Blaise 5 – Is Worth the Wait

Mark M Pierzchala, MMP Survey Services, LLC, United States

1. Abstract

There are 2 possible ways to understand the title of this paper.

- Blaise 5 – Is Worth the Wait?
- Blaise 5 – Is Worth the Wait!

In other words, is the title a question or is it a positive statement of fact? The reason to pose this question is that Blaise 5 has taken a long time to come into production. Additionally, many of the institutes that have tried Blaise 5 so far have struggled with its concepts. This is especially true for layout and devising templates that work across devices and screen sizes.

On the other hand, there are fundamentally important technical and questionnaire design issues with fielding governmental and scientific (*GovSci*) survey instruments in today's environment. The next section describes many of these challenges. Additionally, Blaise surveys are often lengthy and complex, and must be implemented in a methodologically sound manner. After a description of the challenges, the paper demonstrates how Blaise 5 addresses these challenges and how users can use the system optimally.

2. Fundamental Methodological and Technical Issues

MMP Survey Services, LLC (MMPSS) has worked with the Statistics Netherlands Blaise Team for several years to research, document, and validate Blaise 5 concepts. Investigations were also conducted into the types of questions and question structures that are found in *GovSci* surveys. Over the years, the research revealed how nine aspects of the survey environment impact electronic instrument design, especially screen design impact on question formulation. These are referred to as the 'nine multiples'.

The point of this section is to illustrate how modern-day challenges are so complex that (1) new instrument design and production methods are required, (2) no person has the knowledge to deal with all these aspects, and (3) an institute should plan ahead in order to know how to handle its survey world. The last point is explained further below. Some powerful solutions and approaches are also given.

2.1 The Nine Multiple Aspects of Computer-Assisted Instrument (CAI) Design

The list below enumerates the nine technological aspects that impact instrument design. The term *Multiple* derives from the prefix *Multi* as used in the following terms (in alphabetical order).

- **Multi-cultural**: surveys across human cultures
    - There may be different representations of the same underlying survey concepts between cultures. This may result in differently worded questions or response options, or even alternative structures for the same concept. These go beyond linguistic challenges.
- **Multi-device**: surveys on devices ranging from smart phones to tablets to computers
    - Differences in screen size, pixel density, whether there is a physical or virtual (pop-up) keyboard, and other differences between devices, can alter question display.

- **Multilingual**: surveys with two or more languages.
    - Translations must be timely and accurate and decisions have to be made about whether a user can change languages on the fly and how. Alphabetic, Asian, right-to-left, and left-to-right languages all must display properly.
- **Multimode**: surveys that are both interviewer- and self-administered (and variants)
    - Blaise 5 considers *Interviewing* and *Self-Administration* as the major mode division, but also formally accounts for variants such as CATI and CAPI (as interview modes) and web and paper (as self-administered modes). Question statement, instruction, and even question structure can differ between modes. One of the toughest issues is how to handle DK and RF options between modes.
- **Multi-national**: surveys in two or more countries
    - There can be country-specific attributes and references that have to be handled. These include currency, phone number formats, names of institutions and programs, languages to allow in each country, web links, terminology, and other items. Even same-language countries have to account for wording changes. For example, *surname* in the UK and *last name* in the United States refer to the name that comes after the first name. There is a Blaise 5 sample that includes three versions of the North American Industry Classification System (NAICS) as they are used in Canada, the United States, and Mexico.
- **Multi-operable**: surveys having to deal with different operational characteriscs such as touch- versus mouse/pointer- devices; virtual versus physical keyboards
    - The more complex the question or question structure, the more design is impacted by how the device is operated. For example, a battery of scale questions will be operated differently between a computer browser screen, a smart phone, and a Windows computer.
- **Multi-platform**: refers to surveys running as native applications on a device or running across the internet in a connected mode. For a browser, it can also refer to whether JavaScript or pop-ups are allowed.
    - The performance of the instrument, and when the rules are executed, may impact how many questions are placed on a screen, as an example.
- **Multi-structural**: surveys with questions that are structured differently for some manifestations.
    - This may refer to the practice where survey sponsors insist on different question structures or wording between modes. The researcher does this in order to achieve *cognitive equivalence* between modes and/or devices, even in situations where the structure can be rendered identically between modes. But there are also cases where a question structure in one mode cannot be achieved in another (see section 4.2 below).
- **Multi-version**: refers to surveys with many questionnaies that are similar, but have differences.
    - For example, kinds of industries or agricultural regions each require custom questionnaires. The questionnaire versions, and there can be hundreds, are all variations on a theme.

