

# Converting a Blaise2.5 Questionnaire to a Blaise4W Instrument

By

John O'Connor and Jacqueline Hunt, Central Statistics Office (CSO)

## 1. Introduction

The Quarterly National Household Survey (QNHS) commenced in September 1997 replacing the annual Labour Force survey which had until then been accepted as the definitive source of data on employment and unemployment in Ireland. This had been conducted each April/May on a biannual basis from 1975 and annually since 1983.

The QNHS was the first major survey to use Computer Assisted Personal Interviewing (CAPI) in the CSO. CAPI was introduced to ensure an optimum turnaround time from the start of data collection to publication of results. Blaise2.5 was used to write the questionnaire with a front end written in Microsoft Visual Basic and data productivity and quality checks written in Access2.

The QNHS field staff consists of ten regional Co-ordinators and one hundred and thirty Interviewers. Each Co-ordinator is responsible for thirteen Interviewers. Approximately, three thousand households are interviewed each week or thirty-nine thousand over the course of the quarter.

The sample design for the QNHS comprises 2,600 smaller areas or blocks. This set of blocks remains fixed for five years. Each block contains, on average, 75 households, of which 15 are sampled in a quarter. Households participate in the survey for five consecutive quarters and are then replaced. This means that, in each quarter, 20% of the households are rotated out of the survey and replaced by new households in the same block.

Initially, the hardware in the field consisted of 200 identical Interviewer's laptops and 10 Co-ordinator's P.C.s. However, since then due to wear and tear two other types of laptops have been introduced.

At the end of 1998 it was decided to migrate from Blaise2.5 to Blaise4W because of doubts as to the formers Y2K compliance. This paper outlines the work involved in the move and the reaction of the developers to Blaise4W.

## 2. Moving to Blaise4W

### 2.1 *Starting the Project*

We first set up a project team specifically to look at the move to Blaise4W. The reasons for using project management techniques were threefold:

- it was a new project, we had not done a similar type of project before and we therefore lacked the requisite experience
- it was relatively complex
- there was no margin for error, we had to get it right first time.

The project team included people from the IT area, business area and field staff. Although there was a large number of people involved, their roles and responsibilities meant that their input was not required on a full time basis for the duration of the project. The first priority of the project team was to define the scope and objectives of the project. We felt it was important to set out the goals from the start as it provided focus and clear objectives for everybody concerned.

The scope and objectives were to:

- rewrite the system maintaining the structure and functionality of the original system
- successfully convert our existing data collection and productivity and quality checking applications used by the Co-ordinators and in-house CSO team.

Having agreed the scope and objectives we proceeded to draw up a project plan and work schedule.

## 2.2 *Blaise4W Training*

The first requirement was to train the development team in Blaise4W. Both of us went on a week long training course in December 1998. The course consisted of four days Blaise4W training and an introductory day on Maniplus. We were also given time to work with the product in order to become comfortable and confident with the software.

We then began to look at ways of addressing issues such as questionnaire design, data conversion etc. Once we started to form a plan of what we needed to do and how to do it we brought in a Blaise4W consultant from Statistics Netherlands.

There were three reasons for this:

- to confirm our approach was correct
- to provide technical help in achieving our objectives
- to ensure best practice and efficiency.

## 2.3 *Questionnaire*

The new Blaise4W version of the questionnaire was due to go live in Quarter 4 1999. A pilot was scheduled for Quarter 3 1999. This was to be conducted in conjunction with the usual Blaise2.5 questionnaire. The aim of the conversion was to produce a Blaise4W version of the questionnaire maintaining the structure and functionality of the original Blaise2.5 questionnaire.

The Blaise2.5 questionnaire had been in use since Quarter 4 1997 and had evolved to its current complex structure since then. For example, the Quarter 3 1999 questionnaire consisted of 4,447 lines of code split over sixteen files and blocks. The questionnaire was required to accommodate up to twenty people per household which necessitated the use of a block of array type one to twenty within a table. This type of structure was also used to handle separate questions on emigrants which was required to accommodate up to six people.

