

First steps along the Audit Trail

Tim Burrell, Office for National Statistics

1. Introduction

In the last few years the Office for National Statistics (ONS) has started to examine the potential of the Blaise Audit Trail facility. An Audit Trail produced within Blaise records interactions made by the interviewer during computer assisted interviewing (CAI) or the respondent during computer assisted self-interviewing (CASI). The Audit Trail can provide a detailed history of the sequence and timing of navigation, field entries, changes to field entries and other events.

The paper will outline the practical application of Audit Trails on two projects:

- 1) to monitor respondent behaviour during CASI on the General Household Survey (GHS); and
- 2) to observe interviewer actions on a survey of income and living conditions (EU-SILC).

The paper will describe how we have been able to analyse the data produced, with reference to surveys mentioned above, and what progress we would like to make in the future. It will also consider some of the implications and concerns we have on using this tool.

With CAI, and Blaise in particular, the industry standard for the collection of official survey statistics, there is an increased availability of data for the researcher to investigate interviewer and respondent behaviour during an interview. With the help of a package provided by Statistics Canada named ATLAS, we have been able to enhance our use of the Blaise Audit Trail facility and widen the extent to which we have implemented it on surveys. This new software has also led to much easier analysis of the extensive data produced.

2. Previous use of the Audit Trail within ONS

ONS has previously used the Blaise Audit Trail facility on a small-scale project to evaluate the utility of this monitoring method. We chose an ONS survey of children and adolescents in the care of local government authorities, known as the survey of Looked After Children (LAC) as the vehicle for this trial. The Blaise instrument for the survey was designed to collect information about the mental health of children and included a substantial CASI section relating to sensitive or illegal behaviour. The main interest of the trial was to monitor respondent behaviour in the Audio Computer Assisted Self-Interview (A-CASI) section of the interview. (Bumpstead 2001)

This study confirmed in ONS earlier findings (Bumpstead 2001, Hansen and Marvin 2001) that the Blaise Audit Trail can provide useful data about respondent use of CAI instruments which it would not otherwise be possible to capture. The processing and analysis of audit trail data is not generally straightforward. However, it was possible to formulate a strategy for dealing with audit trail data which was not unduly cumbersome or time consuming.

3. ATLAS - A tool provided by Statistics Canada

ATLAS has been used to aid in the analysis of the vast amounts of data produced within an Audit Trail. It has the ability to read and average out across cases the times of individual questions or sets of questions from the Blaise Audit Trail. This proved useful as we could separate out the time spent in particular parts of the interview, such as the interview section of the questionnaire and in the administrative section. Measurement of elements of the interview was the main point of the projects mentioned in this paper.

Before acquiring the ATLAS package from Statistics Canada, ONS had only been able to carry out limited analyses of Audit Trail data. We had looked at information case by case. Data was produced like that shown in Figure 1, which shows the details from one case. This type of data is hard to read and Audit Trail data is hard to analyse in its raw form. Figure 1 shows the interviewer's progress through 4 questions. In this particular case, the interviewer took 11 seconds to pass through this set of questions.

Figures 2 and 3 show how the information is presented using the ATLAS. Figure 2 shows information from the same case as Figure 1. Figure 3 shows the average time over a number of cases (153 cases from the UK's General Household Survey, GHS). It shows that, on average, interviewers took 7 seconds to pass through these questions.

Figure 1. Audit Trail raw data

```
"19/06/2002 18:16:31","Enter Field:QSignIn.StartDat","Status:Normal","Value:"
"19/06/2002 18:16:37","Leave Field:QSignIn.StartDat","Cause:Next
Field","Status:Normal","Value:20020619"
"19/06/2002 18:16:37","Enter Field:QSignIn.DateChk","Status:Normal","Value:"
"19/06/2002 18:16:40","Leave Field:QSignIn.DateChk","Cause:Next Field","Status:Normal","Value:1"
"19/06/2002 18:16:40","Enter Field:QSignIn.IntEdit","Status:Normal","Value:"
"19/06/2002 18:16:41","Leave Field:QSignIn.IntEdit","Cause:Next Field","Status:Normal","Value:1"
"19/06/2002 18:16:41","Enter Field:QSignIn.WhoHere","Status:Normal","Value:"
"19/06/2002 18:16:42","Leave Field:QSignIn.WhoHere","Cause:Next Field","Status:Normal","Value:1"
```

Figure 2. Analysing the Audit Trail using ATLAS for one case

BLOCK	INITIAL		SUBSEQUENT		TOTAL	
	HITS	SECS	HITS	SECS	HITS	SECS
QSIGN	1	11	0	0	1	11

Figure 3. Analysing the Audit Trail using ATLAS for a number of cases

TIAL BLOCK	SUBSEQUENT		TOTAL HITS	TOTAL SECS	AVG	
	HITS	SECS			HITS	SECS
QSIGN	4	7	0	0	4	7

It can be seen from the screen shots above that ATLAS summarises data in an aggregated form which is much more convenient as well as being in a much more readable form . This has enhanced our ability to analyse data produced by the Audit Trail. The data can be shown at the case level (figure 2) or the average over a number of cases (figure 3).

