

Choréo – The Blaise 5 Multimode Management System

Mark M Pierzchala / Mark M Pierzchala, MMP Survey Services, LLC (MMPSS)

This is the status of the Blaise 5 Multimode Management System, now called **Choréo**¹ as of July 2020.

1. Abstract

Blaise 5 Choréo will be demonstrated and discussed. The Choréo architecture closely follows the design agreed by the BCLUB Multimode Management Group as demonstrated at IBUC 2018. Choréo was first prototyped in Blaise 4 and an information webinar was given in August 2019. It was ported to Blaise 5 in November 2019 where work has continued.

While the overall architecture follows the BCLUB design, there were many details to work out. In addition, Choréo is designed to fit into an institute's existing infrastructure. For example, most institutes already have a Sample Management System (SMS). Such a system can be used to send mail and email, and to receipt responses when they arrive, among many other tasks. Choréo does not replicate such functionality.

Choréo does provide a Survey Handling System (SHS) with a Survey Handling Database (SHD). The Survey Handling Database keeps track of survey management statuses, counts, and indicators. It does not contain any Personally Identifiable Information (PII). Such confidential data are held in the institute's own Sample Management System.

The Survey Handling System issues instructions for all survey management actions. To do so, it must be able to communicate with an institute's systems. Further, Choréo should operate the way the institute wants. This is done through specification databases. There are databases for (1) survey design parameters, (2) happenings, and (3) actions. It also works with Blaise 5 CATI. The Choréo should be able to produce reports using an institute's own coding system. Further, there are hooks in the system to implement management responsive design.

How all this is elegantly done will be revealed at virtual IBUC 2020.

2. A Brief History of Choréo

The BCLUB Multimode Management Group was formed in October 2017. It was to define functionality for multimode survey management. The group members represented a wide variety of experiences and practices. The final report (Pierzchala, et al, 2018) was released to BCLUB members in September 2018. At the 2018 International Blaise Users Conference, a high-level description was given (Pierzchala 2018).

A webinar for BCLUB members in August 2019 showed a rough prototype programmed in Blaise 4. Lon Hofman of the Blaise Team at Statistics Netherlands ported the system to Blaise 5 in November 2019. A

¹ Previous system names were MMS and M3. Choréo is a new faux word derived from 'choreography' since multimode management must choreograph many cases and outcomes. 'Choréo' is rooted in Greek, and is present in English, French, German, Dutch, and other languages. The 'é' is stressed. The pronunciation is 'core-ray-oh'.

key feature that made this port possible was the extension of Manipula to handle dialogs. Finally, a status update for BCLUB members was posted on BaseCamp at the end of December 2019.

Recently, there was a first release of the CAPI Management App (CMA) as well as improvements to CATI. The progress of CATI, the CMA, and Choréo make it possible to integrate these 3 modules.

Recently, MMPSS reread the original BLCUB report, the 2005 multimode article by de Leeuw, and the 2016 AAPOR report on outcome dispositions. A new AAPOR report released in October 2019 (Olson, et al, 2019) provided an additional excellent review of multimode survey management needs resulting in new Choréo features or hooks for future features. Another valuable resource was a paper that proposed an outcome coding scheme for UK household surveys (Lynn et al, 2001).

2.1 The Near Future

The combined Choréo, CATI, and CMA sub-modules should be ready for testing in late 2020. Testing such a system, with thousands of possible sequences, requires automated testing running on a network.

3. The Goals of Choréo

Choréo has been designed to achieve the following goals:

- Provide a multimode management system in Blaise, that
- Fits easily into an institute's existing infrastructure, where
- The Choréo capability is value-added to the institute's own systems, that
- Is generalized and built around principles that are common to institutes, while
- Allowing an institute to maintain its own survey management coding scheme, that
- Is mainly implemented through an easy specification exercise for the institute, where
- The system is extensible for the institute by specification or modifying source code, which
- Applies to person- and household-level surveys, and at the same time,
- Collects massive case-level survey paradata that enable the institute to implement
 - Its own reports (in addition to some basic Choréo reports)
 - Its own Responsive or Adaptive Survey Design (Groves and Heeringa, 2006), Kreuter (et al, 2013), and Schouten et al (2018).

☛ Choréo is based on operational codes and statuses. In both AAPOR (2016) and Lynn et al (2001) these are called temporary codes. Choréo formalizes the handling of operational statuses. They contain enough information to map them to final disposition codes. See below for more information.

