

Blaise 5 Paradata Requirements

*Rebecca Gatward, Lisa Wood, Patty Maher and Gina-Qian Cheung
Survey Research Center, University of Michigan*

1. Introduction

Paradata are captured during the survey process (Couper, 1998) and are a valuable source of information about the quality of the survey data and in helping us understand, monitor and inform decision making throughout the survey lifecycle. As has been widely documented, there are various types of paradata. Those which are linked directly to the administration of a survey instrument are usually collected automatically through CAI software, such as Blaise. The source of this type of paradata are primarily audit trails and recordings of part or whole interviews using Computer Assisted Recorded Interviews (CARI), this category of paradata can also include server and client side data for web collection. Another type of paradata are collected or stored in survey and sample management systems. This type of paradata describe the actions and processes involved in data collection, it often includes data about contact attempts, the interview, observations from the interviewer, the sampling frame and respondent related administrative procedures, such as compliance to complete consent forms and payment of incentives.

During the last 15 years, paradata have become an essential tool in survey management and as O'Reilly (2009) describes many survey organizations have developed or enhanced existing management systems with paradata as a central feature.

In 2011 a working group was formed with the purpose of specifying the Blaise community's paradata requirements for Blaise 5 (the new build of Blaise developed by Statistics Netherlands). The group comprised representatives from Blaise User Corporate license holders, including the Survey Research Center at the University of Michigan and Mathematica Policy Research. The requirements paper was presented and discussed at the Blaise Corporate License User Board meeting in January 2013.

The main focus of this paper is to detail the paradata requirements as specified by the user group which, given the composition of the group, are primarily for instrument paradata. However, the paper will also provide some practical examples to illustrate how paradata can be used to inform decision making throughout the survey process and assist in fieldwork monitoring.

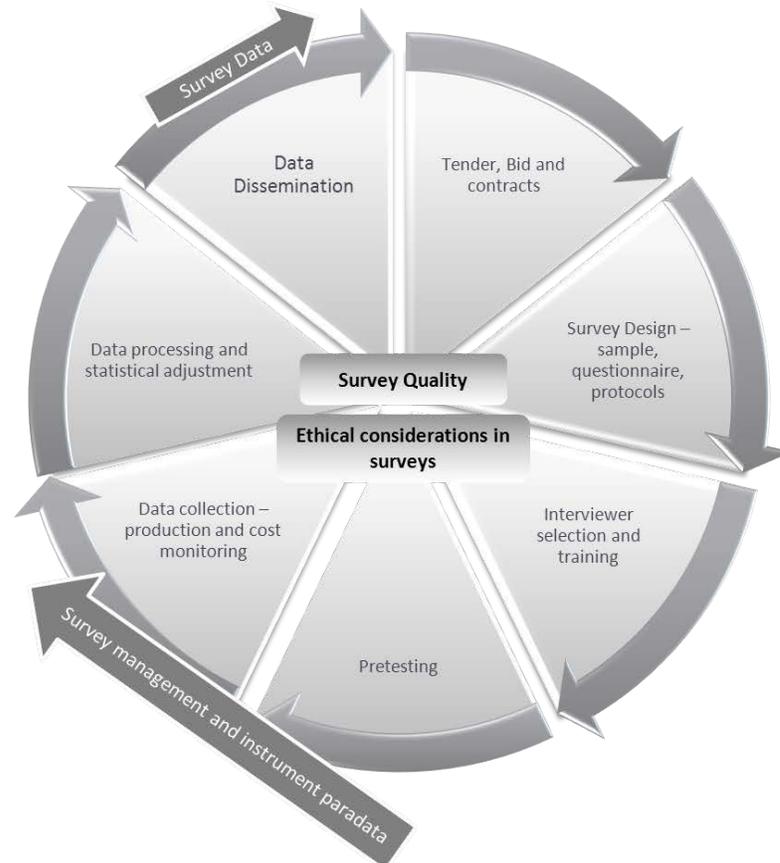
2. Use of paradata

Paradata are collected and used throughout the survey lifecycle – from tender or budgeting , through the design and piloting phase, during data collection and when writing final documentation. The primary uses of paradata are to:

- monitor data collection and other study processes,
- identify methodological error in design,
- reduce survey error,
- improve the quality of survey processes and products,
- aid analysis when reviewing survey measures,
- and improve organizational quality.