The data produced from ATLAS can be easily transformed into SPSS format where data can be analysed further. Other information available in the tool was the ability to single out particularly long hits, or unusually long amounts of time spent on questions. This feature means it is possible to single out any questions where a very long time was taken for an answer to be filled out. This could be due to a break in an interview for a number of reasons. It is then possible to remove such an outlier from the analysis if it would bias the measurement of the overall time spent on that question.

Data can be analysed both at an individual level, looking at the data case by case, or by looking at averages over all cases. Additional features include the ability to look at the number of fields which were answered *don't know* and *refusal*, and changes or edits to questions.

4. A pilot to monitor respondent behaviour during CASI on the General Household Survey (GHS)

The GHS is a multi-purpose survey providing the government with annual information about the major social fields of Population, Housing, Employment, Education, Health and Income to supplement the more specialised best national sources for estimates on these topics, such as the Labour Force Survey. Because all these topics are covered in one survey, it is possible to examine the relationships between them.

The General Household Survey interview comprises two parts: the household and individual questionnaires. The household questionnaire contains questions on demographic characteristics of household members, tenure and accommodation, consumer goods, migration and ethnicity. Questions relating to income are asked in the individual questionnaire on the GHS and the results aggregated for the household. ATLAS has been helpful in measuring the amounts of time spent in these different sections of the questionnaire.

The GHS carries topics which are asked to all members of the household. Some of these topics, such as smoking, drinking and contraception are sensitive, particularly for adolescents if there is a risk of being heard by their parents and other family members. Due to the sensitive nature of these topics they have always been carried out as self-completion sections. In previous years the self-completion sections have been entered onto paper by the respondent and then keyed in by the interviewer at a later stage. This paper-based method has the advantage, compared with CASI, that it allows all members of the household to fill out their self-completion forms at the same time rather than in turn, passing the laptop to each person.

It is possible that some information was being withheld due to this method of data collection. The respondent may see it as less confidential than if they key the data into the laptop themselves. A pilot study was undertaken to find out how long respondents would take to complete these questions by a CASI section within the Blaise questionnaire.

A pilot study was conducted out using CASI for sections of the questionnaire which previously had been completed through the paper form. The Audit Trail was enabled on the questionnaire to allow these sections to be timed. These sections included smoking, drinking, family information and contraception. Interviewers had stated that using the paper form usually took between 1 and 4 minutes for a household to complete. Self-completion also provided a welcome break to the respondent, giving them an active task to break up the long sequences of interviewer questioning .

The GHS pilot achieved 153 interviews. The results are displayed in Table 1.

Table 1. Average times for self-completion sections in the GHS

	Paper (concurrent)	CASI (sequential)	Excess of CASI (sequential) over Paper (concurrent)
Family Information	2 minutes	5 minutes	+3 minutes (150%)
Contraception	2 minutes	6 minutes	+4 minutes (200%)
Drinking	4 minutes	7 minutes	+3 minutes (75%)
Smoking	1 minute	4 minutes	+3 minutes (400%)

Obviously the time spent by household in each of these sections depends very largely on the size of the household and how much they have to say for each section. For example, a household where an individual smokes 20 cigarettes a day will take longer to complete the form than a household of non-smokers.

Due to the nature of CASI, each respondent must enter information sequentially on to the laptop. When the sections are conducted via paper, all members of the household can enter information concurrently. This has a large bearing on the increase in time for respondents to complete these sections. It also means that the two methods of data collection are not directly comparable.

It can be seen that, on average, it took an extra 3 minutes to conduct each section by CASI rather than on paper. However the relative increases are large, as much as 400% for smoking. Over the length of the entire GHS pilot interview, CASI added 13 minutes to the length, taking the mean time from 80 minutes to 93 minutes. This would be a large increase in respondent burden if implemented on the production survey.