The new datamodel was to have a similar structure to the Blaise2.5 questionnaire (see Appendix). Every block in the new datamodel was based on a block in the Blaise2.5 version. The opportunity was taken to rename the blocks with more meaningful names and each block was established as an *INCLUDE* file due to their size. The new datamodel consisted of thirteen blocks, a procedure and a library. The procedure was a new addition and is used to perform checks on the numeric key fields. Using a procedure allowed us to discard several lines of repetitive code. A library of answer types was also created which can be reused every quarter.

When rewriting the questionnaire as a Blaise4W instrument the main objectives were:

- Preserving the layout to keep a familiar look and feel for the business area and field staff.
- Reducing ongoing maintenance. Every three months we release what is essentially a new system with additions and amendments to the questionnaire. We identified and wrote the code in modules which would rarely need to be changed.
- Providing a similar level of functionality for the Interviewers from the available menu options within Blaise4W.

## 2.4 Handling External Files

The questionnaire accesses five different external files sources.

1. **Lookup1**  
Block1 : STRING[4]  
WkNo1 : 01..13  
Rot1 : 1..5
2. **Lookup2**  
Rotn2 : 1..5  
Year2 : 1997..2020  
Quar2 : 1..4  
Rota2 : STRING[1]  
Wave2 : 1..5
3. **Lookup3**  
Year3 : 1997..2020  
Quar3 : 1..4  
WkNo3 : 01..13  
Day3 : 01..31  
Mth3 : 01..12
4. **Refquart**  
SDS4 : STRING[3]  
Quar4 : 1..4  
Year4 : 1998..2000
5. **Previous Quarter's Questionnaire**

The information on the lookup files (1 to 4) combine to give all possible reference data for each block over the entire QNHS sample. This information relates to reference week nos., reference week dates and quarter information. The lookup files are the same on all the QNHS laptops which allows the laptops to be moved between Interviewers and from Co-ordinator area to Co-ordinator area if necessary.

The fifth external file is the previous quarter's questionnaire which is accessed on all repeat calls to participating households. The current quarter's questionnaire is required to access the previous quarter's questionnaire and bring forward any data necessary for routing purposes.

### 2.4.1 Lookup Files

With Blaise2.5 it was possible to run an *External File* option which converted the external lookup files into a format that Blaise2.5 could access. This option generated two new file types with the extensions *.dat* and *.idx*. To access the external files in Blaise4W the files first had to be converted into Blaise4W databases. Blaise4W datamodels were prepared based on the files' record description. We then used the Manipula application to convert these Ascii files into four separate Blaise4W databases.

With Blaise2.5 the record layout of all the external files had to be included in the Blaise2.5 questionnaire code. We were using the *READFILE* command to access the external files which is a complex command involving the use of five parameters i.e.

*READFILE (IndexFileName, Key, DataFileName, InputVariable, Result).*

The fifth parameter is based on the search for the key i.e. *0* if the record was found and *1* if the record could not be found.

In Blaise4W the external files are listed in the *USES* section of the datamodel and in the *EXTERNALS* section where the read is to take place. Blaise4W includes the *SEARCH* and *READ* statements which allows more readable code to be developed. The *SEARCH* is used with an *IF* condition and if the *SEARCH* is successful the *READ* is done i.e.

*IF BlaiseDatabaseName.SEARCH(key) THEN BlaiseDatabaseName.READ.*

The overall effect was to reduce the amount of code required and the end result was more transparent, readable code.