4. Choréo High-Level Design

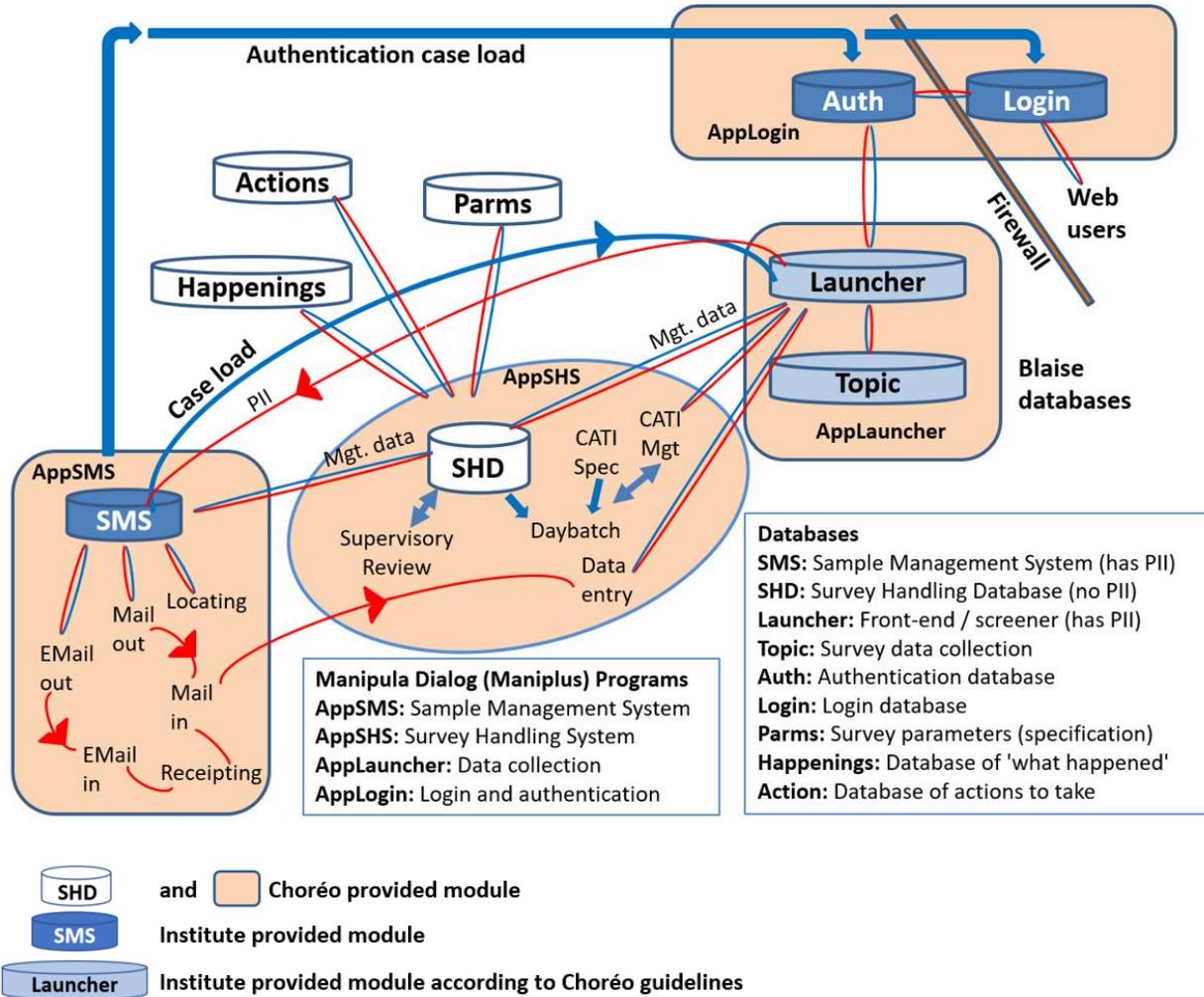
The overall design and principles of Choréo have not changed since the 2018 BCLUB report. Figure 1 shows the overall architecture. Key design principles include:

- The Survey Handling Database (SHD), its external files Happenings, Actions, and ParmS, and the 4 Manipula Dialog (Maniplus) App programs make up Choréo. The Launcher and Topic Blaise instruments would be provided by the institute, according to certain Choréo guidelines.
- The SMS, the web infrastructure, and many subsystems are provided by the institute.
- The internal Choréo source code will not refer to any published coding system.