ATLAS was used to calculate the times in the Blaise Audit Trail from the instrument. This information was used when drawing conclusions from the pilot study. With the help of ATLAS these times were able to be measured very quickly and in a form which enabled research staff to run more in-depth analysis than would have been possible without this package. Times were calculated both for sets of questions and individual questions.

One issue that the analysis investigated was if there was any time saving in using computer-assisted coding (CAC) in Blaise for large coding frames rather than coding from a reference manual. We were particularly interested in occupation coding, which our interviewers code at home after the interview. We did not have a direct comparison of the two methods for the same coding frame, so we explored the issue by an indirect argument. The Audit Trail told us that it took almost exactly the same mean time for interviewers to code occupation by CAC (their normal method) as to code industry from a reference manual (also their normal method): 52 seconds and 53 seconds respectively. However, the cognitive task involved in coding occupation from a reference manual is known to be much greater than the cognitive task in coding industry: occupation is coded to 5 digits as against the 3 for industry, and the coding index for occupation is over 50 times longer and less structured. Therefore we were able to conclude that it would take longer to code occupation than to code industry by the same method of referring to a manual. Since coding occupation using CAC took the same time as coding industry from a manual, it follows that coding occupation using CAC is quicker than coding occupation from a reference manual.

The Audit Trail analysis also allowed us to make design decisions about CASI for the GHS. As the pilot survey showed that CASI required additional time, it was decided that not all self-completion sections of the GHS could be asked as CASI to all sets of individuals. CASI is to be reserved for those sections which are particularly sensitive and, in particular, where confidentiality from the interviewer as well as other household members might be an issue. The Smoking and Drinking sections are mainly sensitive for adolescents in the presence of their parents so, while self-completion remains mandatory for people aged 16 & 17, and optional for all other respondents, it seemed reasonable to continue to use paper rather than CASI. The family information section is potentially more sensitive and for more people: it asks about former as well as present relationships and about all births/abortions and contraception. From 2003, it will be carried out in CASI. Paper self-completion documents are also available for this section for the rare occasions when a respondent is unhappy about using a laptop, or there are several eligible respondents in the household. In the latter scenario, some people can use the paper self-completion forms while others complete the section on the laptop. This will save time in the interview.

A pilot to measure the total time spent on new questions woven into appropriate places throughout an existing Blaise questionnaire

The European Union is planning, through its statistical office, Eurostat, a mandatory survey on income and living conditions (EU-SILC) to be carried out by all member countries. The aim of the survey is to provide information on poverty and social exclusion in the UK that can be compared with the situation in other EU countries. In the UK, the cross-sectional element of the survey will be met by adding questions to the multi-purpose General Household Survey (GHS) which already covers many of the required topics. The information for EU-SILC will be picked up from many individual questions scattered throughout the GHS instrument.

This part of the paper describes the role of the Blaise Audit Trail in the pilot work for the version of the GHS which will deliver the requirements of EU-SILC in addition to the normal GHS requirements⁵. To test the cross-sectional component of the EU-SILC in a pilot study, a CAPI instrument was prepared that integrated the GHS with the EU-SILC primary target variables that it did not already cover. Detailed timing of questions was required for this study to provide Eurostat with information on how well the questions had worked as well and, in particular, how burdensome the sections were to respondents. We needed to isolate and measure the burden that the EU-SILC questions comprised in total. This was a difficult task since the questions were in many groups scattered throughout the combined GHS/SILC pilot instrument. We decided to use the Blaise Audit Trail to provide the level of detail that was needed.

The sample design of the study involved twenty interviewers, covering the range of levels of experience expected in the live survey, to work on the EU-SILC pilot study. Probability sampling methods were used to select 20 Primary Sampling Units (PSUs). Twenty addresses were randomly sampled within the PSU and interviewers were instructed to obtain interviews at 10 addresses in order to achieve 200 household interviews. 203 individual interviews were achieved.

The EU-SILC pilot study was conducted using Blaise, like all ONS social surveys. Table 2 provides information about the length of the combined GHS and EU-SILC

⁵ The GHS is published each year as *Living in Britain*. *Living in Britain 2001* is a web publication at <http://nswbcopy/lib2001/index.html>

interview as well as the length of time taken by the EU-SILC questions alone. The combined GHS and EU-SILC timings include the interview and post-interview administration time (such as calls information and coding at home, e.g. occupation and industry). The time shown as spent on EU-SILC questions, summed by ATLAS from all over the questionnaire, represents only time actually spent in the interview.