## 2.4.2 Previous Quarter's Data

The Blaise4W instrument was due to go live in Quarter 4 1999. It was required to access Quarter 3 1999 data collected using Blaise2.5 if a repeat visit to household took place. The first step in converting the Blaise2.5 data into a format that Blaise4W could read was to produce a full file description of the questionnaire fields we wanted to access. Once we had established the file description we used Blaise2.5 manipulas to convert the Blaise2.5 data into Ascii format. The next stage was to convert the Ascii data file into a Blaise4W database. This was achieved by first creating a datamodel based on the Ascii file layout and then using a manipula setup to convert the Ascii data into Blaise4W format.

Reading the previous quarter's data in Blaise2.5 had been a complex procedure. A full file description of all the fields in the previous quarter's questionnaire had to be included in the *EXTERNAL* section. An external file description of the previous quarter's questionnaire was generated using the *UTILITIES* option from the Blaise2.5 menu. The file description was then incorporated into the new questionnaire.

To access the previous quarter's data the question cluster numbers had to be identified. Question clusters are created in questionnaires where one or more blocks are declared as subfiles. Blaise2.5 creates one cluster for every question referred to as a subfile block. Clusters are identified by sequence numbers corresponding to the order in which subfile questions are referred to in the route paragraph at the highest level. The cluster numbers for each person within each block were identified and then used in the *READBLAISE* statement i.e.

*READBLAISE (External Key, External File Name, Block Name, Cluster No., Result ).*

Blaise4W does not require a full file description of the previous quarter's questionnaire. We included the questionnaire in the list of external files in the *USES* section of the datamodel and in the *EXTERNALS* section where the read is to take place. We could then access the data using the *READ* and *SEARCH* commands as described already. To access particular fields from the previous quarter's data we could use dot notation to identify the exact location, for example,

*PrevQnhs.PersonGrid.Person[Line No].Individual.Name.*

The procedure to convert the Blaise2.5 data into Blaise4W format was eventually combined as an option on the Interviewers' Visual Basic front end menu. This enabled the Interviewers to convert the Blaise2.5 interviews from the previous quarter into Blaise4W format in order to proceed with Blaise4W interviewing in the current quarter. They could then complete any unfinished or late interviews as normal in the previous quarter and when complete convert the data again using the same procedure. This option was only required for Quarter 4 1999 as from that quarter on all the data was in Blaise4W format.

The main considerations accessing the Blaise2.5 data were:

- To convert and/or drop data not required to be carried forward into Blaise4W. This was done by identifying what could be dropped and then writing a Blaise2.5 manipula to convert the data into an ASCII file.
- Creating a Blaise4W instrument to reflect this ASCII file and then converting the ASCII file into a Blaise4W database.
- Rearranging the data in a different layout to reflect the structure of the new datamodel.
- Ease of use for the Interviewer. We set up the conversion procedure to run when "clicked on" by the Interviewer. This option converted any Blaise2.5 data on the laptop from the previous quarter into Blaise4W format.

The biggest problem we encountered with the data conversion concerned the Blaise2.5 manipula. This manipula is a DOS based program and would not run on the Interviewer's laptops due to an insufficient memory problem. This problem occurred because of the amount of memory allocated to DOS based programs by Windows 95. Our solution was to reduce the size of the manipula, however, the effect of this was to limit the amount of data it was possible to carry forward into the Blaise4W questionnaire.

## 2.5 Other IT Applications

Our old applications used a combination of Blaise2.5, Visual Basic, Dos and Access2. With the move to Blaise4W it was necessary to review the use of all the other software tools.

We decided to update the front ends, written in Visual Basic, to compliment the new look of the Blaise4W screen. This was quite a radical change as it involved moving from command buttons to drop down menus and toolbars. We were able to remove all the DOS programs and use Visual Basic instead to connect with the Blaise4W Dep and Manipula applications. We built in additional message boxes to provide the field staff with more information on the outcome of the procedures they were running.