- Tokens (explained below) are used in the source code to refer to happenings and actions.
- Choréo is built around happenings and actions, and survey design parameters.

Figure 1: The High-Level Design of Choréo

Choréo Multimode Management System



The Survey Handling Database (SHD) is a Blaise provided database as are the external files used by AppSHS. The Launcher and Topic instruments may be Blaise datamodels, but this is not required.

The relationship between the Sample Management System (SMS), the Survey Handling Database (SHD), and the AppSHS is as follows:

- All Personally Identifiable Information (PII) is handled by the SMS.
- No PII is held in Choréo. None!
- Launcher holds selected PII that is necessary to identify the respondent or for an interviewer to conduct the interview.
 - The Launcher can collect or modify PII. This new information is sent directly to the SMS bypassing the Survey Handling Database (SHD) though the SHD knows about this.

- There is constant communication between the SMS, the SHD, and other modules via AppSHS.
- AppSHS commands Choréo including the SMS. When there is an independent SMS action (e.g., a paper questionnaire is receipted), it reports the action to the SHD.

5. Building Choréo around Happenings, Actions, and Survey Parameters

A happening is simply anything that can happen to a case during a survey. A happening occurred in the past. Examples include:

- A busy signal
- The respondent slammed a door in the face of an interviewer
- A web respondent logged in
- An email bounced back

An action is anything that can be done with a case. An action would occur in the future. Examples include possible actions to take after the happenings above:

- Call back the respondent in a few minutes
- Record an adamant refusal
- Rest the case for all outbound contact attempts
- Change to a different email address if available

The Parmis database is a survey design database. It is philosophically like the CATI specification module, but its scope is for the entire survey. At some point, either (1) a clear line will be drawn between the CATI specification module and Parmis, or (2) they will be combined. For the Actions above, the Parmis survey specification or the CATI specification file would be used as follows:

- CATI specification: Follow the callback pattern for busy signals
- Parmis: Stop the case
- Parmis: Indicates the length of the rest period
- Parmis: Switches the email pointer to a different email, if one is available

5.1 Why Happenings and Actions

Possible happenings and actions are common to all institutes. A busy signal may sound different from country to country, but a busy signal is a busy signal. Happenings and actions are operational in nature. Choréo provides a standard and extensible happenings database of around 400 records and an actions database of more than 200 hundred records.

The use of happenings and actions avoids using any published outcome coding system during survey operations. However, the AAPOR (2016) outcome coding system was heavily referenced mainly to determine which attributes of a happening or action would be useful for survey operations.

Records in the happenings and actions databases are identified by tokens (primary key). These are character strings that are used (hopefully minimally) in the Choréo source code. For each record (identified by a token) the institute is free to:

- Give any numeric or alpha-numeric code to the happening or action record

- Give any description to the happening or action record
- Change specific attributes of a happening
- Add a happening or action record that has a unique token

The code and label can be based on the AAPOR or AAPOR-like coding system, or it can be based on a long-standing institute coding system. Choréo does not care.

If a new happening or action record is added, the institute may have to modify one or more Manipula procedures that indicate how to handle the new record (see below).

5.2 Generalization of Choréo with Tokens

Figure 2 shows a few records from the Happenings database specification.

Figure 2: Happenings Specification with Token

Happenings Token	InstCode	MetaDescription
NonC_Busy_Phone	11336	Busy
NonC_SomeOneElse	90337	Contact with someone else
NonC_VMNoIDNoMsg_Phone	11401	Voice mail, SM no ID, no left message

The token *NonC_Busy_Phone* contains 3 pieces of information separated by the underscore: *NonC* indicates this happening is a non-contact, *Busy* (a busy signal) is the happening, and *Phone* is the channel of communication. *InstCode* is an institute-provided alpha-numeric code. This entry enables the institute to connect with its own systems and coding schemes. *MetaDescription* is also provided by the institute. The second entry token, *NonC_SomeOneElse*, applies to all channels of communication or modes because it does not have a third component.

The main point of the happening, e.g., *Busy*, is that it can happen to any survey that uses the phone channel of communication. This allows Choréo to adapt to any institute’s survey program without imposing an outcome coding scheme.

Figure 3 shows a tokenized action. The token name uses the same concepts as the Happenings tokens. *ActionCode* and *Action Descriptions* are institute specific while the token is used by Choréo.