Figure 1 illustrates that although specific types of paradata are produced at certain stages of the survey lifecycle, they are used for multiple purposes throughout the survey process.

Figure 1 – Use of paradata throughout the survey process



The following sections provide some examples of the way paradata can be used as a survey and sample management tool, how they contribute to the quality assurance process and the type of analysis that can be undertaken using paradata.

- As a **Survey Management tool**, paradata can be used to:
 - Monitor interviewer’s skills, experience and appropriateness to be staffed on a project, for example, interviewer location, whether they are bilingual, the types of data collection they have undertaken and their success at refusal conversion on current or previous studies.
 - Determine the length of time required to administer a complete interview, an interview session (where the interview is not completed in one session), block of questions or individual questions. This analysis could also be carried out to provide comparisons by mode, by characteristics of the respondent or interviewer. A case study of this type of analysis is provided in a paper by Devonshire (2013).
 - Monitor the data collection progress for each sampling unit (within or across modes), some examples include,
 - contact attempts - time, day and date of contact, mode and outcome,
 - number of times the survey has been started and suspended,
 - point at which a mode switch occurred,
 - or the question at which an interview was suspended or abandoned.

- Identify areas in the questionnaire that are “trouble-spots” through the analysis of key stroke data, such as the frequency of the use of remarks by question or the use of short cut keys assigned to access any question specific help screens.
 - Monitor respondent payments, for example, amount paid to respondents, timing of the payment and total cost. This data could also identify the characteristics of respondents who do not cash their payment or methods that could reduce delays in paying respondents.
 - Track communication with respondents and the effect of any interventions.
 - Finally, paradata can be used to examine progression through the survey, for example, identify if certain questions are revisited multiple times.
- For **Sample Management**, paradata can help inform the following types of key decisions:
 - Planning the number of interviewers required to complete data collection within the desired field period and the assignment of cases to interviewers.
 - Identifying respondent characteristics and contact strategies, for example, the best time to make contact to respondents based on past waves or by characteristics of the household.
 - Determining the cost of the survey (cost per unit) and establishing best collection strategy to move forward.
 - Target collection effort based on stratum response rate.
 - Identifying follow-up strategies based on collection priority or weighting assigned to specific respondents.
 - To identify the need, type and success of a responsive design strategy.
 - Depending on the type of interviewer observations collected, observation data can be used to monitor if interviewers are implementing appropriate strategies for certain situations, identify better listing protocols, reduce coverage error caused by listing or reduce non-response rates by improving tracking methods.
 - Verify interviewer either listed or interviewed at the correct address using GPS coding.
- Paradata can be used as a **Quality Assurance** tool in the following ways:
 - To evaluate interviewer performance – for example, an analysis of the audit trails may detect or support suspicions of interview falsification. CARI files provide recordings which can be used to carry out detailed assessments of interviewers interviewing skills.
 - Where necessary, data can be recovered using audit trails or provides the ‘event history’ of an interview.
 - During testing and can facilitate bug identification, diagnosis and reporting - for example, analysis of paradata can help identify if the source of a reported issue is a specific system, survey instrument or the user. Paradata can also be used to monitor system performance and track the incidence of an issue.
 - To track interaction of the survey instrument with external systems and diagnosis system issues related to this interfacing.
 - Auditing access to survey data (security management).
 - **Adhoc analysis** – paradata can also be used to investigate specific research questions or to explore an issue in greater depth. One key factor to consider is that, although it is useful to collect as much paradata related to the administration of the Blaise instrument as we can, we should aim to do so without impacting the efficiency of the data collection process. Paradata have been analyzed to identify:

- Respondent response pattern and response order effects.
- CAI design improvements – this is particularly useful in the piloting stage of a questionnaire or whilst testing individual questions.
- Respondent preferred data collection method.
- The impact of processes surrounding data collection on the survey outcomes. For example, interviewer characteristics, mode and device used of data collection and the recording process.
- To examine the actions on each page of the instrument and the impact these actions have on survey outcomes. For example, on a web self-administered questionnaire; did the respondent have to scroll? Were questions answered in the order presented (if a grid is used)? Did the respondent use a mouse or keyboard? What was the total number of keystrokes?

