

Challenges and Opportunities of Upgrading Surveys to Blaise 5

David Kinnear, Office for National Statistics UK

1. Background

1.1 Since 2018 the Office for National Statistics (ONS) has been using Blaise 5 to conduct mixed-mode surveys. We initially upgraded the Opinions survey (OPN) to Blaise 5, closely followed by a Transformed Labour Force Survey (TLFS) pilot.

1.2 The two surveys were the first social surveys using web as a collection method, as part of a multi-mode web and telephone (CATI) collection solution. Blaise 5 was the only serious contender for achieving this.

1.3 Although we have been coding Blaise syntax for many years, the introduction of Blaise 5 presented numerous challenges whilst upgrading our legacy questionnaires. However, we also used the transformation as an opportunity to implement improvements using the additional functionality provided by the new software.

2. Challenges to implementing a Blaise 5 survey

2.1 Learning new functionality

2.1.1 One of the greatest challenges to updating our first social surveys to Blaise 5 was understanding the new functionality that was not previously available in Blaise 4.8, and then implementing it to deliver a successful questionnaire. Coding in Blaise 5 was a straightforward process for experienced Blaise 4.8 programmers. However, the layout template designer was a new concept that took a significant time to implement effectively. The challenge was not only to gain an understanding of designing and creating the numerous templates, but also implementing them in an effective manner. We needed to revisit our standards and look at how we could make best use of the interaction between the Blaise source code and the templates. For example, we set up global user-defined types for use across multiple surveys. These could be referenced by the resource database to enable the automatic application of templates for a particular field if a certain global type was used when defining that field. This negated the requirement for manually assigning templates to fields in the layout designer.

2.1.2 We had previously conducted face to face and CATI for our social surveys in Blaise 4.8. These used a well-established grid template for interviewers to conduct the interview. Initially we attempted to replicate this grid approach for our Blaise 5 questionnaires, with limited success. Although we successfully created templates to deliver this design, at runtime the reduced performance when navigating between the input boxes made this approach unacceptable for interviewers.

2.1.3 The introduction of a web element for data collection essentially drove the replacement design for the traditional grid approach, as a key requirement was for the Blaise 5 questionnaires to meet ONS web design standards. We worked towards these design standards, and our initial templates were developed quickly to meet challenging deadlines. Although they have proved reliable in the live environment, they were developed at a time when we were still developing the expertise in using the resource database. As our knowledge has increased, we have identified areas for improvement; one such area is taking advantage of all the accessibility features that Blaise 5 offers. We are currently working

to implement these features, acting on the recommendations of an external review to ensure the surveys are WCAG 2 compliant.

2.2 The hosting environment

2.2.1 Due to the lack of hosting capability at the time within ONS, the web elements of the OPN and TLFS have been hosted by colleagues at the Northern Ireland Statistics and Research Agency (NISRA). This arrangement has remained in place to the present time.

2.2.2 One of the more significant challenges was ensuring the CATI database was consistent with the latest responses from the web. Hosting each mode on separate databases in different organisations was always going to introduce a time lag between the consistency of the web and phone data. A secure data transfer was implemented between NISRA and ONS, and web data is transferred at regular intervals to the ONS CATI database.

2.2.3 Although the modes are collected on separate databases, the data transfer from the web ensures that the CATI scheduling accounts for any online activity, although not in real time. On opening the web questionnaire, after reaching a specific point in the questionnaire a status field is set to partially completed. At the end of the questionnaire the status is set to completed. The scheduler rules are set in the CATI specification file to take account of the case status. If the respondent breaks off from the web questionnaire without completing it, the status will remain as partially completed. Depending on survey requirements, partially completed cases can be included or excluded in CATI scheduling.

2.2.4 Our objective is to move to a one database approach hosted on our own cloud environment. We have progressed with the questionnaire development but have encountered significant issues during performance testing, with the scheduling of CATI cases severely impacted by the same database receiving web case requests. Although this has delayed our own further testing and development, the Blaise team have released a fix in Blaise 5.15 that will potentially address this issue.

2.3 Business processes

2.3.1 Our existing Blaise 4.8 questionnaires were part of a legacy processing environment built up over decades. Many Visual Basic applications built by IT teams interfaced with Blaise, such as sampling and the interface for telephone interviewers and managers to conduct CATI surveys. The introduction of Blaise 5 impacted many of these critical business processes, making legacy software and processes redundant. New processes have been developed to accommodate the new Blaise architecture; this has and continues to be a major resource activity involving many teams to design and implement.

