

# Converting Paper Collection on the International Passenger Survey to Blaise 4 on a Tablet Computer

*Mike Hart, Office for National Statistics, United Kingdom*

## 1. Abstract

The International Passenger Survey (IPS) is a continuous survey that has been running since 1961 by the Office for National Statistics (ONS). This is the main source of information in the UK on international travel and tourism, as well as the associated expenditure.

Between 700,000 and 800,000 face-to-face interviews are conducted every year at the major air, sea and tunnel routes to and from the UK by a national team of 240 interviewers.

At the moment, the IPS data is collected on paper forms which are then manually typed into Blaise 4.8 and transferred electronically to ONS head office where data cleansing and validation take place.

ONS policy is to be digital by default and this coupled with the need to produce efficiencies has placed an emphasis on exploring more efficient of collecting and validating the data we collect.

This paper follows on from the work undertaken in 2014 and will demonstrate some of the challenges we have face and how we have overcome with them.

The vision remains the same:

- Speed up data collection
- Improve data quality
- Develop a collection instrument that is user friendly.

## 2. Introduction

All UK government departments have had a directive from central government to be digital by default and the International Passenger Survey (IPS) is the only Social Survey we still collect the data on paper. So naturally our attention has been focussed on removing paper from the data collection exercise. We have been looking at this for a couple of years now and initial investigations were carried out with Blaise 5, but we are now looking at Blaise 4 as the plan is to reuse as much of the current systems and adapting these systems to work with Blaise 5 was considered too onerous.

In this paper we report on our progress towards transferring the International Passenger Survey (IPS) from a paper form to an electronic questionnaire on a tablet. The paper will reflect on learning and concepts developed in Blaise 5 and how we have reflected this learning in the Blaise 4 instrument. Whilst the directive from central government is digital by default we have to maintain data quality and ensure the collection exercise is not a burden on the respondent.

## 3. What benefits for the survey?

The IPS is a continuous survey that has been conducted by ONS since 1961. Between 700,000 and 800,000 interviews are conducted every year at major airports, seaports and tunnel routes covering both departures and arrivals terminals. This is the main source of information used by the UK governing bodies for planning, monitoring and informing decisions on tourism and immigration policies.

At the moment, the IPS data is collected on paper forms which are then manually keyed into the CADI system in Blaise 4.8. The data is then transferred electronically to an ONS data centre where

editing and validation take place.

**Figure 1. Conducting Survey**



The vision for the project as previously stated in the ONS paper delivered by Lan Benedikt in Beijing 2015 [1] remains the same:

- Ensure data collection at point of collection is no slower than collecting the data on paper
- Take advantage of in-built validation and the routing capability in Blaise to cut down on human error and improve data quality
- Eliminate manual post data collection tasks

#### **4. First prototypes**

The project started in July 2014 and a questionnaire was developed in Blaise 5. A site visit was carried out at Bristol airport to gain an understanding of how the field work was conducted and how it could be adapted to collect the data on a Tablet.

Three main areas for improvement were identified.

1. Survey flow: the IPS questionnaire has complex routing with a Main Questionnaire and various trailers e.g. Immigration Trailer, Employee Trailer, Student Trailer short/long forms [2][3], making it sometimes difficult to quickly identify the next question to ask, especially for inexperienced interviewers.

2. Coding frames: IPS interviewers usually carry a folder with all the lookup tables that they need for filling out the survey. Examples include flight numbers, airport IATA codes, country codes, region codes, UK towns. Other documents they carry include the trailers mentioned above, as well as “show-cards” in 16 different languages. It is difficult to shuffle through the folder to quickly find the correct forms, even more so because IPS data is collected on the move.

3. Data keying (CADI): this process is unanimously perceived as very time consuming and cumbersome. So much so that most field managers have to take work home to complete their tasks on time and manual data keying also has the potential to introduce human error. The elimination of this step is where significant improvement can be achieved, both in terms of efficiency and data quality.

The good news was that most of the problems identified could be solved with routing, in-built checks and lookup features. Thus, a first prototype of the electronic questionnaire was developed and deployed on a tablet computer, Lenovo Miix2 8-inch Windows 8.1, in order to explore what could be done. Examples of screenshots are shown in Figure 2.