Figure 3: Action Specification with Token

Actions Token	ActionCode	Action Descriptions
Follow_Busy_CATI	260	Follow the CATI busy treatment

5.3 Careful Use of Terms

Two of the happenings in Figure 2 use the term *Phone* while the token shown in Figure 3 uses the term *CATI*. This is a deliberate usage of these terms. *Phone* refers to the general use of the phone as a channel of communication while *CATI* refers to the use of the Blaise CATI Management System.

5.4 Attributes of Happenings Tokens

The specification of happenings includes many attributes for each happening. Figures 4, 5, and 6 show additional attributes of happenings specifications.

Figure 4: Happenings Tokens with Attributes (basic information)

Happenings Token	WhoActed	Burden	Cost
NonC_Busy_Phone	NoOne	3	2
NonC_SomeOneElse	NoOne	3	1
NonC_VMNoIDNoMsg_Phone	Interviewer	5	3

Figure 5: Happenings Tokens with Attributes (implications of the happening)

Happenings Token	BurdenReset	Direction	Contact	Identity	Eligibility	ChannelValidity	SPImplication	SampleCategory
NonC_Busy_Phone	No	Outbound	Unknown	No	No	NoInfo	NonContact	UnknownEligibility
NonC_SomeOneElse	No	Outbound	Unknown	No	No	Invalid	NonContact	UnknownEligibility
NonC_VMNoIDNoMsg_Phone	No	Outbound	No	No	No	UnkValidity	NonContact	UnknownEligibility

Figure 6: Happenings Tokens with Attributes (linked actions, the first of several each, there are more →)

Happenings Token	DLA1Token	DLA1Text	Pr1
NonC_Busy_Phone	Follow_Busy_CATI	Follow the CATI busy treatment	Required
NonC_SomeOneElse	SendSupReview	Send case to supervisory review	Required
NonC_VMNoIDNoMsg_Phone	Rest_CATI	Rest case for outbound CATI	Required

The presence of attributes makes the internal Choréo source code easily adaptable to the institute. To the extent possible, Choréo is programmed in terms of attributes.

5.5 Choréo Programming and Adaptability Strategy

Choréo is programmed in Manipula and refers to Blaise databases. There are 4 master programs that start with the acronym App. These are **AppSHS**, **AppSMS**, **AppLogin**, and **AppLauncher**. AppSHS and AppLauncher are part of the distribution. AppSMS and AppLogin stand in for the institute's own systems. For now, they allow thorough testing. Each Manipula program uses many small procedures.

Internally applied procedures: These are procedures that impact the inner workings of Choréo. For example, if only AppLauncher and AppSHS are involved, this is totally within Choréo.

Externally applied procedures: These procedures connect Choréo to the institute's own IT infrastructure. For example, if AppSHS must send something to the institute's SMS, then the institute will want to adapt this procedure to govern this interaction. This might be done by converting some Manipula assignment statements to Manipula API statements.

For each happening, every effort is made to program Choréo solely in terms of happenings attributes. While this is the overall goal, there are times when the Choréo source code refers to tokens.

Choréo programming in terms of attributes: When this is achieved, this part of the Choréo system is completely parameterized. This means you can change the behaviour of the system by changing the attributes of the happenings record, by adding additional happenings or actions records, or modifying the contents of the Survey Parameters Database.

Choréo programming in terms of tokens: When this is necessary, the institute may have to modify one or more of the Manipula procedures. All Choréo Manipula source code is distributed.

5.6 Extensibility of Happenings and Actions

Choréo provides hundreds of specified happenings and actions that apply to almost any institute. However, the institute can add its own happenings and actions tokens along with their attributes. If

Choréo can operate on their attributes, then there is no further work for the institute (in theory). On the other hand, if direct reference to tokens is required, then the institute may have to modify source code.

5.7 Successions of Happenings and Actions within an Attempt

For an attempt, there will usually be successions of happenings and actions. Let us use the example of a phone attempt.

Happenings (past tense) include:

- Case was delivered to the interviewer
- Interviewer reviewed the case
- Interviewer placed a call
- A busy signal was recorded (happening NonC_Busy_Phone)

Happenings are duly recorded in the Survey Handling Database (SHD) because they actually occurred.