Many of the above design aspects are related, but each can impact design by itself.

2.2 Other Important Design Considerations

While the nine multiple aspects of CAI design mentioned above are tough enough to deal with, they do not account for all design challenges. In addition to the *nine multiples* above, other design considerations include:

- Paper questionnaires that never went away. These are still heavily used for some kinds of surveys, and have their own design considerations (mostly limitations). These can result in differences in question wording and construction between paper and electronic modes.

- Making electronic instruments usable to blind, visually impaired, and moter impaired users presents additional design challenges. This challenge of dealing with this issue can be under estimated. One reason is that assistive technology is used in conjunction with the Blaise instrument. Another reason is to truly accomodate the needs of the disabled rather then merely complying with laws such as check-listing against Section 508.
- Handling the following edit hierarchy can present additional design issues:
    - Univariate range edits (user types a number that is not within the field definition)
    - Univariate soft edit (a suspicious, but possibly valid value is entered)
    - Soft or hard edits between two or more fields on the same page.
    - Soft or hard edits between fields on different pages, including situations where the fields are physically located many pages apart.
- Legacy issues arise in surveys that have been fielded for many years. Survey sponsors may be reluctant to change how questions are asked because they do not want to disrupt data series.
- Mobile devices give new opportunities for data collection. By definition, some collections with smart phones and tablets are not even possible for other platforms.

2.3 The First Overall Goal

The first overall goal of Blaise 5 development is to be able to handle all of these aspects with one Blaise instrument, one database, and one well-conceived specification.

3. Blaise 5 Features give New Opportunities

This section gives only a hint of Blaise 5 features that are designed to handle the above challenges. Enough description is given here in order to describe below how to bring all these features together.

3.1 Layout Features and Modules

There are two design tools and three language features worth mentioning with respect to achieving layout. When used properly, and with adequate pre-planning, one reward is achieving instrument layouts that work across modes, devices, and screen sizes. A second reward is the large savings of time and effort in implementing desired layout.

3.1.1 Two Layout Design Tools

The **Resource Database** allows the institute to define fonts and styles (and many other aspects of layout) and to design display templates that are suited for each mode, device, and screen size. The elements of the Resource Database are available as design options in the Layout Designer. The Resource Database is a stand-alone module separate from the Control Centre.

The **Layout Designer** is found in the Control Centre. Here, design elements from the Resource Database are applied to the questionnaire. Screens for all modes and devices can be seen in the Layout Designer.

3.1.2 Three Language Layout Features

A display keyword called **GROUP** indicates a multi-field display. A group can be handled similarly or differently between modes, devices, and screen sizes.

The keyword **MODES** can be used in the source code itself and can (should be used to) map to Layout Set Groups in the Layout Designer. This makes it possible to marry source code to layout.

Naming conventions for types, groups, and blocks can be defined in the Resource Database templates and used in the source code. The result is the automatic application of layout elements; a possibility that is especially useful, and even necessary, for large instruments.

3.2 Programming Language Features

New keywords in Blaise 5 address several of the nine multiple aspects listed above. These include:

- MODES: In addition to mapping to Layout Set Groups, this keyword can be used in the Rules.
- LANGUAGES: This keyword is also present in Blaise 4, but in Blaise 5 it is used for natural spoken languages. Thus it is liberated from tasks that are now assigned to ROLES.
- ROLES can be used for tooltip display, screen reader text, variable names for downstream statistical packages and so forth. ROLES text can be defined in all the LANGUAGES.
- SPECIALANSWERS: This keyword allows users to assign additional non-value statuses to fields. These are in addition to the still-standard DK and RF. For example, it is now possible to define NA for 'Not Applicable'
- SPECIALANSWERSETS: This keyword allows users to allow different special answers for each mode. This formalizes a practice often seen in multimode surveys, especially the differential handling of DK and RF and EMPTY between interviewing and self-administration.
- ATTRIBUTES: Assigns the desired SPECIALANSWERSET for the instrument.