The Co-ordinator and Office Applications used Access2 for productivity and quality checks on work done by the Interviewers. After the review it was decided to upgrade the software to Access97 mainly because Access2 was not Y2K compliant. The upgrade incorporated a redesign of the screens compatible with the new look of the Blaise4W and Visual Basic screens. The redesign also had to handle the new AsciiRelational files created by the Manipula applications which we used for transferring data.

## 2.6 Data Transfer

With the old applications the data was transferred from the Interviewer to the Co-ordinator and then to the CSO. This was done by converting the data to AsciiRelational files on the Interviewers' laptops, sending the data to the Co-ordinator who loaded the data to a Blaise2.5 database on their PC. The Co-ordinator then checked for productivity and quality and if satisfied sent the data to the CSO also in AsciiRelational format. This was done on a weekly basis and only "new" and "changed" forms were sent. There was also an option to send cumulative data by both the Interviewer and the Co-ordinator. This option would be used primarily as a security procedure.

We needed to replicate this scenario in Blaise4W using the manipula application. This required careful consideration, as there are subtle differences between the old convert features of Blaise2.5. and the new Blaise4W manipula application.

We used the *HISTORY* method to select the interviews. The result of the method is "new", "changed" or "unchanged". Using this method we could select only the interviews that were new or had been amended. We then used the *RESETHISTORY* method to mark all the interviews as "unchanged". For example,

```
IF (InputFile1.HISTORY = NEW) OR (InputFile1.HISTORY = CHANGED) THEN
    OutputFile1.WRITE
    InputFile1.RESETHISTORY
    InputFile1.WRITE
ENDIF.
```

We also included an option to send cumulative data by both the Interviewer and the Co-ordinator. Again this option would be used primarily as a security procedure.

### **3. Preparing for Live Release**

#### **3.1 Testing**

From a testing point of view we regarded the applications as completely new and as such we would need to test them in the same way as we would test any other new application. To do this we drew up test plans. These test plans were designed to provide a structured systematic approach to testing.

Using the test plans we developed test data. Testing was first done in the IT area where each program was tested individually. We then carried out suite tests and finally the whole system was tested. At this stage, the systems were passed to the business area for further testing (this included both in-house and field staff). This paid off as no major bugs were found after live release.

#### **3.2 Pilot**

A pilot was scheduled for Quarter 3 1999 which was to be conducted in conjunction with the usual Blaise2.5 questionnaire.

The objectives of the pilot were to:

- assess the Interviewers' reaction to the new applications
- identify areas of difficulty for field staff e.g. problems with layout, clarity of messages, procedures etc.
- identify any software problems from source Interviewer laptop to CSO e.g. problems with routing, data transfer etc.
- assess training requirements for the field staff.

#### **3.3 Training Field Staff**

All field staff were trained centrally in the CSO. The training course was developed based on the existing knowledge base of the Interviewers and the experiences of the Interviewers involved in the pilot.

Groups of approximately thirty Interviewers were brought together for a two day training session. They were first given an opportunity to see the new application and then they were brought through a number of sample interviews. They also had an opportunity to run each procedure available on the menus.

The training sessions provided an opportunity for team building and standardisation of work practices throughout the 10 different Co-ordinator areas. Sample applications were loaded to the Interviewers' laptops which gave them time to practice and get familiar with the new system at their own pace.

#### **3.4 Live Release**

The live release was scheduled for September 1999. Final training sessions took place in the 10 Co-ordinator offices and the final versions of the applications were loaded on the Interviewer's laptops. These training sessions built on the earlier training and provided an opportunity to deal with any further questions by the field staff.

Some new laptops also being introduced at this stage resulted in an unexpected problem. The Blaise2.5 manipula would not run on the new laptops due to an insufficient memory problem. This was due to the amount of memory being allocated to DOS based applications on a Windows 95 machine except in this case because the machines were new they had a later version of Windows 95 which had further reduced the amount of memory available. The data conversion from Blaise2.5 to Blaise4W had to be done on the old laptops and then the converted data was copied onto the new laptops. Unfortunately this put the trainers under pressure having to deal with the extra workload.