Figure 2. Look-and-feel of the First Prototype of the IPS Questionnaire in Blaise 5. Different background colours were used to differentiate various trailers and the main questionnaire.



At the end of August 2014, the IPS research team organised a focus group in Newport during which the prototype was demonstrated to a group of interviewers and field managers from Bristol, Manchester and Liverpool airports. The main aims of this meeting were to gauge users' first reactions to the new survey instrument and to gain feedback on what could be improved. The results were broadly positive and no major blockers to implementation identified (the results are available in the ONS paper delivered by Lan Benedikt in Beijing 2015 [1]).

## 5. Change of direction from Blaise 5 to Blaise 4

Since late 2015 there has been a change of direction within the project and we have decided to implement this survey with Blaise 4 rather than Blaise 5. The driver behind this decision is that ONS is undergoing a large transformation project and significant resource is being devoted to developing new systems and it was considered that moving to Blaise 5 would divert resource away from the transformation project.

ONS has recently moved to a new methodology when developing software, AGILE. There are 12 principles underpinning the Agile approach (The Agile Manifesto [4]) and the first states that 'our highest priority is to satisfy the customer through early and continuous delivery of valuable software'.

So with that principle in mind our first task was to identify the customer. For a statistical office you would think the customer would be the end user of the data but in this instance the customer is the interviewer collecting the data. This is because they are the user of the software and without acceptance of the software from them then the final product, the statistical outputs, may be flawed because the interviewer is not able to perform the function required of them.

Another key concept of AGILE software development is 'fail fast, fail often' [5], whilst failing may sound undesirable it is a fact of life. The point to be understood about 'fail fast' is the second word, it is about reducing delay. If a failure is going to take place you want to reduce the time lag in detecting the failure and relaying the detection back to the developer(s). Failing fast and often also gives you an opportunity to terminate a project without expending lots of resource and money to a project that is never going to deliver should an insurmountable problem emerge.

To complete the new development framework ONS is moving to DevOps teams where teams are made up of multi-skilled individuals to help remove silos and where the people understand they are all on the same side, working together to build and develop software fit for purpose.

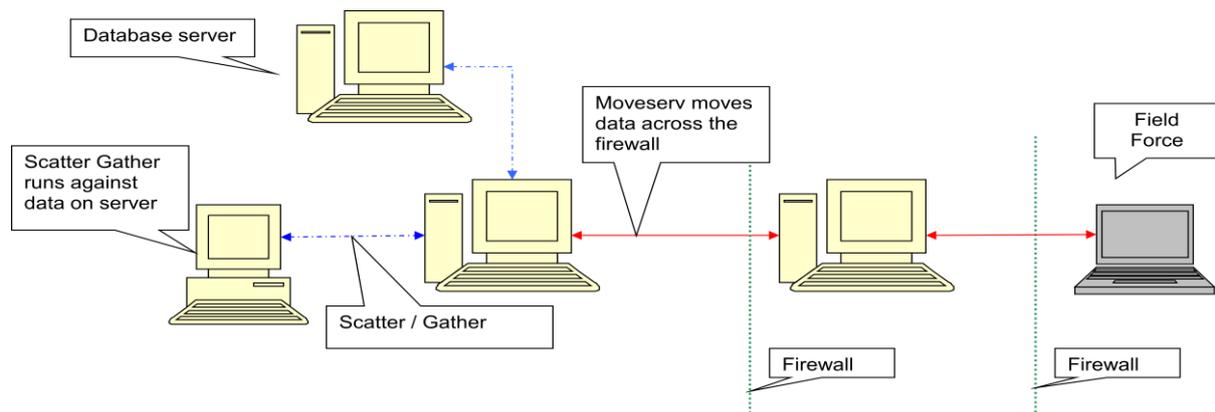
## 6. Testing

With the AGILE concepts fresh in our minds and looking to build on the previous work done in Blaise 5 we adopted a new approach to delivering the project. A series of Sprints were defined, each being designed to build on the previous one and add incremental value to the project.

### Sprint 1 – Using Legacy systems can we transmit and receive data from the Tablet

This sprint focussed on plugging the Tablet into the existing transmission systems (see Figure 3). As we were reusing existing systems we were not expecting any issues and none were encountered. The test was still valuable as this is a key part of the system and if data could not be transmitted to and from the Tablet then this needs addressing before progressing to the next stage.