Listing 1. Model Source Code showing some New Language Features

```
DATAMODEL B5Complex "Blaise 5 Complex"

  LANGUAGES =
    ENG "English",
    NED "Nederlands"

  ROLES =
    Help "Help",
    ToolTip "Tool tip hint",
    ScReader "Screen reader text",
    StatVar "Stat variable name"

  MODES =
    Self "Self",
    Interview "Interview",
    Data "Data"

  SPECIALANSWERS =
    NA "Not applicable",
    OoR "Out of Range"

  SPECIALANSWERSETS
    EmptyAllowed = Self : EMPTY
                   Interview: DK, RF
                   Data : DK, RF, NA, OoR
  ATTRIBUTES =
    EmptyAllowed
```

In Listing 1, three MODES are recognized. These are *Self*, *Interview*, and *Data*. These are used respectively for self-administered mode, interviewing mode, and post-collection processing. The SPECIALANSWERS include *NA*, and *OoR*. These are in addition to the standard DK and RF.
The keyword SPECIALANSWERSETS allows differential handling of special values by mode. For example, for the *Self* mode, EMPTY is allowed for all fields by default. For *Interview*, NOEMPTY is the default value, but DK and RF are allowed for all fields by default. The *Data* mode allows by

default *NA* and *OoR* in additions to DK and RF. All these defaults can be overridden on the field level if necessary.

MODES, SPECIALANSWERS, and SPECIALANSWERSETS are all defined independently of device, screen size, and settings. They allow different handling of special answers between modes. Thus Blaise 5 formalizes survey taking as it has been practiced in governmental and scientific surveys.

3.3 Settings

Settings determine when rules are invoked, when special answers are displayed, how to enforce routing, and how to handle edits. That is, settings further adapt the operation of an instrument to different modes. Table 1 shows four model Settings, *Interview*, *Self*, *App*, and *Data*.

Table 1. Selected Settings for Setting Names Interview, Self, App, and Data

Setting Name	Letter	DK/RF & Special Answers Display	When Rules	Dynamic / Static Execution	Run Mode
Interview	A	Yes	Datum change	Dynamic	Thick client
Self	B	No	Page change	Dynamic	Client-server
App	C	No	Datum change	Dynamic	Thick client
Data	D	Yes	On demand[1/]	Static	Thick client

The settings shown above reflect typical uses. For example, *Interview* is intended for thick client instruments where rules execution is instantaneous. The rules are executed every time a data item is changed. The DK and RF are displayed for the interviewer.

On the other hand, *Self* is intended for a browser where the rules are executed remotely from the user. Here, it is appropriate to execute the rules on a request for new page. This often leads to fewer items per screen and it can slow down the overall instrument execution.

3.4 Other Configurable Aspects of Blaise 5

There are other aspects of modern survey taking that are accounted for by Blaise 5. Here is a short list.

- **Screen size:** Resource Sets in the Resource Database allow users to design templates for each screen size. These are applied in the Layout Designer.
- **Runtime parameters:** Allow users to match mode to layout set to settings to screen size.
- **Styles:** These allow users to easily change the look and feel of your instrument.
- **Screen layout:** Resource Sets in the Resource Database also allow users to decide between full-, split-, and entry- screen designs as shown in Figures 1 through 3.

Figure 1. Full-Screen Display for Self-Administration on a Web Browser on a Computer

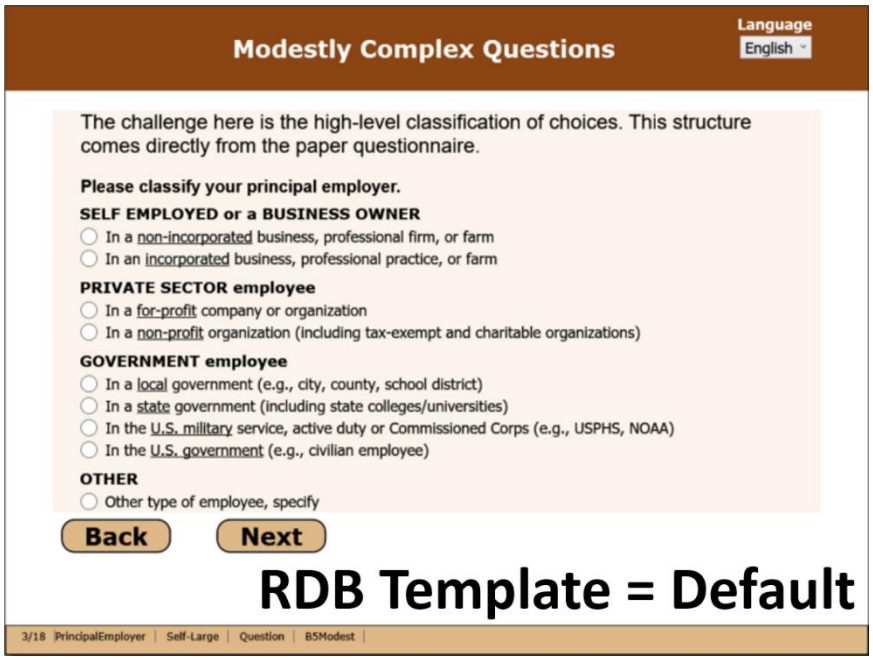


Figure 2. Split-Screen Display for CATI in Windows for a Computer Desktop or Laptop