Actions (future tense) for the happening NonC_Busy_Phone are recorded as attributes of the happening record (with identifier NonC_Busy_Phone). Actions to execute will or may include:

- Refer to the CATI management system busy treatment specification (required)
- Call back in a few minutes (if indicated by the CATI Management system)
- Retire the case for the day (if the action above is not applicable)
- If there is a succession of busy signals, refer the case to supervisory review (if the action above is not applicable and if indicated by survey parameters)
- Switch to a new phone number if available (if the action above is not applicable).

The sequence of actions includes required, possible, and/or contingent actions. The requirement status of each action is recorded in the happening record. See Figure 6. Note that Figure 6 shows only 1 of several actions that are possible for each happening.

5.8 Implied Operational Statures – Derived from a Succession of Attempts

Some cases will have lengthy and complicated attempt histories. Choréo anticipates these situations and provides adaptable procedures to assess the true situation. For example, a succession of No Answer happenings may indicate the respondent is travelling and the case can be put on rest for a while.

6. Survey Parameters

The survey parameters file, *Parms*, is where you design your survey. Figures 7, 9, and 10 give an idea of the many possible survey parameters.

Figure 7: Page 1 of the Parms Specification

Figure 7 shows that Choréo is designed to handle person-level and two-stage surveys. This is enabled by a two-part ID number. Figure 8 shows a UnitID and a CaseID. Unit 10002 holds 2 cases.

Figure 8: Two-Part ID Structure

UnitID	CaseID
10001	1
10002	1
10002	2

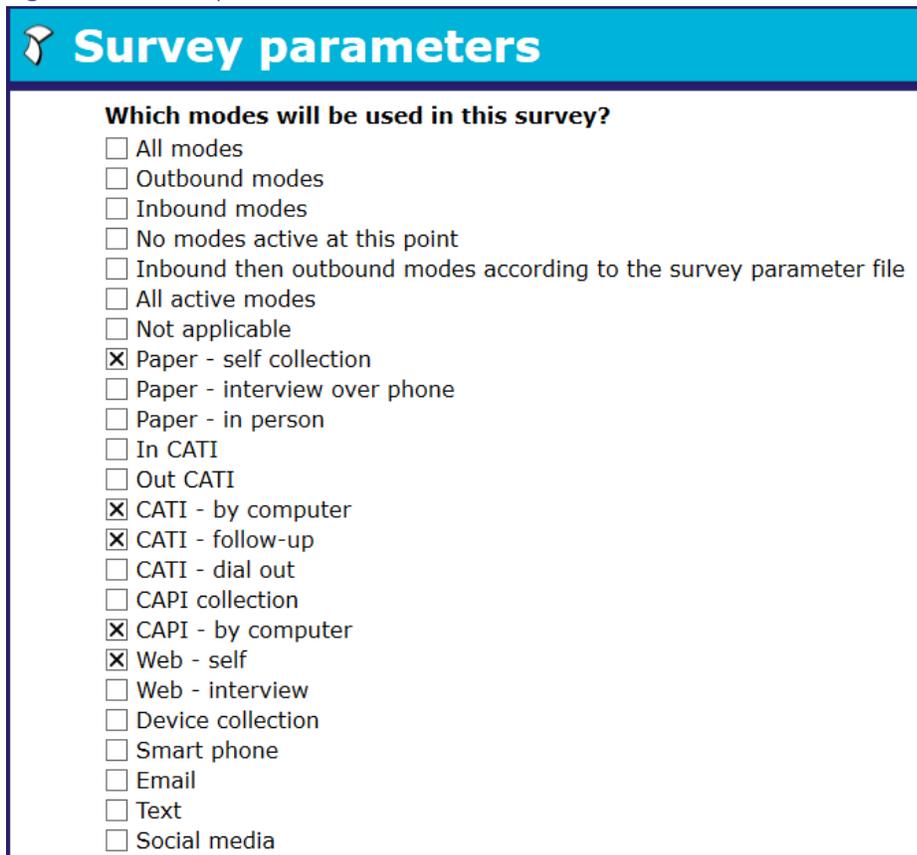
Figure 8 shows the situation, known ahead of time, where 2 people from the same household are selected for a survey. This can happen in a survey of program participants where 2 or more adults from the same household participated in the program. This is operationally important to know. For example, if you interview person 1, then after this first interview you can ask if person 2 is available. If so, then person 1 can hand over the phone to person 2 and it is possible to get the second interview immediately.