3. Paradata requirements

The following is a summary of the requirements as specified in the Blaise 5 paradata white paper. These requirements are specific to paradata related to the administration of the CAPI and Blaise IS instruments, those for the Blaise CATI Survey Management System are not included.

3.1 General requirements

Users requested the following requirements which relate to the collection and storage of the paradata rather than specific elements of paradata.

- First, and most importantly, the level of detail included in the current Blaise ADT files should be retained.
- Security of the paradata is a concern.
 - Paradata should be stored securely (i.e., secure database or encrypted text file).
 - Users should be able to ‘turn-off’ collection of various paradata elements.
 - It should be possible to decouple or store the paradata separately but still be able to merge it with the questionnaire data for analysis.
- The data structure must be consistent across all modes of data collection although the data collected may vary by mode.
- Users should be able to define key project-specific variables to include in the paradata such as stratum, project name, or user-defined status or outcome variables (i.e., variables preloaded into Blaise or calculated within Blaise). These variables should also be read- and write-able in Rules, Manipula and API.
- It would be useful for Blaise 5 to include a ‘paradata viewer’ for all modes: The paradata viewer should provide a survey summary viewer for information at the survey level available at run-time for all modes. This can further be enhanced to provide customizable reports on the survey progress.
- Current and Historical Paradata Viewer. User should be able to view the current paradata in list view for all modes in one interface. Clicking on a selecting case should provide users with the historical paradata transactions associated to the case.

3.2 Specific requirements for Computer Assisted Personal Interviewing

Users determined that instrument paradata for Computer Assisted Personal Interviews (CAPI) should be collected for each respondent at the variable, page, block, session and interview level. The specific elements users requested are detailed in an appendix to this paper (Appendix A). The following are paradata which are required at an interview or specific part of the interview.

- Mode of data collection at a page or item level as well as at a session or form level.
- Geographic information about location of the respondent (e.g., GPS coordinates).
- Paradata should include CARI specific information (e.g., log of questions recorded, name of sound file(s) and/or screen shot(s)).
- Paradata should include information about the CARI set up or recordings (e.g., log of questions recorded, name of sound file(s) and/or screen shot(s)).

3.3 Specific requirements for web data collection

The following is a summary of the paradata required for surveys administered using Blaise IS. Again, the detailed elements required by users are listed in the appendix to this paper (Appendix A).

- When survey data is collected via the web, paradata about the environment in which the survey was administered should be captured - for example, the type of device, operating system, browser, connection speed and screen or browser size.
- Paradata for web surveys should capture data from both the server and client side:
 - Server side includes submissions to server, page-level times and break-offs.
 - Client side involves embedded java script code that captures user actions like changed responses, response latencies, keystrokes, and even mouse movement - for example, mouse coordinates every 'nth millisecond, details about whether and if so how the keyboard and scrolling keys were used and whether horizontal or vertical scrolling used.

4. Conclusion

As is evident from the user requirements and their extensive use throughout the survey lifecycle, paradata continue to be important, especially as mode of data collection impacts quality and cost. Consistency in format of paradata and the ability to turn on and off various elements will help tailor the analysis and usefulness of paradata.

Paradata will continue to be a central part of survey collection operations, the further development of its use and the developments or efficiencies gained through the analysis of paradata are something we can share amongst the Blaise community.

References

- Couper, M. (1998). Measuring survey quality in a CASIC environment. *In proceedings of the Section on Survey Research Methods of the American Statistical Association.*
- Devonshire, J. (2013). 'Adding Business Intelligence to Paradata: The Blaise Audit Trail'. *Presented at the 15th International Blaise Users Group Conference, Washington, DC, USA.*
- O'Reilly, J. (2009). 'Paradata and Blaise: A review of recent applications and research'. *Presented at the 12th International Blaise Users Group Conference, Riga, Latvia.*

Appendix A

Blaise Paradata – Specific Elements

Specific data elements currently used by Survey Research Center for Blaise and BlaiseIS surveys are listed in tables below. In addition to the current elements, the following additional elements should be included:

- Mode of Administration (page level)
- Location of Interview (session level) – IP Address and/or GPS coordinates
- Operating system and device used (session level)
- More detailed information for keystrokes (see number 28 listed below)

BlaiseIS Paradata (this is collected for each page visited)