2.3.2 Prior to the introduction of Blaise 5 and web collection, individual interview cases were built with their relevant sample information loaded into the case object built with IT systems. The shift to web collection placed the emphasis for loading samples into the Blaise database onto Research teams. For the TLFS and OPN web surveys the Blaise package file is built with the sample already pre-loaded for installation on the NISRA hosting server.

2.3.3 For the TLFS and Opinions CATI surveys, hosted in the ONS cloud environment, the questionnaire packages (excluding any pre-loaded sample data) are installed using newly developed web user interfaces. Manipula is then used to load in the sample through a Blaise interface file connecting to the MySQL database used to store the Blaise data. Another web interface has been developed to enable telephone survey managers to access the Blaise 5 CATI dashboard, whilst telephone interviewers can access their scheduled cases.

2.3.4 With all the above changes to business processes, one of the biggest challenges was allocating sufficient time for training teams alongside their existing workloads. Although new interfaces had been provided to make tasks such as the installation of questionnaires into production as user-friendly as possible, some of the production effort had shifted to Research teams. Tasks such as sample loading using Manipula required the teams to gain more of an understanding of Blaise 5 in areas other than updating questionnaire source code.

3. Opportunities

3.1 Improving business processes

3.1.1 Upgrading the OPN and TLFS questionnaires to Blaise 5 provided the opportunity to review existing processes and introduce improvements.

3.1.2 The web element of the mixed-mode pilot would increase the number of respondents contacting the Survey Enquiry Line (SEL) team in Telephone Operations (TO). The SEL deals with any number of calls from respondents, ranging from queries, complaints, refusals to arranging appointments. For a previous unrelated pilot SEL operatives would record outcomes of respondent calls relating to the web survey in a Blaise 4.8 instrument.

3.1.3 For the OPN mixed mode survey the decision was made to incorporate the SEL instrument into the main telephone survey questionnaire, ensuring that the TO Blaise database would always have the latest case information available to the scheduler. This would include, for example, those respondents contacting SEL after receiving the advanced letter to confirm their refusal to take part. The SEL team can access the relevant CATI treatment parallel in the main questionnaire to record the refusal, ensuring that the case has the correct outcome code assigned. The CATI treatment block code has also been standardised for other CATI surveys to incorporate. This promotes the flexibility of TO staff to switch between surveys and be aware of the questionnaire functionality available.

3.2 Revisiting standards

3.2.1 The upgrade gave us the opportunity to introduce new template designs in a radical overhaul of existing Blaise 4.8 questionnaire standards. This was primarily driven for the need to meet Government and ONS design principles for web surveys.

3.2.2 Existing screen design standards had been in place for many years, covering question text formatting, interviewer guidance and question help. These were applied to our Blaise 4.8 questionnaires and designed for use in an interviewer led environment, either face to face or by phone. Data input was by keyboard, and questions were usually displayed in a grid format, with a row designated for each person in a household (see Figure 1).

Figure 1. Blaise 4.8 standard questionnaire layout

I would now like to ask how the people in your household are related to each other.

Code relationship - JIMMY is JOAN's...

Treat relatives of Civil Partners as though the Civil Partners were married.
Also, treat relatives of cohabiting members of the household as though the cohabiting couple were married, including same sex couples.

- 1. Spouse
- 2. Cohabitee
- 3. Son/daughter (incl. adopted)
- 4. Step-son/daughter
- 5. Foster child
- 6. Son-in-law/daughter-in-law
- 7. Parent/guardian
- 8. Step-parent
- 9. Foster parent
- 10. Parent-in-law
- 11. Brother/sister (incl. adopted)
- 12. Step-brother/sister
- 13. Foster brother/sister
- 14. Brother/sister-in-law
- 15. Grand-child
- 16. Grand-parent
- 17. Other relative
- 18. Other non-relative
- 20. Civil Partner
- 99. (Office use only)

| NameRel | R[1] | R[2] | R[3] | R[4] | R[5] | R[6] | R[7] | R[8] | R[9] | R[10] | R[11] | R[12] | R[13] | R[14] | R[15] | R[16] |
|------------------|------|------|------|------|------|------|------|------|------|-------|-------|-------|-------|-------|-------|-------|
| QAskRel[1] JOE | | | | | | | | | | | | | | | | |
| QAskRel[2] JOAN | | 1 | | | | | | | | | | | | | | |
| QAskRel[3] JIMMY | | 3 | 3 | | | | | | | | | | | | | |
| QAskRel[4] | | | | | | | | | | | | | | | | |

3.2.3 As mentioned earlier, we initially attempted to replicate the grid approach in an early build of Blaise 5 but encountered performance issues. A combination of this and the need to meet ONS design standards for planned web collection provided the drivers for revisiting our questionnaire layout designs. Our first surveys upgraded to Blaise 5 were mixed mode web and CATI surveys. The same questionnaire instrument would be used for both collection modes, so the proposed web design influenced the design templates for the interviewer led mode.