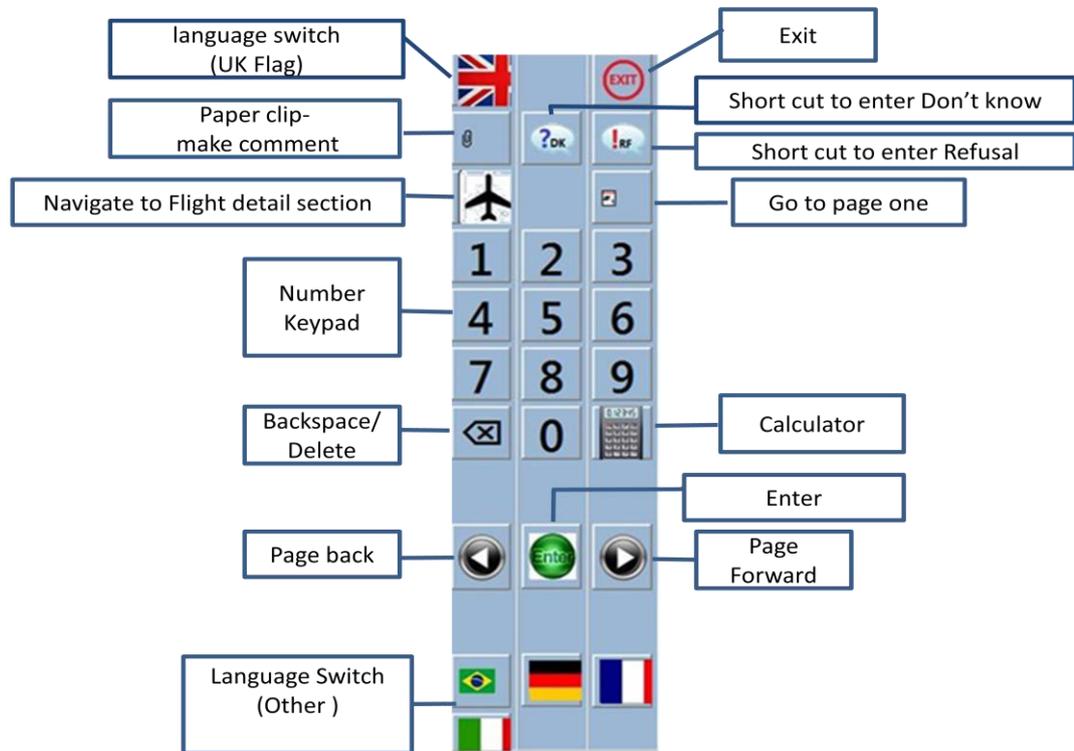
**Figure 3. Schematic Representation of Transmission Architecture**



### Sprint 2 – Can we design a user interface to enable data collection?

Using the interface developed for Blaise 5 was not an option as the features available in Blaise 5 are not available in Blaise 4. After some head scratching and asking other NSI's how they use Tablets we came up with the following design (see figure 4), a big thanks to CSO Ireland who sent us some example code. This navigation panel coupled with the on screen keyboard native to the tablet means the interviewers do not have to switch between numbers and alpha inputs during data entry and gives them the complete range of functions that they need to have to perform an interview. We have been able to make use of some of the design principles identified when we were designing the survey for Blaise 5 making use of Don Normans principles "The Design of Every Day Things" [6].