#	Variable	Type	Size	Description
1	PostID	Num	8	Project-specific variable; preloaded
2	ServerName	Char	7	Name of web server
3	ID	Num	8	Primary Key
4	HashID	Char	8000	
5	PAGESTARTTIMESTAMP	Char	30	Date/Time the current page was started
6	PREVPAGETIMESTAMP	Char	30	Date/Time the previous page was started
7	PREVPAGELENGTH	Num	8	PREVPAGETIMESTAMP - PAGESTARTTIMESTAMP
8	SESSIONID	Char	20	Unique key for each web session
9	PRIMARYKEYVALUE	Char	200	Project SampleID; preloaded
10	VPROJECTID	Char	40	ProjectID; preloaded
11	MODE	Char	10	Mode Used (web SAQ or interviewer-administered)
12	SCREENSIZE	Char	50	Screen Size as "width, height"
13	BROWSERSIZE	Char	50	Browser size
14	JAVASCRIPTENABLE	Char	10	JavaScript enabled = true in the browser
15	CONNECTIONSPEED	Char	20	A value of bytes/second throughput, calculated
16	SURVEYNAME	Char	50	Name for the survey
17	BROWSER	Char	500	Browser name
18	CURRENTPAGENO	Num	8	Interview Page: Cstr(SessionState.StoredPageIndex)
19	PREVPAGENO	Num	8	Like CurrentPageNo, filled in by Blaise IS ASP
20	VERSIONDATE	Char	20	Version date of survey
21	VERSIONTIME	Char	20	Version time of survey
22	ACTION	Char	20	Values are: InterviewStarter, Next, Other, Previous, Start, Submit
23	PREVPAGEQUESTIONS	Char	32767	Like PageQuestions, filled in by Blaise IS ASP
24	PREVPAGEANSWERS	Char	32767	Like PageAnswers, filled in by Blaise IS ASP
25	CURRENTPAGEQUESTIONS	Char	32767	List of questions on current page, ^ delimited
26	CURRENTPAGEANSWERS	Char	32767	List of responses on current page, ^ delimited
27	SUBMITSTATUS	Char	20	Values are Completed or Missing (if "Completed" then Action="Submit")

28	FLASHENABLE	Char	10	Flash enabled
29	OTHERPLUGINSENABLE	Char	10	Other plugins enabled
30	KEYSTROKE	Char	32767	information added about client-side events listed in order that they occur on the page, (i.e., javascript that picks up keystrokes on the page); currently includes data indicating if scrolling occurred (H=Y/N or V=Y/N), but no scrolling details, includes information about use of keyboard vs. mouse, but does not include mouse-click x/y coordinates

ADT Database

The following tables and data elements are compiled from Blaise ADTs into a SQL database;

- tAdtField (this comes directly from the ADT files)

Field Name	Data Type	Size	Description
ProjectID	Text	30	Unique project identifier
CaseID	Text	30	Sample ID within a project
FldSeq	Long Integer	4	Sequential counter; increments by one for each ADTrow
MetaName	Memo	-	Name of Blaise data model
MetaTime	Text	200	Date/Time Blaise data model was created
UserID	Text	30	Interviewer ID for this particular instrument entry
FldName	Memo	-	Blaise long field name
BlockName	Text	100	Blaise block name
FormVstNum	Long Integer	4	Instrument visit number; increments by one for each entry
FieldVstNum	Long Integer	4	Field visit number; increments by one for each entry
PrevLeaveLineNo	Long Integer	4	Previous field Blaise line number
EnterLineNo	Long Integer	4	Current field entered Blaise line number
LeaveLineNo	Long Integer	4	Current field exited Blaise line number
ISLEAVEFORM	Long Integer	4	Interviewer exited Blaise at this field (Yes/No)
EnterDate	Text	30	Date interviewer entered the Blaise field
EnterTime	Text	30	Time interviewer entered the Blaise field
LeaveDate	Text	30	Date interviewer left the Blaise field
LeaveTime	Text	30	Time interviewer left the Blaise field