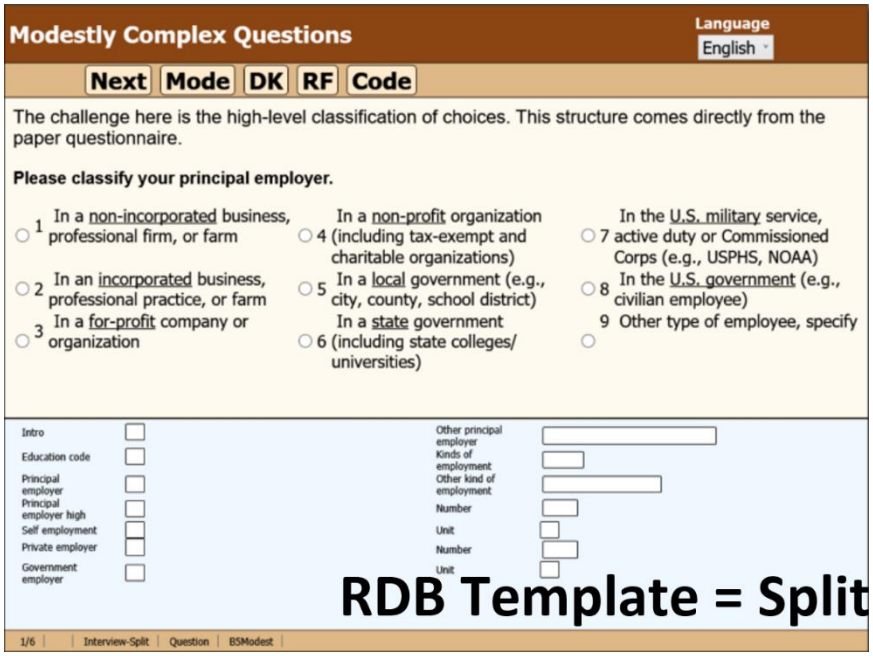
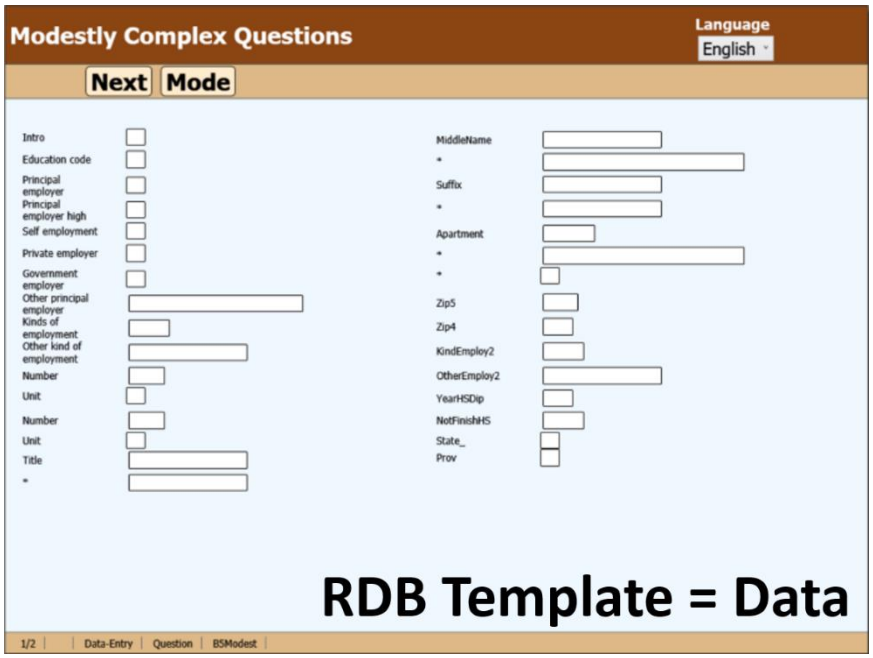


Figure 3. Form-Pane Display for Post-Collection Processing



3.5 A Second Overall Goal

A second overall goal is to have the questionnaire designer focus on questions (Figure 4). This person should not worry about methodological and technical details; it is just too complicated and overwhelming. These details should already have been decided by the institute as described below.