Figure 4. Navigation Panel

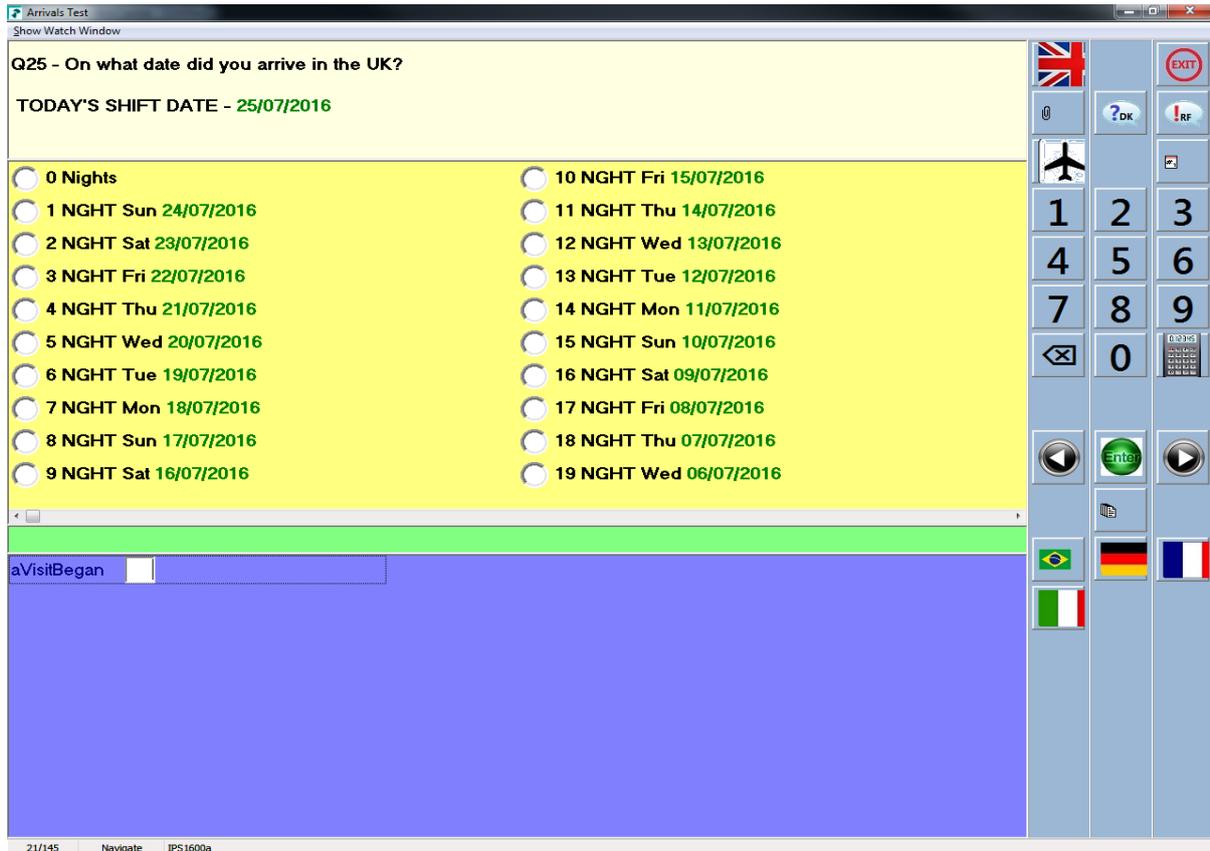


### Sprint 3 – Demonstrate user interface to interviewers

During this sprint we gathered a small group of interviewers together to review some of the challenges facing us during data collection: how to collect dates, how to collect information displayed on maps and how to best use lookups.