3.2.4 Although the overall look and feel of the main page templates would look similar, Blaise 5 offered the functionality to design template variations to meet differing requirements dependent on collection mode. We have been developing a resource database to use across our typical social surveys, incorporating mode-specific layout templates that are automatically applied to ensure the corporate standard is adhered to.

3.2.5 Figure 2 is an illustration of the layout set for interviewer led surveys, with features enabled to assist the interviewer. These include question type prompts (such as ask or record), additional question guidance, input boxes for those who prefer keyboard data entry and additional remarks to a question response. Field names are also displayed (as they were in the Blaise 4.8 display) to help interviewers in reporting issues for questions that respondents have difficulty with.

3.2.6 To aid interviewer navigation, a side panel listing key areas of a questionnaire is displayed. These key areas are defined in the parallels section of the questionnaire source code. Additionally, CATI surveys have access to parallel sections to enable the recording of call details such as refusals, appointments, line busy etc. These are positioned in a parallels area at the top of the display.

Figure 2. Typical telephone survey display

The screenshot shows a web-based survey interface for 'Opinions and Lifestyle Study' with 'Case: 10030'. At the top, there are navigation tabs: Answering Service, Appointment, Busy, Disconnected, No Answer, Refusals, and Others. Below these, a 'Previous' link is visible. The main heading is 'About your household. Enter your details.' with a blue link 'ASK OR RECORD'. The section is titled 'Person 1' and contains three input fields: 'Enter title of respondent' (Title), 'Enter first name of respondent' (FstName), and 'Enter surname of respondent' (SurName). The 'FstName' field contains 'Joe' and the 'SurName' field contains 'Bloggs'. At the bottom left, there is a green 'Save and continue' button and a '>>' button. On the right side, there is a sidebar menu with 'OPN2310_OSH' selected, and other items: 'Information Sheet', 'About the household', and 'Admin'.

3.2.7 Designing the web survey templates drove the move away from the traditional Blaise 4.8 grid display shown in Figure 1. Several factors were taken into consideration in the design. The look and feel of the web page had to meet ONS design standards. The page has a straightforward presentation with none of the navigation features or additional functionality present in the CATI display (see Figure 3).

Figure 3. Typical web survey display

The screenshot shows a web-based survey interface for 'Opinions and Lifestyle Study' from the 'Office for National Statistics'. At the top right, there is a 'Save and sign out' button. Below the header, there is a 'Previous' link. The main heading is 'About your household. Enter your details.' The section is titled 'Person 1' and contains three input fields: 'My Title', 'My First or Given name', and 'My Surname or Family Name'. The 'My First or Given name' field contains 'Jo' and the 'My Surname or Family Name' field contains 'Bloggs'. Below these fields is a green 'Save and continue' button. At the bottom, there is another 'Previous' link and a footer with links for 'Contact us', 'Accessibility', 'Confidentiality', and 'NHS Mental Health Resources'. At the very bottom, there is a copyright notice: '© OGL All content is available under the Open Government Licence v3.0, except where otherwise stated'.

3.2.8 The traditional grid display for our legacy interviewer led surveys was not suitable from a security perspective for web completion by individuals. Household members would have their responses quite clearly displayed in the table rows. With respondents self-completing on the web, the new design of enabling a respondent to complete their own individual section of the survey with an isolated display of typically one question per page helped fulfil the security requirement.

3.2.9 Introduction of web collection also necessitated the implementation of templates for small screen devices. It is imperative that enough resource is assigned for the design and testing of templates on different devices and browsers to optimise displays and uncover issues. Our testing processes have been significantly enhanced because of the introduction of web collection.