Figure 4. One Specification handles Modes, Devices, and Screen Sizes

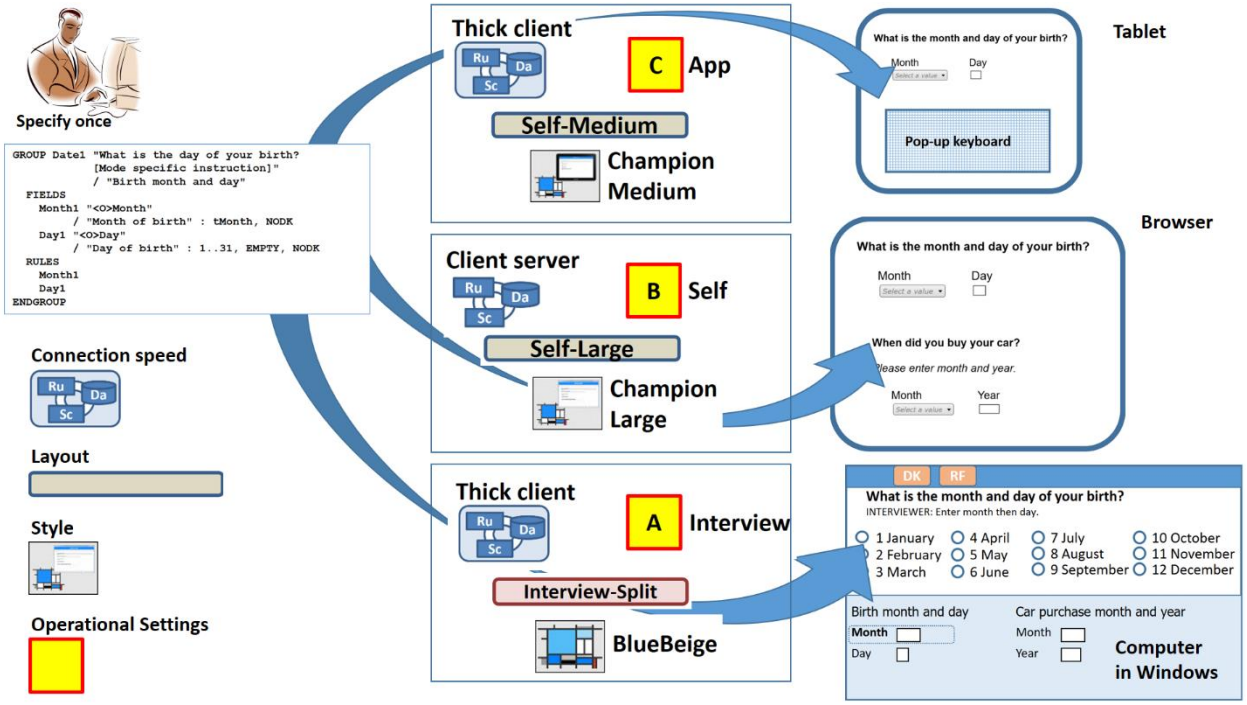
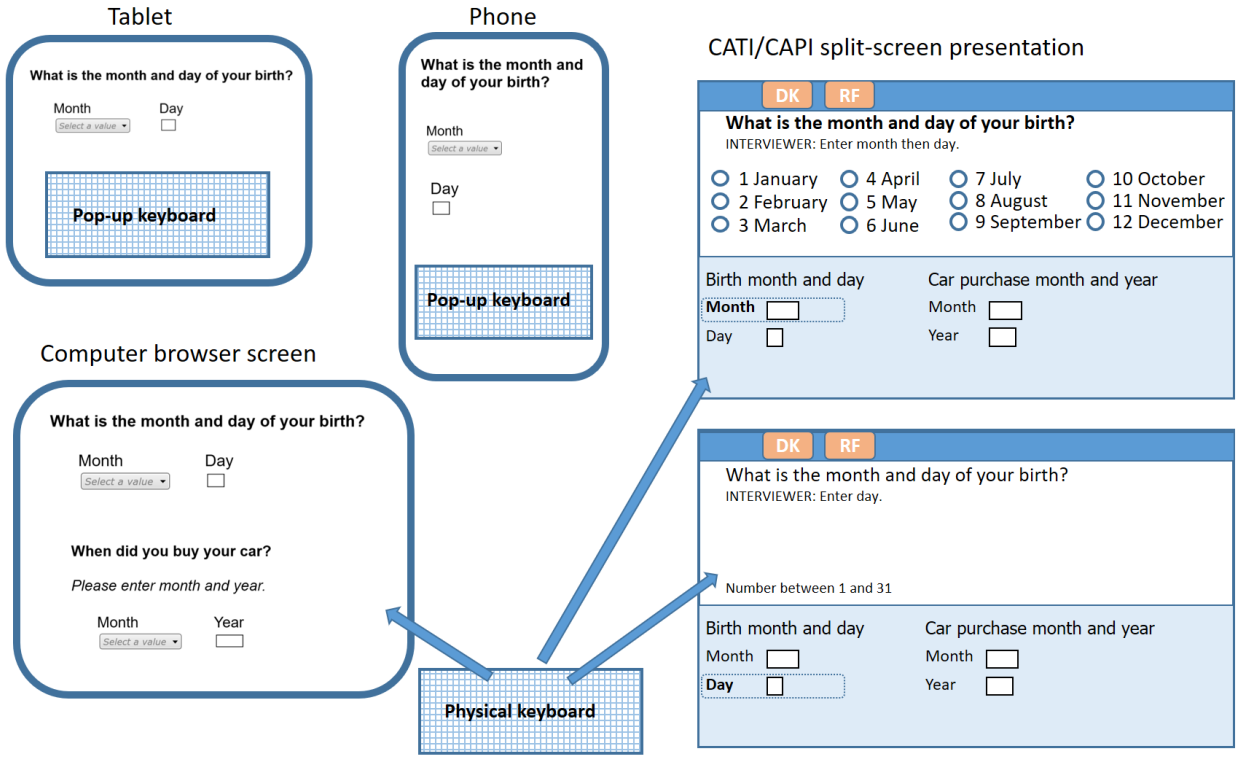