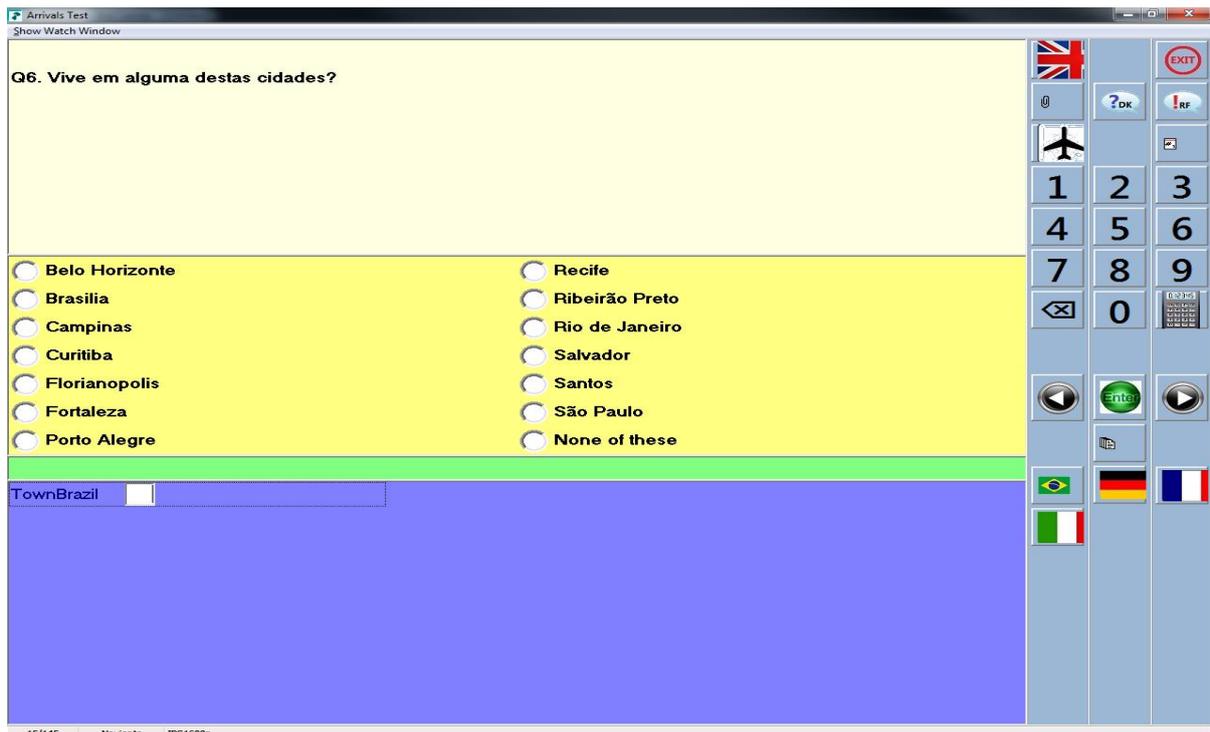
Collecting dates raised an interesting problem, we looked at implementing a date picker in Blaise 5 but this feature in Blaise 4 does not display large enough on a Tablet device to collect the information speedily and accurately. So what to do? We looked to increase the size but even with an upgrade of the software from CBS size remained a problem. We therefore looked at the problem from a different perspective. What data are we trying to collect and how to collect it in the most efficient way? So whilst a date picker is an industry standard, based on the data we collect we know most trips last less than 22 days, why use a date picker when you can visually present this data with radial buttons for easy selection (see figure 5). We also added in additional information to aid collection, such as the number of nights and the day of the week of the end date displayed in full to aid people's memories. If the date is out of range then we have the ability to collect this at another question where we collect the date in its most basic form of DD/MM/YYYY.

Figure 5. Date Information displayed as a List of Valid Dates with Associated Aid Memoir Information alongside the Data Items



The next issue to tackle was how to collect information currently displayed on a paper map about which city or region a respondent lives in. In Blaise 5 we looked to develop interactive maps with a complex network of grids with each cell designed as a selectable button. Whilst this was effective this functionality is not available in Blaise 4 and it would also carry a maintenance burden when new maps need to be added. We looked at displaying the maps in the screen size available but doing this made them unreadable for countries with lots of individual regions, for example Switzerland. Again we looked at what data it was that we were looking to collect from the respondents, city of residence or region of country. Given that respondents already know this information presenting a simple list with the answers and question in the native language solves the problem without the need for anything more complex (see figure 6). Sometimes the simplest idea is the best idea.

Figure 6. City Information displayed as a List, Question Text displayed in Portuguese



Finally we looked at the use of lookups. We use them extensively within the questionnaire to collect data on destinations, nationality, airlines and towns within the UK. It might seem obvious to display a list of the most commonly used codes to collect the desired information; however the survey collection environment is all about speed and accuracy. This method can lead to more pages having to be displayed to collect the information, so currently we are experimenting with both methods of collection one being a list (see Figure 7) followed by a lookup and the other being a straight lookup (see Figure 8).