Alternatively, where you access a household, enumerate the members, then determine which are to be interviewed, Choréo can easily handle the spawning of cases. In a household, each would have the same UnitID but a different CaseID and its own SHD record. This latter scenario also applies to a corporation with subsidiaries (complex corporations can be represented using a cleverly constructed 2-part ID). Longitudinal surveys can be handled similarly. Unit-level reports can be generated by UnitID.

Figures 9 and 10 show that Choréo distinguishes between data collection modes and channels of communication. This distinction is due to de Leeuw (2005). For example, 2 modes shown in Figure 9, paper and CAPI, rely on a physical address that underlies the mail and visit channels of communication

(figure 10). Choréo knows if a happening applies to a mode or to a channel. Thus, for example, the paper mode is distinguished from the mail channel.

Figure 9: Modes Specification

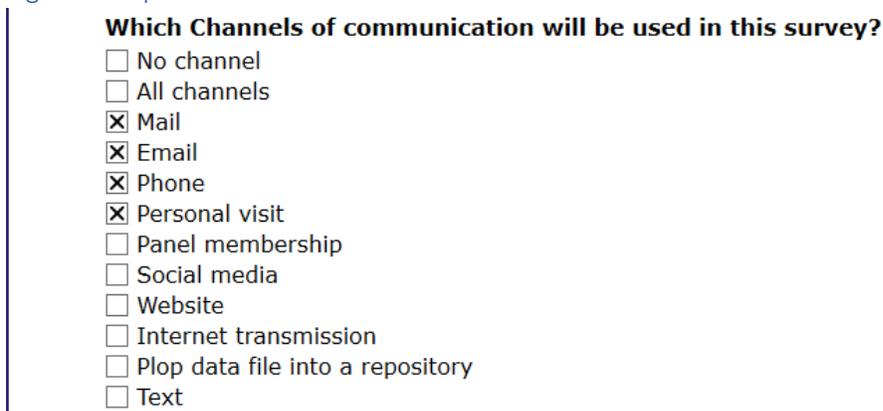


Survey parameters

Which modes will be used in this survey?

- All modes
- Outbound modes
- Inbound modes
- No modes active at this point
- Inbound then outbound modes according to the survey parameter file
- All active modes
- Not applicable
- Paper - self collection
- Paper - interview over phone
- Paper - in person
- In CATI
- Out CATI
- CATI - by computer
- CATI - follow-up
- CATI - dial out
- CAPI collection
- CAPI - by computer
- Web - self
- Web - interview
- Device collection
- Smart phone
- Email
- Text
- Social media

Figure 10: Specification of Channels of Communication



Which Channels of communication will be used in this survey?

- No channel
- All channels
- Mail
- Email
- Phone
- Personal visit
- Panel membership
- Social media
- Website
- Internet transmission
- Plop data file into a repository
- Text

Survey parameters allow many design options including whether the survey is staged (e.g., one mode follows another) or is racehorse (all modes all the time), the length of kinds of rest periods, how many attempts to allow per case, and so on. Operationally, Choréo records happenings, refers to indicated actions, considers sequences of attempts, refers to Parns to fill in details, then acts through a procedure called *HazLo*, Spanish for do it!

7. Operational Codes versus Final Disposition Codes

Operational codes and final disposition codes are highly related, but they are not the same. An operational code, or a sequence of operational codes for a case, must map to a final disposition code as the survey closes. Operational codes are useful when they help survey-conducting staff optimize effort. According to Lynn et al (2001) the mapping to final disposition codes should be objective and as automated as possible. However, there are times when the final code is assigned by survey staff through a manual process. These tend to be cases where there is a complicated history. In Choréo, this history is stored in the SHD. Table 1 below gives some high-level operational status scheme implemented in Choréo.

Table 1: High-Level Operational Status Scheme

<p>Working: Engaged Active NoChange</p> <p>Refusal: RSoftRefusal RHardRefusal OSoftRefusal OHardRefusal AdamantRefusal</p>	<p>Appointment: RDefAppoint ODefAppoint RAmbAppoint OAmbAppoint</p> <p>Breakoff – Resistance: BreakoffIntro BreakoffAfterIntro BreakoffOther</p>	<p>Rest case: RestToRefield ProcessToRefield PendLocating</p> <p>Supervisory review: SupReview</p> <p>Stop: StopOutbound StopAll</p>
--	--	---

6.1 Statuses, Counts, and Indicators

Choréo keeps track of several kinds of statuses, counts, and indicators. These values are stored in the SHD. To map to final dispositions, each case tracks (1) whether R is ever identified, (2) whether R is ever contacted, (3) whether R has ever been engaged, (4) whether R is eligible, and (5) completion status.

