Use of Blaise in Today's Survey Research Environment

Jane Shepherd and Ray Snowden, Westat
International Blaise Users Conference, April 2012, London, UK

Summary: *Blaise is used in a variety of survey research environments by Westat and many other organizations. Surveys typically involve multi-mode data collections and organizations are required to use commercial off-the-shelf (COTS), integrate multiple platforms and applications to solve complex problems, and assure the security and confidentiality of the respondents and collected data. Software selection is a key issue as organizations have to invest efforts in training staff and supporting and managing software and systems infrastructures over time, while offering cost-effective solutions. These factors are discussed in relationship to the use of Blaise in this evolving environment.*

Dynamic Survey Research Environment

One of the major challenges for software developers today is the evolving and dynamic nature of the survey research environment. In recent years there has been a continuing increase in the number of data collection modes and the types of activities and measures being incorporated into these data collections. In conjunction with these needs survey researchers have seen increased pressures for real time reporting and integration across modes for the exchange of data and performance of quality control tasks.

Some of our most common modes include CAPI, CATI, CAWI, PAPI, CASI/ACASI, and IVR. These modes are often inter-mixed with other presentation media such as brochures, mail outs, show cards, permission forms or other materials, and the data collection software may need to interact across platforms and operating systems. The need to support integration with other types of software, for example GPS or biomedical components, is increasing. The role of the interviewer in administering these various modes of data collection may vary from the traditional model. Interviewers may be required to do more or to change roles during the data collection. The completion of the survey may be interviewer mediated or respondent mediated or a combination of both. For example, the interviewer may arrive at the household and do the initial contact and consent task, and then move the data collection into a CASI, IVR or other mode. In other situations a hard copy questionnaire or IVR may be used to initially solicit the respondent's participation, and then the data collection mode may switch to some type of interviewer mediated mode. Survey designers are taking advantage of the multiple available modes to maximize respondent compliance with the data collection and offer respondents the option to complete the data collection at the time of their choice and in the mode of their choice.

In addition to the complexities described above, the administration of the multiple data collection modes may be sequential, concurrent, or variable (for example, changing from sequential to concurrent at some point in time). Adaptive designs that take advantage of feedback during the data collection can alter the operational plan for the survey. Paradata captured during survey administration and survey results including refusals and item nonresponse rates provide valuable feedback to researchers looking to optimize the use of resources and maximize response rates.

In short, the survey research environment is very complex and changing rapidly by incorporating new technology and techniques to improve the quality of research data, improve response rates, reach more respondents, and lower costs. No single product can meet all of these emerging needs. In addition to a strong set of core survey support features, survey research organizations require software with other key characteristics that allow products and tools to be adapted, extended, and integrated in creative new solutions. We define some of these characteristics in this paper, the manner in which Blaise reflects many of these characteristics as a leading survey research tool, and examples of how Westat has integrated Blaise capabilities in numerous custom applications.

Key Characteristics of Survey Research Software

Data collection software has to meet the needs of the current multimode landscape and be flexible enough to adapt to future challenges. The following characteristics represent both the core capabilities required in survey research software plus those characteristics that permit managers and developers to creatively extend and adapt the software for new applications:

- Fulfills core survey functions - the tool allows the user to author and execute survey instruments that support a wide range of question types, complex data structure such as rosters and loops, and complex routing while maintaining data integrity through the interviewing process.

  Blaise is perhaps the leading product of its type with respect to these core features. The Blaise data model records the metadata that defines questions, responses, and flow and is interpreted by the Blaise rules engine. These two components are central to the Blaise solution and ensure a high level of functionality and consistency in the collection and interpretation of data collected using any of the Blaise supported modes, platforms, or tools.

- Provides a rich set of features, tools, and utilities – this includes tools to author and customize instruments, deploy and execute instruments in different operational and technical scenarios, and access and manipulate data.