Table 2. Length of interview and of EU-SILC module

Pilot study elements	Household Questionnaire	Individual Questionnaire	Complete interview
GHS & EU-SILC*	17 minutes*	1 hour and 48 minutes*	2 hours and 5 minutes*
EU-SILC	4 minutes	42 minutes	46 minutes

*Including administration.

Table 2 shows that the EU-SILC components (whether already covered in the GHS or covered by additional questions) accounted for an average of 46 minutes of the total interview length in the pilot study. However, it should be noted that because (as we knew from other data from the study) respondents had so much trouble understanding the EU-SILC household income questions⁶ the interviewers did not feel able to do what they normally do on ONS income surveys and encourage respondents to find documentation. Hence the questions did not take as long as they would have done if implemented in a production survey. As a result, the final report to Eurostat stressed that these data underestimated the true burden of the EU-SILC components and recommended changes necessary to reducing it. It also stressed that the GHS component was over-estimated in these results, partly through consequential effects of poor EU-SILC questions and partly through the inclusion of administration time. These findings point to the importance of careful interpretation of the results. The Audit Trail and ATLAS can provide measurement but not analysis.

Detailed timing of the entire interview and of the selected areas of interest was provided by the Audit Trail. ATLAS enabled the research team to easily divide the EU-SILC module from the rest of the questionnaire.

The ATLAS tool proved invaluable when analysing data over a number of cases with the way it aggregates timings. In this manner we were able to look very quickly at the length of the interview and make rapid changes where necessary. Once again, the data in this readable form was able to be transferred into SPSS where further investigation could be conducted. As mentioned above, it was important to look at specific sections and how each section related to each other. This can give an indication of whether particular sections are influenced by other parts of the questionnaire.

5. Conclusions

The Audit Trail facility has become a more usable and adaptable tool which can now be used for analysis. When we first began to use the Audit Trail there was a fear that the amount of data produced would have a negative effect on the performance of the laptop. There has been no evidence of this. We are, therefore, continuing to use this function where detailed timing information will prove useful when analysing data or there is a pressing need from the client for detailed times in a questionnaire.

⁶ Not designed by ONS!

ATLAS is a very useful analysis tool but further development of the software (and other systems if necessary) would be required to fit ONS requirements for routine/automatic use.

These developments are:

- ability to distinguish between different types of case (e.g. responding, non-responding cases);

- outputs should have min, max and range of values (field, block, subsections, datamodel) as well as averages;

- data should be able to feed into other information management systems (e.g. response rate, interviewer performance monitoring both for individually and divisionally);

- the user interface could be improved, so that you can look at different reports at the same time (e.g. different months of same survey).

Despite our wish to see these desirable additions, ATLAS has already greatly enhanced our ability to utilise the Audit Trail facility and we will continue to use it for future analysis. We are very grateful to Statistics Canada for permission to use it. Audit Trails have proved useful in both the timing and testing of questionnaires as well as trying new procedures and ways of carrying out data collection. It has been interesting to see what sections take the interviewer a particularly long time to get thorough and to assess where, if any, changes can be made to make interviews less burdensome to the interviewers and respondents.

As the Audit Trail has, as yet, had no visible adverse effects on data collection, there is the ability to collect information routinely on all SSD surveys. With the help of ATLAS this data can now be quickly transformed into a readable form for analysis. The fact that data can be quickly extracted has led to its increased use on designing surveys and questions. We have recently used it to aid the design of the *People, Families and Communities Survey* which is intended to find out about the role individuals and families play in their local community and to explore issues related to social capital.

The Audit Trail has been used on the pilot of this survey to give detailed timing information to the research team. From this information they were able to provide feedback to clients about if there were any particularly long sections or any subject matter where it took more time than expected to collect data.

There is also the potential to use Audit Trail data to give an indication of interviewer performance. For example, it is possible to compare the times interviewers take to complete individual sections of an interview. with due attention to our previous warnings about the need for careful interpretation, this method can be used to judge an interviewer's effectiveness. Audit Trail data can be used more straightforwardly for probity checking.

We would also like to take further advantage of some of the other features of ATLAS such as looking at the number of fields which were answered *don't know* and *refusal*, and at changes or edits to questions. It will be interesting to find out if there is a certain type of question which often has its value changed or edited.

We are continuing to increase our use of the Audit Trail facility. It has become a standard tool for the design of surveys.

6. References

Bumpstead, R., (2001) "A practical application of Audit Trails", 7th International Blaise Users Conference

Hansen, S. and Marvin T., (2001) "Reporting on item times and key strokes from Blaise Audit Trails", 7th International Blaise Users Conference