Figure 5 shows how a Quantity/Unit question (two fields) might appear on a tablet, a smart phone, a computer browser screen, and in the traditional Blaise 4 split-screen presentation. Note that the computer browser screen shows two Quantity/Unit pairs whereas the tablet of a similar screen size shows just one pair. This is because for the tablet, it might be necessary to reserve room for the pop-up keyboard.

In any case, the goal is to have all these displays (and more) appear, as if by magic, by one well-conceived specification, written by an individual who is concerned with just the questions.

Figure 5. A Quantity/Unit Pair displayed different Ways on different Devices



4. Institutes should Establish Layout and Operability Standards

This section shows how an institute can establish layout and operability standards for multi-device and multi-mode production of instruments. As explained above, there are so many details to think about that the institute should convene a standards-setting group instead of making up ad hoc survey standards.

By design, Blaise 5 has the capacity to pre-package these agreed institute standards. Layout standards are encoded in a standard Resource Database. Other standards are held as model source code, model settings, and model layout sets and groups. Additionally, the institute should establish type, block, and group naming conventions, that when used in the source code, automatically generate layouts.

4.1 Design Assertion

The methods described below rely on the following design assertion:

> "The design of layout across modes and devices depends only on question format and structure."

This means that it is possible to conduct all necessary design and methodological research through investigations of question format and question structures. All resulting institute findings and standards are thus independent of subject matter, which is a strong asset.

4.2 Unimode versus Generalized Mode Design

In multimode surveys, there are two design traditions. The first is called *Unimode* design. This tradition holds that question statement and structure should be presented the same across all modes. This might mean that the question is presented less than optimally in one or more modes. The benefit is that the measurement of the concept is as same as possible across modes.

The second tradition is called *Generalized Mode* design. In this tradition, the aim is to achieve cognitive equivalence across modes. This often means that the question structure is presented differently for one or more modes in order to optimize the presentation for each mode.

4.2.1 Discussion of Design Traditions

A good description of *unimode* versus *generalized mode* design is given by de Leeuw (2004). Pierzchala et al. (2004) give an example of a survey that used generalized mode design for a multimode CATI, Web, and paper survey. For some questions there were three question structures across the three modes. Pierzchala (2006) concludes that as the number of modes goes up in a survey, the harder it is to maintain a unimode design. This paper defines the concept of *degrees of disparity* between modes on several design criteria. However, a recent volume by Dillman, et al. (2014), advocates for unimode design wherever possible, even in surveys mixing paper, CATI interviewing, and Web.

Most of the original discussion took place before smart phones and tablets emerged as popular survey devices. As Blaise 5 was developed, it became clearer that the nine multiple aspects of CAI Instrument Design (section 2.1 above) make it harder to consistently achieve unimode design. The strategies and techniques described (at a high level) in this paper allow the institute to methodically investigate these issues, and where generalized design is necessary, to achieve the necessary displays across devices.

4.2.2 Impact of Difficult Questions on Instrument Design

Blaise is used for the world's most difficult governmental and scientific surveys. These surveys often feature lengthy and complex questions. They also incorporate difficult question structures such as tables. Our research shows that lengthier and more complex questions also encourage generalized mode design.