  Blaise provides a wide variety of features, tools, and utilities including developer tools and documentation, multimedia support, multi-language support, customizable layout, ancillary programming languages such as Manipula, Maniplus, and Basil for data manipulation and interface control, a CATI management system, audit trails, support for data entry verification, and integrated user help. These features and tools all build on the core datamodel and rules engine of Blaise and provide a wide variety of options to creatively and efficiently extend the core capabilities of the product and allow Blaise-based functions to be used throughout the survey life cycle.

- Reliable performance and scalability – in order to maximize response rate and interviewer efficiency it is critical that a large complex survey performs well in a single user environment (e.g., CAPI) and that a survey intended for a large audience functions well in a multi-user environment (e.g., CATI or CAWI).

  Blaise has a long history of reliable performance in a wide variety of large and complex surveys. The Blaise datamodel and rules engine has a huge capacity for large numbers of questions, edits, and hierarchies and can apply these rules to responses with sub second response time. The Blaise architecture has evolved to a more flexible multi-tier architecture that can support large concurrent population of respondents.

- Integration with other technology, products, and applications – no single product can provide all of the support required for a large scale, end-to-end, survey operation. The ability to interoperate and share data with other applications and platforms, often in real-time, is an essential characteristic to permit survey research organizations to build complex and unique solutions.

  Blaise provides numerous features and mechanisms for integrating with other applications including the use of RDBMS for data storage, APIs to invoke Blaise functionality, and the use of Active X components and DLLs to extend Blaise functionality.

- Supports a secure data collection environment – given that confidential data is frequently collected in survey research projects, organizations may fall under numerous legal and regulatory mandates including FISMA, HIPAA, FDCC, and others relating to information security and data confidentiality. This is a complex and growing area of concern and can consume significant time and resources through the survey life cycle.

  Blaise provides numerous features and capabilities that directly support or are compatible with many of the controls required to meet these requirements (Rhoads and Snowden, 2010).

- Availability of support – when undertaking large and complex survey that may be active for years, it is important that the technology and tools used to support the operation are supported and will be supported in order to adapt to new technology, build in new features, fix bugs, train staff, answer questions, and solve problems. Support of this type is particularly important for a product as complex and extensive as Blaise.

  There are a wide variety of support mechanisms in place for Blaise including technical support for the core product by Statistics Netherlands, the International Blaise Users Group (IBUC), the Blaise Corporate Users Board (BCLUB), and 3rd party support providers such as Westat.

Blaise in the Survey Research Environment

Blaise is a core system used at Westat throughout the survey lifecycle. In addition to the core features and capabilities listed above, Westat has extended Blaise and integrated Blaise with various applications and technologies to provide full-service support for a wide variety of survey modes and scenarios and to support the full life cycle of the survey life cycle. Some of these extensions are shown in the following diagram and described below:

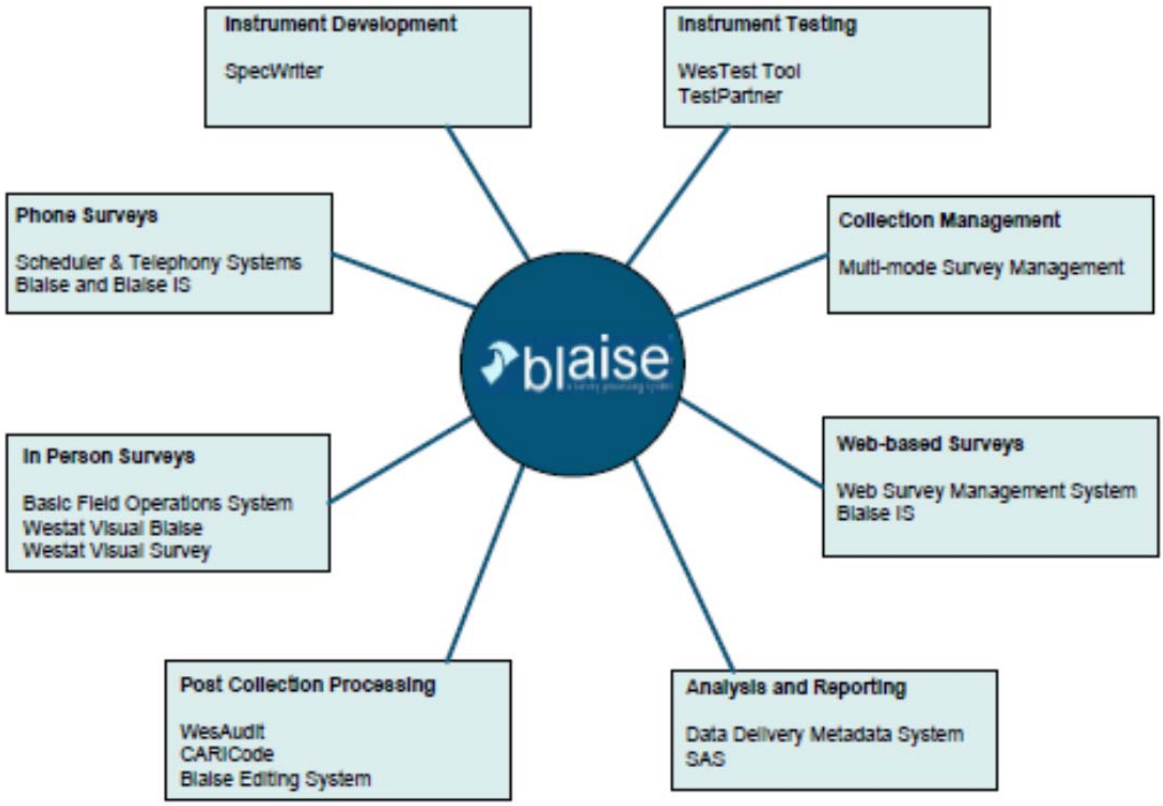

Figure 1. Survey instrument life cycle: Blaise and Westat tools, extensions and applications

Instrument development – the core Blaise product includes a full set instrument development suite that permits developers to specify the data model and customize the look and feel of the Blaise survey. Westat has modified a custom specification authoring tool to generate output that is fully compatible with the Blaise authoring process.

- Specification development and maintenance – A SpecWriter tool is used at the specification stage in the development life cycle. It provides designers with dynamic templates for considering and specifying all required elements for key item types, where items are defined as questions, boxes, or edits. SpecWriter allows hard-copy generation of the specifications and generates Blaise block templates that Blaise programmers use to create a Blaise instrument (Gowen and Clark, 2007). The output generated from this tool is customizable and contains the detailed specifications for the instrument in terms of question text, response text and values, questionnaire flow, and whether interviewer help is available for the question. SpecWriter creates HTML files, Blaise template code, and detailed specifications used by coders. It also streamlines the process of documenting and updating instrument specifications. Testers can use detailed specifications to compare data entry screens and functions against requirements and expected results. It increases efficiency by automating some of the questionnaire programming and helps reduce errors in creating and implementing the instrument.

Instrument testing - the testing of Blaise instruments is a crucial part of the lifecycle to insure that the instrument will collect and store data correctly and that the instrument flow properly reflects the survey design. This is particularly important with a large and complex survey instrument where the number of test scenarios required to reflect the numerous navigational paths through which the instrument may pass can grow quite large. Automated testing support tools are absolutely essential to

permit comprehensive testing on a reasonable schedule and cost. Westat has used the following COTS and custom tools in the testing of Blaise instruments:

- Testpartner is a COTS testing product that can be used to automate testing for Blaise Internet projects including Westat Visual Survey. Automated testing provides a more cost effective way of conducting effective regression testing and eliminates much of the manual testing work (Brenner and Manoj, 2010). Repeatable scripts can be executed against code after updates are made to assure that the data are being captured correctly.

- The WesTest tool has made problem reporting during testing easier by implementing an automated reporting system that is integrated with a COTS change management system. When a problem is discovered during execution of a Blaise instrument, the user can invoke an executable that produces a dialog box that the user completes to report a problem. The user need only enter a summary, description, and "type" of the problem. Items such as date, field tag, primary key, field value, instrument version, etc., are automatically recorded.