Field Name	Data Type	Size	Description
Fld_Time	Text	30	Time (sec) spent within the Blaise field
Fld_SS	Text	30	Time (sec) spent within the Blaise field (different format)
Fld_Time_Mss	Long Integer	4	Time (milisec) spent within the Blaise field
Btw_Time	Text	30	Time (sec) spent between last Blaise field and this one
Btw_SS	Text	30	Time (sec) spent between last Blaise field and this one
Btw_Time_Mss	Long Integer	4	Time (milisec) spent between last Blaise field and this one
Adj_Time	Text	30	Fld_Time + Btw_Time
Adj_SS	Text	30	Fld_Time + Btw_Time
Adj_Time_Mss	Long Integer	4	Fld_Time + Btw_Time
RspLat_Time	Text	30	Time (sec) between entering field and first keystroke
RspLat_SS	Text	30	Time (sec) between entering field and first keystroke
RspLat_Time_Mss	Long Integer	4	Time (milisec) between entering field and first keystroke
Key_Count	Long Integer	4	Number of keystrokes while in Blaise field
Enter_Value	Memo	-	The value of the Blaise field upon entry
Leave_Value	Memo	-	The value of the Blaise field upon exit
Leave_Cause	Text	50	Action that initiated interviewer to leave Blaise field
Leave_Status	Text	50	Field leave value is normal or DK/RF
Prev_Lang	Long Integer	4	Language was switched to previous while in field (Yes/No)
Next_Lang	Long Integer	4	Language was switched to next while in field (Yes/No)
Set_Lang	Long Integer	4	Language was set while in field (Yes/No)
CtrlL_SetLang	Long Integer	4	Language was changed with hot key while in field (Yes/No)
ALTXExit	Long Integer	4	Alt-X interview suspension was initiated at this field
RemClk	Long Integer	4	Interviewer remark was initiated at this field (Yes/No)
RemChng	Long Integer	4	Interviewer remark was changed at this field (Yes/No)
QHelp	Long Integer	4	QXQ Help was initiated at this field (Yes/No)
BlaiseHelp	Long Integer	4	Blaise Help was initiated at this field (Yes/No)
Error_Esc	Long Integer	4	Blaise check was closed at this field (Yes/No)
Error_Esc_Text	Memo	-	Text of the Blaise check encountered before closing
Error_Supp	Long Integer	4	Blaise check was suppressed (Escape) at this field (Yes/No)
Error_Supp_Text	Memo	-	Text of the Blaise check encountered before suppressing
Error_Jmp	Long Integer	4	Blaise field that interviewer jumps to after Blaise check
Error_Jmp_Text	Memo	-	Text of the Blaise check encountered before jumping
Media_Start	Long Integer	4	Blaise launched media file while in this field (Yes/No)
Mouse_Click	Long Integer	4	Any mouse click detected while in this field (Yes/No)

Field Name	Data Type	Size	Description
F1	Long Integer	4	F1 hot key was pressed while in this field (Yes/No)
F2	Long Integer	4	F2 hot key was pressed while in this field (Yes/No)
F3	Long Integer	4	F3 hot key was pressed while in this field (Yes/No)
F4	Long Integer	4	F4 hot key was pressed while in this field (Yes/No)
F5	Long Integer	4	F5 hot key was pressed while in this field (Yes/No)
F6	Long Integer	4	F6 hot key was pressed while in this field (Yes/No)
F7	Long Integer	4	F7 hot key was pressed while in this field (Yes/No)
F8	Long Integer	4	F8 hot key was pressed while in this field (Yes/No)
F9	Long Integer	4	F9 hot key was pressed while in this field (Yes/No)
F10	Long Integer	4	F10 hot key was pressed while in this field (Yes/No)
F11	Long Integer	4	F11 hot key was pressed while in this field (Yes/No)
F12	Long Integer	4	F12 hot key was pressed while in this field (Yes/No)
CtrlD	Long Integer	4	Ctrl-D hot key was pressed while in this field (Yes/No)
CtrlR	Long Integer	4	Ctrl-R hot key was pressed while in this field (Yes/No)

- tSuspendVariables (calculated using the adt files).