3.2.10 Our definition of standards not only applied to screen designs, but naming conventions for fields and type names. We created standard field, auxfields and types blocks to be included in all Blaise 5 questionnaires. Incorporating these blocks ensure that layout templates are automatically applied to fields or auxfields that have a named user-defined type referenced in the resource database. Inclusion of these blocks across surveys also maintain a consistent naming convention and definition for key variables used in questionnaire functionality and outputs.

3.3 Implementing a standard structure to surveys

3.3.1 We used the upgrade to Blaise 5 to apply a generic format to our most widely used questionnaires. To achieve this a set of core Blaise question code blocks have been created, accessible to all our surveys. Applying these ensures that a harmonised set of key variables is collected, such as household data, questionnaire completion status and outcome codes.

3.3.2 Applying core code blocks across different surveys involves considerable planning and testing. Whilst standardising the code was the objective, we were mindful that there would be instances where survey specific code or structures would be required to meet customer requirements. To achieve this requirement, we used a combination of coding techniques.

3.3.3 Firstly, we used conditional routing based on the survey name to apply certain logic to control survey specific variations in question routing or text.

3.3.4 Next, we used prepare directives to determine the structure based on the survey type. The TLFS questionnaire is asked of all eligible people in a household, whereas the OPN survey is targeted to a pre-sampled individual. The household array is therefore defined to accept a maximum of 10 people for the TLFS, and one person for the OPN survey. This eliminates the creation of a substantial number of columns that will never be populated in surveys targeting just one sampled person.

3.3.5 Finally, we used the resource database to apply appropriate layout templates to reflect any survey specific differences within the core code blocks. For example, the TLFS page that launches the targeted section for a selected respondent (Figure 5) has different fieldpane templates applied in contrast to a survey where one sampled person is targeted. This is the case for the OPN surveys, where a text only display is presented to the respondent (Figure 6).

3.3.6 The generic format of our standard questionnaires can be broken down into three key components: household details, individual section launcher and the individual section containing topic blocks. The first section of a questionnaire collects household information such as person details, relationships, occupancy details (Figure 4). The information provided in this section populates key values such as names for re-use in question text fills throughout the questionnaire and the household size, to control the maximum number of times the individual section loop must be repeated.

Figure 4. Household details

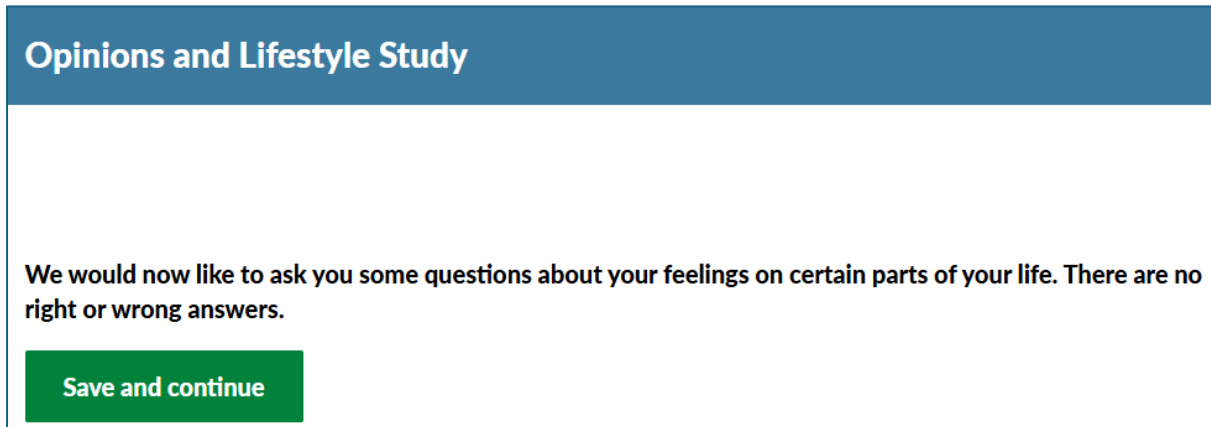
3.3.7 The next section, the individual section launcher, displays the list of individuals recorded in the household section (Figure 5). The interviewer or respondent can then launch the individual section of the survey dedicated to that person. Additional information on the status of an individual’s section of the survey is also displayed to interviewers when CATI mode is applied. An individual refusal can be recorded here by clicking the warning icon, and the refusal status is displayed in the button text. This can be reversed if required.