Collection Management – systems to manage the overall data collection, particularly in multi-mode studies, are essential. These systems are sometimes provided within products, but more often the user needs to provide a layer of integration across independent, mode-based platforms and applications. Westat has integrated Blaise with a multi-mode management system to support these needs (Frey, et. al. 2012).

Phone Surveys – Blaise is used in a wide variety of CATI surveys at Westat. Westat has integrated Blaise with custom CATI management software which permits Blaise to be integrated with the existing Westat CATI operational model. In order to provide web-based access to Blaise instruments for our remote CATI interviewer workforce, Westat uses two different technical approaches.

- Blaise integration with Westat CATI Scheduler and Telephony Systems -- Westat has implemented Blaise for CATI instrumentation and interviewing under a custom CATI Scheduler through the use of our Blaise Mainline software, which provides the link between Blaise CATI instruments and the CATI Scheduler. Operationally, the interviewer sees only Blaise screens, including screens for receiving or requesting cases, viewing interview management information about a case, setting status codes, etc. The integration of these Blaise screens with the CATI Scheduler is transparent to the interviewer and also to the supervisors, who are provided with Westat-created Blaise utilities for examining cases combined with their respective scheduler information, all within a Blaise user interface. Another critically important system for CATI surveys is the telephony system that supports all telephone operations in coordination with interviewer workstations. Westat has adopted a commercial computer-telephony integration (CTI) system for use in telephone interviewing. Through the Blaise Mainline, Westat has integrated this CTI system with the interviewer workstations.

- Browser-based access to Blaise – Westat provides browser-based access to Blaise instruments for our remote CATI staff through Blaise for Windows launched within a Citrix environment and through the use of Blaise IS. These two options provide flexibility in terms of platforms, performance, integration, and end-user functionality. This flexibility permits us to choose the environment that best meets the needs of a particular project.

In Person Surveys – Blaise is also used in a wide variety of in person surveys including CAPI, CARI, and ACASI. Westat has integrated Blaise with custom CAPI management software which permits Blaise to be integrated with the existing Westat CAPI operational model (Hill, 2004; O'Reagan et.al, 2010). In addition, Westat has developed several significant functional extensions using the Blaise interoperability features to meet complex and unique needs.

- Blaise integration with Westat CAPI management system-- Westat implements CAPI studies with our Basic Field Operations System (BFOS), using a web server and a web browser on field staff computers. Westat's BFOS system is designed for use with Blaise, and all management information coordination between Blaise instruments and the BFOS management system is handled through a software component known as the BFOS Mainline. This Mainline software resides on field laptop computers and provides the interface for field staff to review, select, update, and start Blaise instruments for all cases on the laptop computer. BFOS also provides functions to synchronize data between the laptop and the home office.

- Westat Visual Blaise (WVB) is a Blaise-based CASI software system that includes features for self-interviewing, with audio (ACASI) and video file incorporation. WVB has a user-friendly screen design suitable for touchscreen operation by survey respondents.

- Westat Visual Survey (WVS) takes advantage of the underlying technical architecture that captures the benefits of the Blaise open architecture and uses the Blaise rules engine to incorporate the features of Blaise, while writing the data to a COTS database such as SQL server. The systems displays an enhanced user interface via an Internet browser. This architecture results in a multimode product, allowing WVS to run on an individual laptop computer, networked workstation, or via the web. Using this architecture facilitates more complex data collection by allowing a virtually unlimited number of records to be stored interactively at any database level during program execution. WVS provides better programmer control of instrument navigation, enabling the system to fulfill more complex navigation and data flow requirements, by allowing the developer to override the Blaise route and force the application to go to a specific point in the instrument. This also provides for more robust functionality by incorporating summary screens and jumpback navigation screens and provides for better navigation options for interviewers because of enhanced programmer control over what appears on the screen. These features allow developers to program navigation menus in more complex instruments (Segel and O'Reagan, 2012).

