisors use a Manipula setup to read data into the edit data model. They can edit and code data while data collection is still in progress, thus shortening the data delivery time.

### **Use of Reports**

The reports in Cati Manager are supplemented with reports run out by a Manipula setup and processed by SAS in a batch file at night. Summary reports are written to a text file and printed out; reports dealing with individual cases with appointments (Figure 3), cases with dial results such as refusals, nonresponse, other dial, tracking, and needs reset, and cases which have been called at least eight times are written to html files which are viewed in a browser and printed as necessary. Information in the reports includes case ID, interviewer, date of call, comments, and other appropriate information.



Figure 3. HTML Current Appointments report

Finally, the output data set is processed to fill in blank fields in the data and conform to the standard Stat Lab format. The Cameleon SAS script has been modified to write SAS statements to fill in empty fields and change the coding for refusals. All “not applicable” results are filled in with 8s, “don’t know” with 9s, and “refuses” with 7s.

### Summary

In the future, we hope to refine our use of the case management features already available to us, such as appointment parameters; to explore options such as using select and sort fields and time slices; and to learn how to use Manipula to run updates in management fields.

There are some refinements we would like to see in the CATI Specification parameters, including the ability to set a different maximum number of dials for cases with default priority and for cases with appointments. Similarly, we would like to set a different number of minutes between no answer for cases with default priority and for cases with appointments. We would also like to set a different route back option for appointments and for other cases, such as refusal conversions.

Blaise has been very well received at the Stat Lab. Interviewers, some with minimal computer skills, find it very easy to work with. Supervisors are very satisfied with the editing and coding abilities and the possibility of conducting complex projects in Blaise. Data are processed quickly, allowing projects to be

finished quickly. It has been critical to have Blaise's efficient routing tools, which enabled us to create questionnaires with extremely complex skip patterns.

Overall, we have found Blaise to fulfill our requirements of simplicity while having the capability to do complex work.