Figure 7. Radial Buttons to Show List of Countries

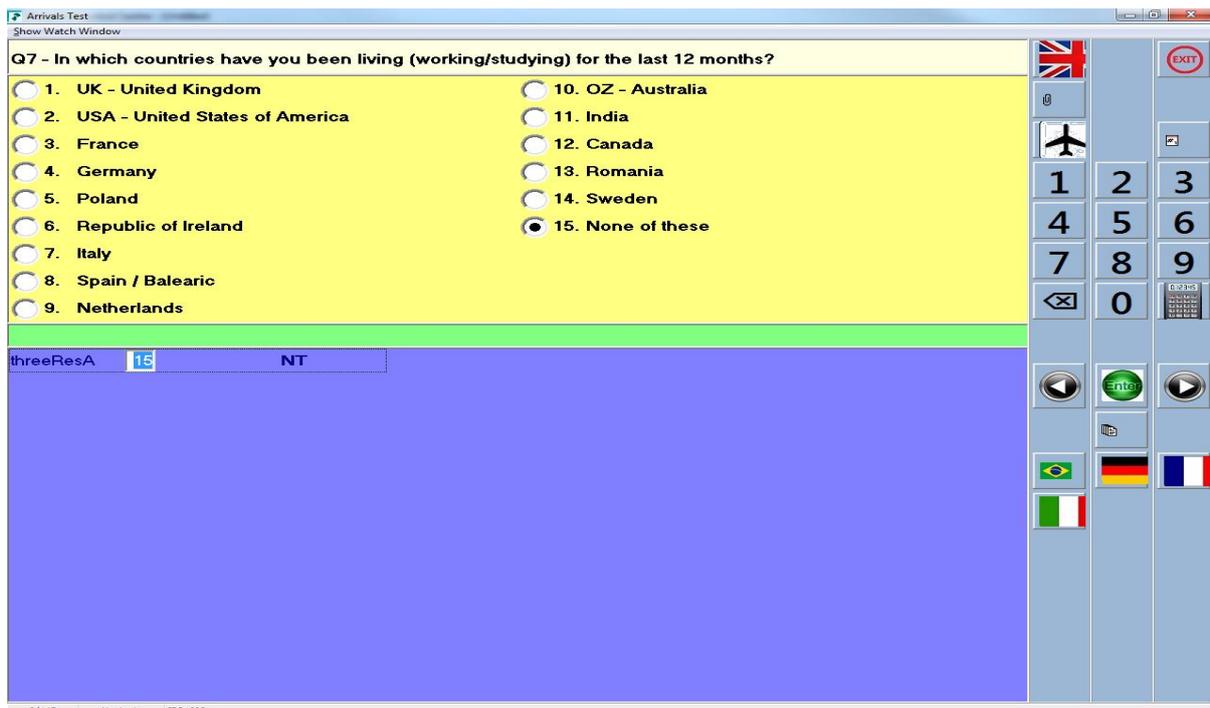
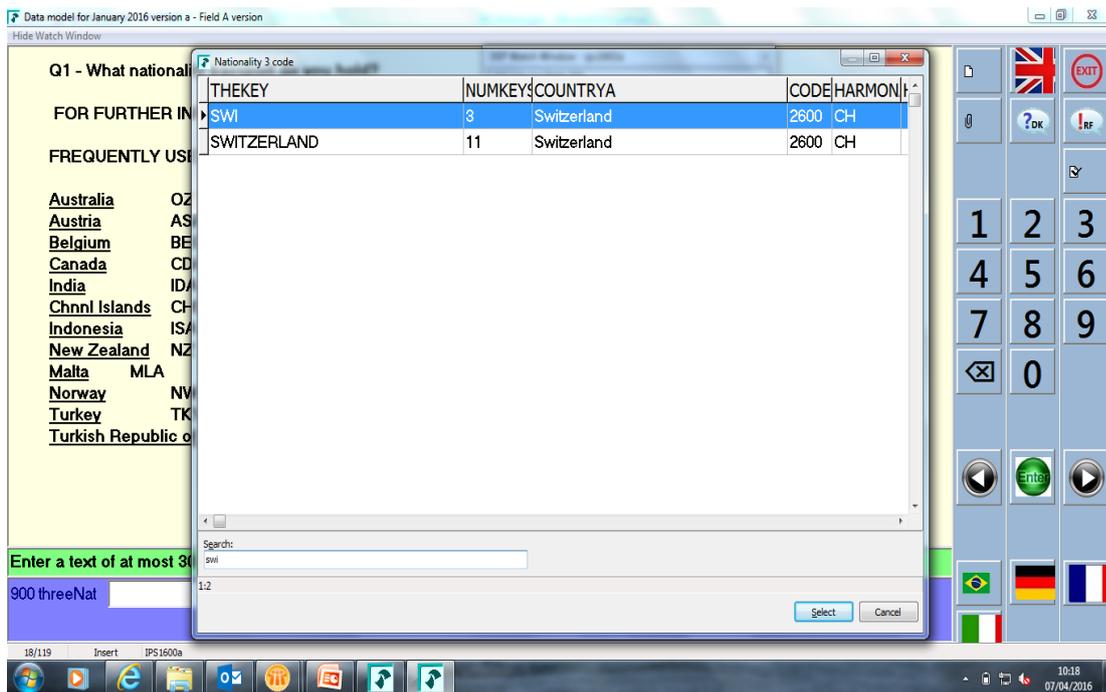
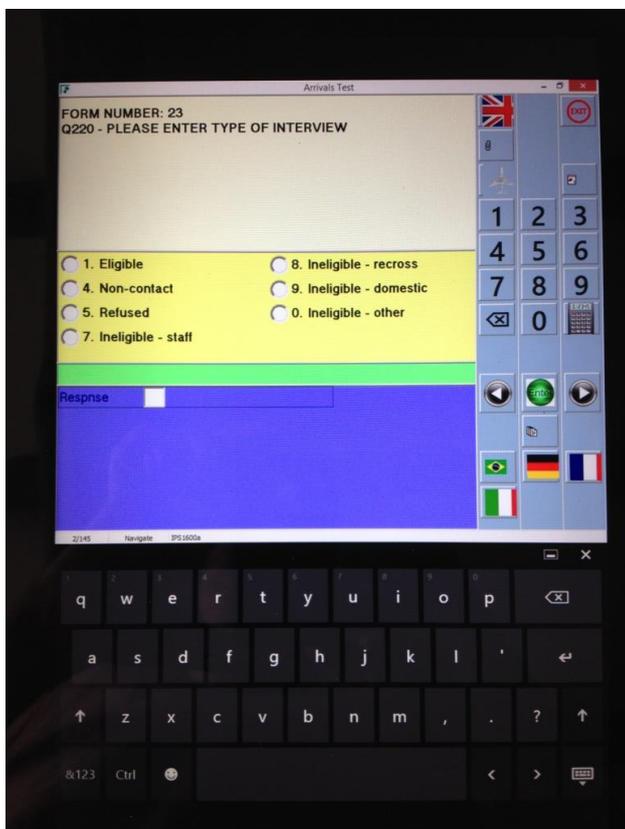