## **4. Review of Move**

### **4.1 Conclusions**

The project began in early December 1998 with the initial training in Blaise4W. The pilot was successfully conducted from June - August 1999 and the new system was released live in September 1999. A considerable effort was required to support the field staff from the time of the live release until March 2000.

By adopting a project management approach to the development we were able to provide a less stressful environment to work in and also provide a focus for everybody involved. It was an invaluable tool in providing an early warning system if deadlines began to lag or problems were encountered.

A project like this needs to be well thought out and planned before the start or there is a huge potential for failure. Active involvement by the business area and the use of "RAD" techniques were essential in providing a system which the business area were happy and comfortable with.

The pilot was important for a number of reasons. It provided an opportunity to assess the field staff reaction to the new application. It helped identify training requirements and it allowed us to monitor the new applications in the live environment.

A good training programme was important to ensure that the Interviewers were both comfortable with and had a good understanding of how the new system works resulting in relatively smooth transition from the old to the new applications. Bringing everybody together for the training reinforced good work practices and standardised work practices countrywide.

We were also fortunate to have good support from Statistics Netherlands.

Overall the project went well in spite of some "minor hiccups" towards the end. The transition went according to plan and was relatively smooth.

### **4.2 Some points worth noting from our experiences**

- "Feel good" factor of pilot Interviewers from their involvement in the project.
- Attachment of Interviewers to certain features in the old system that were missing from the new system.
- Information messages not as clear as we thought, different interpretations possible.
- We originally used the standard Blaise4W screen, however, the Interviewers found the screens difficult to look at e.g. background colour, font size, etc. We changed the background colour and font in the subsequent quarter.
- Data transfer using the Manipula application had to be carefully thought out as there are differences between the Blaise4W and the Blaise2.5 manipula and convert applications.
- Regular occurrence of the "Hospital" error message. Blaise4W seems to be susceptible to data corruption if there is either a hardware or software(Win95) problem on a machine. At this stage, due to the frequency of the occurrence of the "Hospital" error we are moving to a later release of Blaise4W for Quarter 3 2000.
- Mysterious appearance of `~lc` and `~la` files. We could not find any reference to these files in the manuals.
- Different hardware. Applications working on one platform and not on another e.g. different models of laptops.
- Different versions of Windows 95. The amount of memory available to run DOS based applications varies depending on the version of Windows 95 on the machine.
- Developing Visual Basic applications in a WindowsNT environment for roll out to a Windows95 environment.
- Duplicate forms appearing on the Browse Option. This is still under investigation by the CBS.

### **4.3 Additional Benefits**

- New laptops unable to run large DOS based applications which resulted in Blaise2.5 falling over. This is not an issue with Blaise4W.
- Reinforce good practice during training (backups procedures etc.).

## *Appendix*

### *1. Original Blaise2.5 Questionnaire Structure*

```
Questionnaire Qnhs
  Block Main1
    Table Main2
      Block OverAll : Array [1..20]
        Block Din2
        Block Res1
        Block Hist
        Block Age1
        Block Rel
        Block Header
        Block Weekend
        Block Q2B
        Block Q2Ba
        Block Q3B
        Block EduMod
      End Table
    Block Hmod
    Block Smod
    Block Emig
      Table Emig
        Block Emig : Array [1..6]
      End Table
    Block Appointment
  End Questionnaire
```

### *2. New Blaise4W Datamodel Structure*

```
DataModel Qnhs
  BlockMainDetails
  Table PersonGrid
    Block Person : Array [1..20]
      Block Individual
      Block History
      Block Relation
      Block Employ
      Block Employment
      Block IndOcc
      Block HrsWorked
      Block JobSearch
      Block Education
    End Table
  Block Housing
  Table EmigGrid
    Block Emigrants: Array [1..6]
  End Table
  Block MakeAppoint
End Model
```