In order to more fully understand the impact of question structure on instrument design, MMPSS put together a suite of instruments called the Blaise 5 Champion Instruments. These are described below and in Pierzchala (2016). These instruments hold many different question types and structures and are used to validate Blaise 5 layout. They are available for Blaise institutes to use in their own investigations.

4.3 Institute Display and Operability Standards

Blaise institutes often have display and operability standards for their survey instruments. As institutes adapt their questionnaires to the modern survey-taking world, it may be time to re-assess these standards. A few standards used in the production of the Blaise 5 Champion Instruments are given here.

4.3.1 Web Screen Display Standards

A few of the many web display standards are described below. These are meant for a browser on a computer. One aspect of a web-based instrument is that the rules are executed on a remote server and there can be a delay whenever rules are executed.

- Use a full-screen display (Figure 1 above).
- The base font is Arial.
- Question text is bold.
- Instruction text is unbolded italics.
- Choice text (for enumerated fields) is unbolded and not in italics.
- DK and RF are not displayed when the respondent first enters the screen.
- DF and RF are displayed if the respondent leaves a field empty and tries to go to the next screen.
- If one question is used as a condition to route a second question, the second question is placed on the next screen.
- When routing is not an issue, the use of multi-item displays is encouraged to speed up the survey.
- Horizontal scrolling is not allowed. Vertical scrolling is permitted but not liked.
- Edits involving 2 or more fields should be used only if all involved fields are on the same screen.

4.3.2 Interviewer Screen Display Standards

A few of the interviewer display standards are described below. These are meant for an interviewing instrument using MS Windows®. Here, the display and the rules reside on the same computer and rules execution is instantaneous.

- Use a split-screen display (Figure 2 above).
- The base font is Arial.
- Question text is in mixed case bold.
- Instruction text is mixed case, unbolded and is preceded with the text 'INTERVIEWER:'.
- Optional text is mixed case unbolded text.
- Choice text that is to be read is mixed case, unbolded.
- Choice text that is not to be read is upper case, unbolded.
- DK and RF choices are always available where allowed. The interviewer does not offer them, but can use them if necessary.
- Question routing has no impact on page breaks.
- High data density is encouraged in order to facilitate navigation.
- All edits are available to be used regardless of the location of the involved fields.

4.3.3 Discussion of Web versus Interviewing Display Standards

The standards above are illustrative of those used by many Blaise institutes, even if the details differ. On one hand, an institute should try to align display standards across modes. On the other hand, experience teaches that this is not always possible, or desirable. There really is a difference between aural versus visual presentation of questions. There is also a profound difference between the untrained one-time self-respondent and a trained interviewer who may conduct hundreds of interviews on the same survey.

Once the institute has re-assessed its presentation standards, its specification protocols and tools should easily be able to handle these differences.

4.4 Three Layout Design Strategies

One of the goals of Blaise 5 has been to establish model design standards. These are held in the default Resource Database and as model source code, settings, and layout sets and groups. These resources will also be used in the Blaise 5 Champion Instruments. Given these resources, there are three layout design strategies for a Blaise 5 Institute.

- Use Blaise 5 out-of-the-box defaults. This is easy.
- Modify Blaise 5 defaults to create your own standards. This is moderately easy.
- Start from scratch. This is difficult and time consuming.

Deciding on a strategy to use depends on many factors. A good first step is to assess how these standards appear when used in the Blaise 5 Champion Instruments.

4.5 Blaise 5 Champion Instruments

The next step is to establish a test bed of different question formats and question structures. To this end, ten so-called Blaise 5 Champion instruments have been established and will be part of the Blaise 5 distribution. These are described below.

The ten *Champion* instruments demonstrate a range of display capabilities. Six instruments demonstrate specific layout capabilities and four are survey instruments.

4.5.1 Champion Instruments that Demonstrate Layout Capabilities

These six instruments are called B5 instruments.

- **B5Basic**: basic question layouts
- **B5Modest**: modestly complex questions and structures
- **B5Complex**: group questions such as tables
- **B5Scales**: a variety of scale types and display possibilities
- **B5PanEuropean**: multicultural aspects of surveys with over twenty languages
- **B5CodeFrame**: Coding kinds of questions (Windows for interviewing only)

These B5 instruments demonstrate about 100 kinds of items and data-collection structures. These questions, screens, and structures come primarily from government and scientific surveys.

4.5.2 Champion Survey Instruments

Two of the four instruments are based on real surveys (Census and ASI). The other two are made-up but realistic surveys. Two of the survey instruments were programmed from scratch in Blaise 5 (Retail Trade and Census). The other two (NCSPerson and ASI) were Blaise 4 surveys converted to Blaise 5.