Web-based Surveys – CAWI has become an increasingly important mode as a cost effective means of reaching a large and distributed population of respondents. Westat has used Blaise for instrumentation support on web-based surveys particularly in cases where multi-mode support is required and the same instrument can be deployed for multiple uses.

- Blaise integration with Westat custom web survey management systems – Westat implements web surveys with our management system that provides support for respondent authentication, survey task workflow, and instrument integration.

- Use of Blaise IS – with the changes to the architecture of Blaise IS, the product now can provide a level of performance with increased concurrent usage that is required to support complex web-based surveys with a medium to large respondent population.

Post Collection Processing – there are numerous back-end processes through which instrument data may pass before it is ready for final analysis and reporting. In many cases, the ability to utilize the Blaise data model and rules engine to help guide and validate these efforts is crucial for ensuring data consistency and quality throughout the survey lifecycle. Westat has been able to utilize and extend core Blaise capabilities to meet a variety of post collection processing needs.

- Managing audit trail data – The Wesaudit tool is an extension of the Blaise standard audit trail feature that tracks movements in the instrument and records field values. The tool's visual interface provides an organized means for viewing audit information recorded by the basic Blaise audit feature. WesAudit makes it easier for programmers to debug code and follow flow through complex instruments. It also facilitates the process of data recovery if Blaise

data files develop problems that cannot otherwise be fixed. In addition, WesAudit can be used as part of interviewer quality monitoring and to facilitate the exploration of audit trail data in order to set up more extensive methodological analyses of audit trails

- CARI coding – CARICode is a web-based coding tool that works in conjunction with the CARI feature of Blaise. Coders listen to digital recordings of interviews while viewing an image of the Blaise screen and complete the coding tasks. Westat is using this technology to address several different components of survey error thus improving the quality of our data collection efforts. Field data collections use CARI technology as an alternative to supervisor observations to review interviewer performance, identifying any skill concerns in coding of audio recordings, and providing detailed feedback and coaching to individual interviewers or more global re-training on survey specific concepts across the broader field staff. Westat also pairs the audio recordings with other survey paradata (e.g., interview timings, case disposition history) as a monitoring tool for identifying possible falsifiers. The tool integrates applicable paradata and enables coders to quickly and comprehensively perform their work. Other uses of CARI include monitoring data quality, validating interviews, recording open-ended responses, conducting usability reviews, and testing instruments.

- Post-collection editing – the Blaise Editing System (BES) is designed to take advantage of the 'replay' features of Blaise COTS to perform editing and coding operations (Laidlaw, et. al. 2010; Allan, et. al., 2001). As data are received from the field, the CAPI data are loaded into Westat's BES. Comments are auto loaded into a Data Decision Log (DDL) module and can be viewed by the editor. As editors review the comments, they can enter a status for the comment, reason for data update if needed, and launch the Blaise instrument from the DDL module. The Blaise instrument will 'replay' in interviewer or editor mode, so that the editor may post a data update. The Blaise replay mode assures that data updates do not alter the integrity of the Blaise datamodel and associated skip, range and formats in the dataset. The metadata for the study are also maintained throughout the editing process and available for review by the BES editors. A Blaise audit trail is automatically updated when data changes are made so that all data changes are tracked and known. The BES system reduces editor and post-processing labor significantly by assuring that the data updates are made in conformity with the skip and flows within the CAPI instrument and CAPI metadata. The reasons for data update are noted in the decision log. Edited and archival versions of the data files are maintained throughout the process. BES generates reports by editor ID, case ID, and variables updated for QC review. A verification module assures that the work of editors is reviewed and QC processes are observed. Coding of other specify responses or other items can also be handled by the system. The BES system also generates frequencies and cross tabulations of data for review.

Data Analysis and Reporting – wherever the information in the Blaise data model can be used to automate ancillary processes or validate data this results in substantive benefits in both productivity and data quality. Westat utilizes the information in the Blaise data model to drive a variety of analysis and reporting functions.