Figure 8. Entry to Lookup Coding Frame



Time will tell which is the best solution and it maybe that we develop lists for certain collection locations but we would like to develop a generic approach to reduce maintenance burden in the future.

Figure 9. Screen Shot of Navigation Panel, Native Keyboard, and Questionnaire



Having developed the solution to a point where we felt able to conduct an end to end interview now was the time to test our designs on the public and introduce a larger pool of interviewers to the solution. We conducted tests at the following sites:

Sprint 4 – Conduct Trial at Bristol to assess usability on a departure shift

Sprint 5 – Conduct Trial at Heathrow to assess flow on a departure shift

Sprint 6 – Conduct Trial at Manchester to assess flow on an arrivals shift

Sprint 7 – Conduct Trial at Portsmouth to see if location has any special requirements

Between each test we conducted reviews/retrospectives of what went well and what did not and made changes as and when required. The interface had some functions added, others removed and the interviewers made lots of suggestions about the questions themselves and certain sections were moved to aid navigation. The results from the tests have been very encouraging; no show stoppers to implementation have so far been identified. The key findings are summarised below:

### Positive

Ease of use - Tablet is easy to use and the experience was enjoyable

Speed of use - Time taken to complete comparable or quicker than paper collection

Professional - Use of electronic data collection device makes us look professional