These two instruments were initially programmed in Blaise 5 using the latest multimode features.

- **Retail Trade Survey**: collects financial data from firms and features non-linear page navigation. This instrument is meant for PC browsers or tablets.
- **Census-type Listing Instrument**: This instrument is meant for self-completion on device or browser and for interviewing on a device such as a smart phone or tablet.

These two were converted from Blaise 4 which necessitated some code restructuring and broadening of perspective in order to operate on new platforms.

- **NCSPerson**: a self-completion instrument meant to be used on PC Browser and devices.
- **Annual Survey of Industry (ASI)**: a complex CATI and PC Browser instrument that collects economic data from firms. There are differences between web and CATI display, edit handling, and wording. It features the use of metadata statements in external files in order to drive the instrument differently for each industry.

4.5.3 Use of the Champion Instruments by your Institute

Institutes should inspect the Champion Instruments to see if they represent the entire range of question types and structures used in its survey program. If there are additional question structures that are used, these can be added to the Champion Instruments. The goal is to establish a ready-to-used test suite of instruments to assess display and operability.

5. Improved Specification and Instrument Generation

In this conference there is a presentation by Colectica (2016), a company that is producing a Blaise generator. This system has a specification interface and upon pushing a button, a Blaise instrument can be produced. The generation of Blaise instruments has been demonstrated several times in the seventeen International Blaise Users Conferences. The first was by Pierzchala in 1992 when he demonstrated how the National Agricultural Statistics Service (NASS) could generate over forty versions of the same instrument. It could be, that in order to more easily handle the modern survey world, generating Blaise questionnaires may become more common or even necessary.

Regardless of whether a Blaise generator is used, or all source code is typed into the text editor, the Blaise institute should have considered all its design practices across modes. It is easy for a Blaise programmer to follow standards. It is very difficult to invent these on the fly.

6. Summary

The modern survey world forces instrument designers and developers to re-assess their practices and ways of working. It is such a complex world that no one person has all the answers. Best results for an institute is to plan ahead and standardize its way of handling the nine multiple aspects of CAI design.

Blaise 5, for its part, allows users to encode these standards in the Resource Database and in model source code, settings, and layout sets and groups. Then it is possible to easily bring together all these standards for each device, screen size and mode. The questionnaire designer should focus on the questions. With proper planning and preparation, it should be possible to easily produce instruments that (1) work across all devices, screens sizes, and modes, and (2) be methodologically defensible.

To finally answer the question posed in the abstract: **Blaise 5 – Is Worth the Wait!**

The new features, layout possibilities, language features, and settings, can be combined in any way that is necessary. However, the institute will yield much larger benefit if it re-evaluates its standards and practices as described above then encode its standards and practices in the Resource Database, model source code, model layout set definitions, and model settings. The Blaise 5 Champion instruments give the institute a good test bed to test out these encoded standards. At the same time, Blaise 5 gives you the flexibility to adapt these standards to new situations as they arise.

7. References

De Leeuw, E. D. (2005). To Mix or Not to Mix Data Collection Modes in Surveys. Journal of Official Statistics, 21, 2, 233-255.

Dillman, D. A., Smyth, J. D., and Christian, L. M. (2014). Internet, Phone, Mail, and Mixed-Mode Surveys: The Tailored Design Method, 4th Edition. New York: John Wiley & Sons.

Iverson, J. and Smith D. (2016). Blaise Colectica Visual Survey Designer. Presentation at the 2016 International Blaise Users Conference, The Hague, The Netherlands, 2016.

Pierzchala, M. (2016). The Blaise 5 Champion Instruments Suite. Poster Session presented at the 2016 International Blaise Users Conference, The Hague, The Netherlands, 2016.

Pierzchala, M. (2006). Disparate Modes and Their Effect on Instrument Design. Paper presented at the 2006 International Blaise Users Conference, Papendal, The Netherlands, 2006 (at http://www.blaiseusers.org/2006/Papers/207.pdf).

Pierzchala, M., Wright, D., Wilson, C., Guerino, P. (2004). Instrument Design for a Blaise Multi-Mode web, CATI, and Paper Survey. Paper presented at the 2004 International Blaise User's Conference, Gatineau, Canada, September, 2004 (at http://www.blaiseusers.org/2004/papers/24.pdf).

Pierzchala, M. (1992). Generating Multiple Versions of Questionnaires. Paper presented at the 1992 International Blaise Users Conference, Voorburg, The Netherlands, 1992 (at http://www.blaiseusers.org/1992/papers/GENERATING%20MULTIPLE%20VERSIONS%20OF%20QUESTIONNAIRES.pdf).