Figure 5. Individual section launcher – household survey

3.3.8 Figure 6 highlights how Blaise 5 can be used to display a completely different look dependent on the type of survey. The individual section launcher is instrumental in maintaining a standard structure for our surveys and enabling standard flags and derived variables to be set. However, for those surveys targeting a sampled individual, such as the OPN survey, a list of household individuals is obviously redundant. Applying survey-specific routing, prepare directives and templates to the same Blaise code

block, we have influenced the flow without compromising on the standardised structure and behind-the-scenes functionality of the Blaise questionnaire.

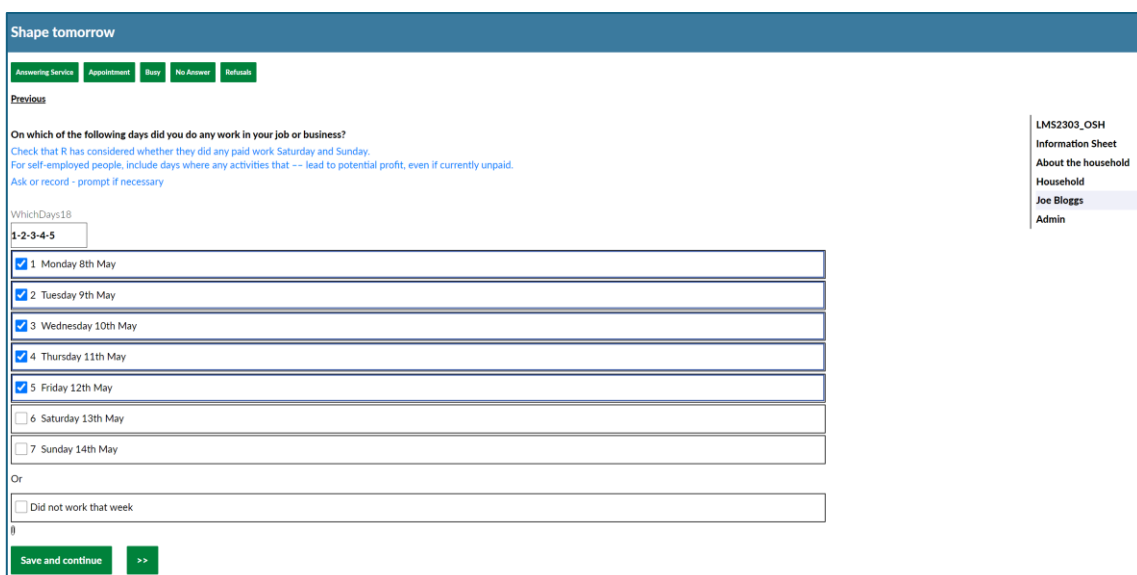
Figure 6. Individual section launcher – sampled person survey



3.3.9 Finally we have the individual section containing topic blocks. For household surveys, everyone recorded in the household section populates an element of an array, up to a maximum of 10 people. The individual array block defined in the source code contains all the individual topic blocks for the survey. Nesting each subject block within a parent individual code block made it easier to implement some of the key functionality requirements around person level response security for web collection and navigation.

3.3.10 Figure 7 shows a typical individual section question page with the CATI templates applied. Although the same structure is used for both web and CATI collection, the enhanced security features that apply to web are not invoked for interviewer led questioning. Derived variables are populated to identify if a topic within an individual's section of the survey has been completed. If so, conditional routing applying to web only ensures that the subject questions are no longer asked, but the responses are kept. CATI interviewers have unrestricted navigation throughout the questionnaire.

Figure 7. Typical individual section question page



4. Conclusions

There were many challenges to implementing Blaise 5. The impact on legacy systems and processes was significant. Resource had to be planned to not only deliver the technical solutions, but also train the many people impacted by the upgrade. However, it also provided the opportunity to revisit these processes and rather than simply replicate, identify where new and more efficient methods could be introduced. Survey teams can now install questionnaires with user-friendly web interfaces, load samples without IT involvement, phone operatives handling incoming refusal calls and recording details directly in CATI instruments are all positive examples of improvements delivered from the upgrade.

Although we have a team of experienced Blaise programmers, it became apparent that significant time had to be invested in understanding the resource database to implement it in an optimal fashion. Even now we are not using it to its full potential, so our continued development in understanding its functionality is essential in our drive to improve data collection.

Of course, many of the above challenges would not have been resolved without the ongoing support of the Blaise team, who have been excellent in responding to our many queries and requests. The ongoing development of the software will present more opportunities to improve our data collection processes in ONS.