- Metadata and Delivery Data Support – the Data Delivery Metadata System (DDMS) is a web-based tool that supports delivery of metadata associated with datasets including raw, analytic, and public use files. Blaise data files and XML can be used with the tool which provides descriptive information about the context, quality and condition and characteristics of the data. This information helps one to understand how the data were collected and any post-collection handling that might have affected its use and interpretation. The tool also provides documentation of each delivery, version control of data deliveries, and searchable linked metadata.

- SAS software is often used for aspects of post-collection processing. Blaise data are easily transformed into SAS or other database systems (Allan, et. al., 2012).

Conclusion

In this paper we have discussed and illustrated two important dimensions of Blaise that allow it to function as a core technology for survey research organizations:

- Blaise provides a stable and secure platform with extensive and powerful core functionality, a large set of features and tools, and numerous support services. As a result, there are a wide variety of solutions that can be developed using out of-the-box Blaise capabilities.

- The Blaise open architecture and related features permit Blaise functionality to be extended or integrated with other applications to create unique and innovative solutions that go beyond the core Blaise functionality. As a result, the significant value of the Blaise data model and rules engine can be reused to improve efficiency, reduce cost, and ensure a very high level of data consistency and integrity.

The survey research environment will continue to evolve and change at an increasingly rapid pace. Some of the factors that will drive this change include:

- Multi-mode surveys will become increasingly common in order to reach a greater number of respondents, control costs, and achieve desired response rates.
- Real-time feedback loops between instrumentation and management systems will become increasingly important to implement more effective adaptive design and multi-mode control.
- Different types of data, media, embedded links, etc. will become more commonly integrated into survey instruments.
- Different types of end-user devices, e.g., tablets, smartphones, etc. will be used to access CAI systems.
- Browser-based access to surveys will become a more common access mechanism for surveys of all modes.
- The use of virtualized and cloud-based infrastructure to support survey systems will continue to expand.
- Security will continue to be a central issue, particularly with Federal or other governmental surveys, and will continue to evolve as the underlying technology and platforms change.

Blaise has performed extremely well in the survey environment most organizations face today and is strongly poised to meet the challenges of the future. Blaise will continue to be an excellent platform for survey research organizations in the future due to its strong set of core features, sound and flexible architecture, and the commitment of the Blaise support team to continue to extend and evolve the product.

References

Allan, B., O'Reagan, K., and Lohr, B., Dynamic ACASI in the field: Managing all the pieces, Proceedings of the 7th International Blaise Users Conference, September 2001.

Allan, B., Frey, R., and O'Reilly, J., Using metadata in Manipula and Maniplus, Proceedings of the 14th International Blaise Users Conference, April 2012.

Brenner, K and Manoj, S., Automated regression testing of Blaise Internet: A case study, Proceedings of the 13th International Blaise Users Conference, October 2010.

Frey, R., O'Reagan, K., and Brown, N., Servicing Blaise instrument needs in a multi-mode environment, Proceedings of the 14th International Blaise Users Conference, April 2012.

Gowen, L. and Clark, P., Lifecycle processes to insure the quality of Blaise interview data, Proceedings of the 11th International Blaise Users Conference, September 2007.

Hill, D., Deploying Blaise to tablet PCs for mobile use, Proceedings of the 9th International Blaise Users Conference, September 2004.

Laidlaw, M., Rhoads, M., and Shepherd, J., Post-collection processing with Blaise in a distributed environment, Proceedings of the 13th International Blaise Users Conference, October 2010.

O'Reagan, K., Frey, R., and Robbins, K., Impressions of Basil, Proceedings of the 13th International Blaise Users Conference, October 2010.

Rhoads, M. and Snowden, R., Security considerations in Blaise environments: Options and solutions, Proceedings of the 13th International Blaise Users Conference, October 2010.

Segel, P. and O'Reagan, K., Extending Blaise capabilities in complex data collections, Proceedings of the 14th International Blaise Users Conference, April 2012.