Interaction - Increased engagement from the respondents

### Needs improvement

Navigation - Difficult to go back to certain sections

Weight - Potential strain on the hand and arm

Use of radio buttons - Buttons too close together

Questionnaire - Needs stream lining and adapting for new collection method

### Other suggestions / requirements

Peripherals - Need stylus pens to aid answer selection

Peripherals - Need some type of carrying device to reduce strain

## **6.1 Measuring Flow**

A key element of the IPS Survey is how we select respondents; this is by means of an imaginary counting line, every time a respondent crosses the line (as determined by collection location) a click is recorded on a clicker and after so many clicks (changes by location e.g. 1 in 20) an interviewer is tasked to interview the selected respondent. To prove that the use of the tablet would not affect the sampling interval i.e. could we keep up the pace of respondent selection in different environments we planned two tests, one at Heathrow on a departures shift and another at Manchester on an arrivals shift. Different shift types were selected for the test as the flow rate is different, on departures the flow of respondents tends to be steady whilst on arrivals then tends to be peaks and troughs.

The test at Heathrow produced mixed results in that we could easily keep up with the flow but this was in part down to the fact that the shift was unusually quiet. It did however throw up an issue with the navigation panel, in that the buttons changed places an error that was traced to having different versions of Blaise installed on the hardware (Dep.exe versus compiled Blaise code).

The test at Manchester was a lot more successful in measuring flow but we did take steps to avoid a repetition of the Heathrow test. Firstly we started the test when we knew we would have 4 or more planes arriving within 20 minutes of each other and we doubled the sampling rate from 1 in 20 to 1 in

10 to stress the test. The test was a great success and accounting for interviewers still getting familiar with the hardware, the new way of working and the increased sampling interval, results were comparable with normal shift conditions.

## **6.2 Increased Respondent Engagement**

One of the more interesting findings from the three live tests is how respondents are now engaging with the interviewers. Currently the interviewers have a clipboard on which to rest the paper questionnaire but due to the small typeface and the fact that the paper form is cluttered the respondents take little notice of the questions on the form and the interview is often conducted facing each other. On the Tablet however the respondents are engaged in the interview process from the start as they are able to read the questions themselves, often quicker than the interviewers. In some cases the respondents are almost self interviewing either by providing answers before the interviewer has finished asking the question or selecting the answers on the Tablet themselves. This has highlighted the redesign element of this project as what we display to the interviewer must also be clear to the respondent.

## **7. What next?**

We have developed a Maturity Roadmap which shows what functions/services we will be delivering out until September 2017, when we hope to have the full service delivered. It is based on AGILE principles where you deliver incrementally, at each stage there is an element of 'walk away ability', so if for some reason the project is stopped then you are able to demonstrate that some value has been delivered.

**Figure 10. IPS Tablets Roadmap**

The key activity we need to deliver by the end of this year is the methodological work that needs to be undertaken. This is required to gain an understanding of any discontinuity in the time series which

may occur as a result of switching collection modes and to understand if switching modes introduces any bias.

Other activities to be concluded by the end of year include:

Review of the IPS Questionnaire to ensure the questions and layout are enhanced for Tablet collection  
Identification of hardware – currently we are using Surface 3 Windows Tablets

Review of the coding, validation and edits – we are aiming to get the data as clean as possible at point of collection

## 8. References

[1]. Lan Benedkit. Developing the UK Passenger Survey in Blaise 5 on Tablet Computer, Office for National Statistics, UK, Statistics. IBUC 2015, 16th International Blaise Users Conference

[2]. User Guide (Volume 1) Background and Methodology. International Passenger Survey Overseas Travel and Tourism [http://www.ons.gov.uk/ons/guide-method/method-quality/speci\\_c/travel-andtransport-methodology/international-passenger-survey/index.html](http://www.ons.gov.uk/ons/guide-method/method-quality/speci_c/travel-andtransport-methodology/international-passenger-survey/index.html), 2014

[3]. User Guide (Volume 2) Background and Methodology. International Passenger Survey Overseas Travel and Tourism <http://www.ons.gov.uk/ons/guide-method/method-quality/specific/travel-andtransport-methodology/international-passenger-survey/index.html>, 2014

[4] The Agile Manifesto <http://www.agilemanifesto.org/principles.html>

[5] Fail Fast, Fail Often: How losing can help you win, Ryan Babineaux & John Krumboltz

[6] The Design of Every Day Things, Don Norman