Name	Type	Size	Description
ProjectId	Text	30	Unique project identifier
CaseID	Text	30	Sample ID within a project
vTotalSuspend	Long	4	Total number of suspends for this sample line
vCaseComplete	Yes/No	1	Is the case completed?
vSuspendedVariable1	Memo	-	Blaise field at which first suspension occurred
vSuspendedVariable2	Memo	-	Blaise field at which second suspension occurred
vSuspendedVariable3	Memo	-	Blaise field at which third suspension occurred
vSuspendedVariable4	Memo	-	Blaise field at which fourth suspension occurred
vSuspendedVariable5	Memo	-	Blaise field at which fifth suspension occurred
vSuspendedVariable6	Memo	-	Blaise field at which sixth suspension occurred
vSuspendedVariable7	Memo	-	Blaise field at which seventh suspension occurred
vSuspendedVariable8	Memo	-	Blaise field at which eighth suspension occurred
vSuspendedVariable9	Memo	-	Blaise field at which ninth suspension occurred
vSuspendedVariable10	Memo	-	Blaise field at which tenth suspension occurred

Name	Type	Size	Description
vDate1	Text	30	Date on which first suspension occurred
vDate2	Text	30	Date on which second suspension occurred
vDate3	Text	30	Date on which third suspension occurred
vDate4	Text	30	Date on which fourth suspension occurred
vDate5	Text	30	Date on which fifth suspension occurred
vDate6	Text	30	Date on which sixth suspension occurred
vDate7	Text	30	Date on which seventh suspension occurred
vDate8	Text	30	Date on which eighth suspension occurred
vDate9	Text	30	Date on which ninth suspension occurred
vDate10	Text	30	Date on which tenth suspension occurred
vTime1	Text	30	Time at which first suspension occurred
vTime2	Text	30	Time at which second suspension occurred
vTime3	Text	30	Time at which third suspension occurred
vTime4	Text	30	Time at which fourth suspension occurred
vTime5	Text	30	Time at which fifth suspension occurred
vTime6	Text	30	Time at which sixth suspension occurred
vTime7	Text	30	Time at which seventh suspension occurred
vTime8	Text	30	Time at which eighth suspension occurred
vTime9	Text	30	Time at which ninth suspension occurred
vTime10	Text	30	Time at which tenth suspension occurred
vLastSuspendedVariable	Memo	-	Last Blaise variable at which a suspension
vLastSuspendedDate	Text	30	Date on which the last suspension occurred
vLastSuspendedTime	Text	30	Time at which the last suspension occurred

- **tCariField** (this from the DRI/CARI log)

Field	Type
ProjectID	varchar(30)
CaseID	varchar(50)
FormNum	int
FieldSeq	int
FileNum	smallint
FieldType	smallint
FieldName	varchar(400)
StartLineNum	int
EndLineNum	int
Mss	bigint
Ss	bigint
AccuMss	bigint
AccuHHMMSSMSS	varchar(50)
StartTimeStr	varchar(50)
EndTimeStr	varchar(50)

- **tCaseRecord** (this from the DRI/CARI log)

Field	Type
ProjectID	varchar(30)
CaseID	varchar(30)
RecordID	int
vLastChangeVarName	varchar(400)
vLastChangeDate	varchar(30)
vLastChangeTime	varchar(30)
vLastAccessVarname	varchar(400)
vUserID	varchar(30)
vDataModelName	varchar(30)
Adj_Time_Mss	int

Specific elements for BlaiseIS (organization two)

Interview Level Paradata: (via the Journal or journal like tool)

Record for each page presented to the respondent and also after the submission or movement from away from the current page thus recording a history of all respondent transactions.

At entry to page, record the following:

- Primary Key
- Referring Page (where they came from)
- Blaise IS page number
- Blaise .bmi version
- Blaise language #
- Time page was initially displayed (mm/dd/yy hh:mm:ss.00)
- Screen resolution (# x #)
- Browser resolution (# x #)
- Java enabled (Y/N)
- Question(s) displayed (Blaise Variable names)

At exit of page, or movement off the page (a hyperlink or button clicked, browser closed)

- Primary Key
- Blaise IS page number
- Blaise .bmi version
- Blaise language # (especially if they can toggle during interview)
- Time page was submitted (mm/dd/yy hh:mm:ss.00)*
- Screen resolution (# x #)
- Browser resolution (# x #)
- Java Enabled (Y/N)
- Question(s) displayed (Blaise Variable names) and answer(s) submitted
- Page action (page submitted, back button, closed browser, etc.)
- Note: if the session timed out/closed browser, record time as submitted (or at the least when the IIS session ended).