For each happening, it is possible to assign a relative burden value and a cost value. Both use scales from 0 to 100. The concept of burden is rarely if ever defined well. Perhaps it is easier for an institute to say that a Busy signal is no burden to R, a left message has relative burden 5, making an appointment has relative burden 10, and completing an interview has relative burden 30. These are accumulated. It is possible to reset the overall burden for some happenings. For example, if you have been pursuing a phone number only to find out that it is not connected to R, then Choréo can reset R's burden to zero.

Similarly, relative costs are used. Placing a phone call may have relative cost of 2 while locating a new phone number may have a relative cost of 15. This gets away from the problems of knowing exact cost per happening and tying the system to a currency.

It is possible in the Parms survey specification database to establish burden, and cost limits. When these are reached, the case is stopped or sent to supervisory review.

8. Current State of Affairs for Choréo

Work over the past few years resolved many difficult conceptual issues and details. Discussions after the 2019 webinar clarified how to make Choréo generalizable using tokens and other measures. The emergence of the CMA as well as further development of Choréo and CATI now enable full-scale prototype testing.

Testing of Choréo will be on a network of computers using automated scripts that can randomly execute sequences of attempts for a case. The Survey Handling Database (SHD) holds a wealth of survey management information. Listings and reports will be generated that will allow an analysis of the sequences of events and whether they were handled correctly. An interesting test will be to see how well Choréo maps to the AAPOR (2016) coding scheme and the scheme proposed by Lynn et al (2001).

9. References

The American Association for Public Opinion Research. (2016). *Standard Definitions: Final Dispositions of Case Codes and Outcome Rates for Surveys. 9th edition.* AAPOR. Found at:

[http://www.aapor.org/Standards-Ethics/Standard-Definitions-\(1\).aspx](http://www.aapor.org/Standards-Ethics/Standard-Definitions-(1).aspx)

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

Groves, R. M. and Heeringa, S. G. (2006). Responsive Design for Household Surveys: Tools for Actively Controlling Survey Errors and Costs. (c) Royal Statistical Society, *J. R. Statist. Soc. A* (2006) 169, Part 3, pp. 439-457

Kreuter, Frauke (editor). *Improving Surveys with Paradata – Analytic uses of Paradata*, © 2013, John Wiley and Sons.

Lynn, P., Beerten, R., Laiho, J., and Martin, J (2001). Recommended Standard Final Outcome Categories and Standard Definitions of Response Rates for Social Surveys, Working Papers of the Institute for Social and Economic Research (ISER), paper 2001-23. Colchester: University of Essex.

<https://www.iser.essex.ac.uk/research/publications/working-papers/iser/2001-23.pdf>

Olson, K., (Chair), Smyth, J. D. (Co-Chair), Horwitz, R., Lesser, V., Marken, S., Mathiowetz, N., McCarthy, J., O'Brien, E., Opsomer, J., Steiger, D., Sterrett, D., Su, J., Suzer-Gurtekin, Z. T., Turakhia, C., Wagner, J. (2019). Report of the AAPOR Task Force on Transitions from Telephone Surveys to Self-Administered and Mixed-Mode Surveys, found at: <https://www.aapor.org/getattachment/Education-Resources/Reports/Report-of-the-Task-Force-on-Transitions-from-Telephone-Surveys-FULL-REPORT-FINAL.pdf.aspx>.

Pierzchala, Mark M., Hart, M., Suresh, R., Rodriguez, G., Filippenko, L., Sjodin, S., Boodhumeah, N., Miceli, C., Degerdal, H., Båshus, T., Haslund, J., Maher, P., Frey, R., Bilhorn-Janssens, A. Abernathy, T., Hofman, L., and Carati, T. (2018). *Blaise 5 Multimode Management – A Report by the BCLUB Multimode Group*, available to BCLUB members on BaseCamp.

Pierzchala, Mark M. (2018b). Blaise 5 Multimode Management – A Report by the BCLUB Multimode Management Group. Paper presented at the 2018 International Blaise Users Conference, Baltimore, U.S., October 2018. This is a condensed version of BCLUB report. http://blaiseusers.org/2018/papers/4_1.pdf

Schouten, B., Peytchev, A., and Wagner, J. Adaptive Survey Design, © 2017 Chapman & Hall.