If scripting is enabled, researchers might be interested in movement throughout the page, the order items were selected (or unselected), but we haven't come across many requests for that level of detail. For the most part, they've been interested in how many times a page was viewed by the respondent and how the data changes per each submission. Since screen size (browser or machine) could change while the interview is active, it should be recorded on the interview level rather than the session level.

One note on collecting the browser resolution: while we can capture the size of the window, we can't capture if they are utilizing the zoom feature on their browser, or how they browser is configured (for example, which task bars are visible), therefore it is almost impossible to represent exactly what the respondent is seeing with 100% accuracy and this data may or may not be useful.

Session Level Paradata

While some of this data can be collected via IIS logs, it would be ideal to have it linked to the primary key in the Blaise instrument, especially for debugging purposes. This data would be collected from point survey session begins until the survey session ends either by the respondent properly closing the session

or by the respondent reaching a time out limit. Ideally it should be maintained separate from the interview data, as to maintain confidentiality.

- Primary Key
- Start time of session
- End time of session (recorded from session time out or via proper exiting of the questionnaire)
- IP Address (this should be optional to record and not mandatory to collect)
- Location of host (country or state, but this should be optional to record and not mandatory to collect)
- Browser & Version
- Operating System
- Status on exit (Submit, Abort, Timed out)

Specific Elements for Blaise 5 (organization three)

1. System Level: This records system information like the Diagnostic Tool.
 - Blaise Software version
 - Component installed
 - Services installed
 - OS of server
 - License
2. Survey Level. This records survey information like information in CATI Specification, *.log, and datamodel parameters.
 - Survey_ID
 - Survey Name
 - Version of data model
 - Version of Blaise software compiled
 - Modes (CATI, Web, Capture, etc) defined
 - Mode specific specifications
 - Language (English, French, etc) defined
 - Collection start date
 - Collection end date
 - User Define Management Variables:
 - Max 3 Stratum variables (used for filtering cases and to summarize performance indicators)
 - Max 10 Key variables (used for filtering cases in the dataviewer)
 - Key Performance Indicators (could also be by stratum):
 - Completion Rate (% of cases completed=finalized or extracted)
 - Response Rate (% of finalized or extracted “in-scope” “response cases)
3. Sampling Unit (a.k.a Case) Level (e.g. a household or an enterprise). This records general information and status about the Case.
 - Survey_ID
 - Case_ID
 - Link_ID to other cases
 - Collection mode assigned
 - Collection priority or weight
 - Time Zone
 - Geographic Identifier

- Respondents information (could be an array)
 - Name
 - Address and Geographic Identifier
 - Telephone
 - Email
 - Time Zone
 - If multiple Unit of Interests (UOI) (UOI can be members in a household or locations in an enterprise)
 - #Unit of Interests
 - Unit of Interests completion status
 - User Define Status or Outcome variables
 - Sampling Unit Resolution Status (No started, In progress, Pending, Disabled, Finalized, Extracted)
 - Sampling Unit Validity status (unknown, in-scope, out-of-scope)
 - Sampling Unit In Progress Sub-Status (No Contact, Validated Sampling Unit, Contacted with Sampling Unit, Screening Completed, UOI collection started)
 - Sampling Unit Finalized Sub-status (Response [full or partial case], non-response)
 - History of events associated to the case
4. Per Event (also per Interview or Intra-case) level . This records events that make changes to the cases data. This includes information in the BTH for CATI or the Journal for Web. There is a need to add event paradata via Manipula and API.
- Survey_ID
 - Case_ID
 - Mode
 - Language
 - Date/Day/Time of event
 - Duration of event
 - If Interviewer assisted
 - Interviewer information
 - Interviewer observation
 - Mode specific event outcome (complete, partial, refusal, appointment, change mode, etc)
 - Event Flags (see attached document)
 - Respondents Information changes log if applicable
 - Audit Trail for the event (same info as in ADT file)
5. Question Level. This is similar to a field in the ADT file or the CARI.
- Survey_ID
 - Case_ID
 - Question_ID
 - Mode
 - Language
 - Question text asked
 - Answer category asked
 - Answer Entered or key pressed to exit question
 - Time entered question
 - Time exit question
 - Recording (Sound and/or Screen.)