

Survey Specifications Management at Statistics Canada

Pamela Ford, Statistics Canada

1. Introduction

Computer Assisted Interviewing (CAI) plays an important role in Survey Collection at Statistics Canada. Over the years, various software packages have been used for CAI instrument development with Blaise being adopted as the standard tool in 1998.

CAI has resulted in increasingly complex questionnaires and survey instrument developers have noted that the preparation of the questionnaire specifications is one of the most time consuming aspects of instrument preparation. Moreover, multi-mode collection, which would require specifying and developing additional instruments, is being viewed as a solution to some of the current collection problems and issues. Under existing practices, adding a second or third mode of collection will result in a significant increase in the total development costs. Consequently, it was determined that the processes and methods used for instrument preparation needed to be significantly revised.

As part of Statistics Canada's Collection Modernization Initiative, the Survey Specifications Manager (SSM) is being developed to facilitate a streamlined specification development process that encourages the standardizing and recycling of questions across multiple modes of collection.

2. Problem Statement

The following items were identified as problems to be addressed within the scope of the SSM project:

- Survey development is currently stove piped often resulting in duplication of specification and development effort.
- Questionnaire requirements are difficult to specify and the specifications process is perceived to be taking too long.
- The current specification process allows for too much change through the development cycle which increases the time it takes to program, test and re-test.
- Survey input file structures are not always clearly specified and validation of such files is limited and often non-existent.
- Survey output from different collection media is formatted differently which increases complexity of subject matter processing systems.
- Not all survey metadata is readily available for dissemination products.

3. Survey Specification Process

It is necessary to formalize a process for the development of survey specifications making use of the SSM application where appropriate. At the time of the writing of this paper, the aforementioned process is far from refined but several key steps have been identified.

3.1 Mode-independent Survey Specification

One of the early steps in the process is the identification of concepts to be measured and the variables and level of detail required for analysis and dissemination. The list of required variables is essentially a mode-independent survey specification. Each variable in the specification is associated with a particular concept which in turn can be attributed to an overlaying theme.

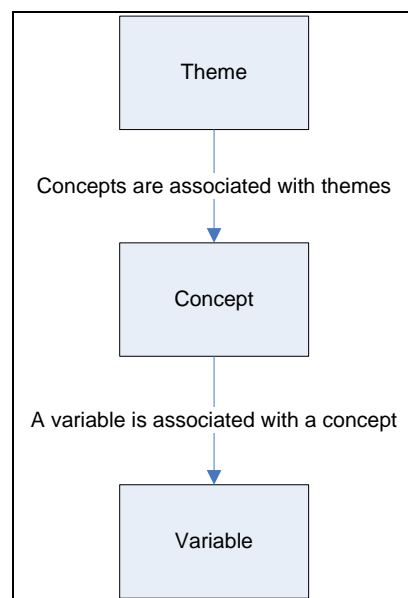


Figure 1 Relationship between Variables, Concepts and Themes

3.2 Survey Collection Requirements

Once the basic content and methodological requirements of a survey are determined it is necessary to assess the available collection options. The following questions might be asked and answered:

- Which collection methods are most suitable given the survey requirements?
- When can collection feasibly take place?
- Can the information be collected by proxy?
- What type of sample frame(s) will be used?
- Will information be fed back from previous iterations of the same survey?

3.3 Survey Instrument Specifications

Once a mode-independent specification and basic collection requirements (Sample frame, mode of collection, proxy/non-proxy etc.) are ascertained, it is possible to proceed with the specification and development of survey instruments.

Survey instruments consist of components each of which address a particular segment within a questioning process. Examples of Social Surveys components include:

- Contact – making contact with an appropriate respondent, determining if the sample unit is in scope
- Household – collecting household-level variables, rostering the household, selecting one or more respondents for which subject matter content will be collected
- Subject Matter Content
- Exit – questionnaire closing (eg. Thank you statements) and coding (registering the outcome of an attempt)

Components, in turn, consist of blocks of related questions with each block collecting data for variables associated with a common concept.

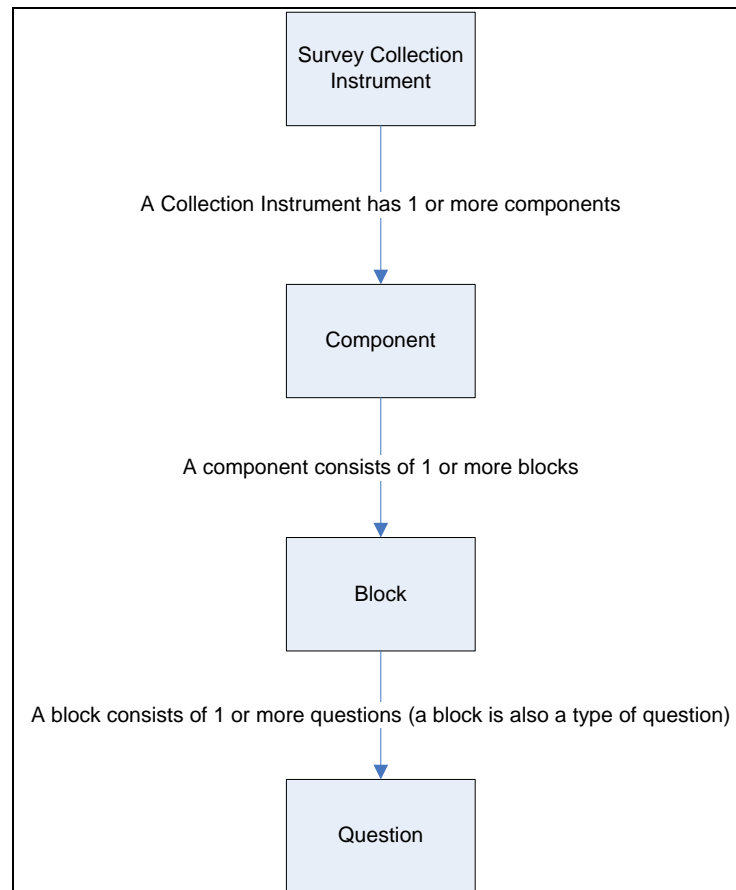


Figure 2 Survey Instrument Design

At this stage, the survey specification process should encourage and possibly even mandate the re-use of existing or standardized question blocks.

3.4 Collection Input

The survey specification process should allow for the identification and definition of collection input other than the survey instruments. This includes but is not limited to:

- Defining the structure of survey input or sample files
- Identifying key (strata) fields to be used in the active management of collection
- Identifying targets for collection at national, local, and stratum levels

3.5 Structured Metadata File

Following collection there is a desire for a structured metadata file including survey instrument metadata to be included with dissemination products. Examples of such metadata might include variable names and descriptions, proxy/non-proxy indicators, date and time of collection for each variable, response/nonresponse indicator by variable, flags to identify if edits had been triggered, etc.

4. Survey Specifications Manager

At the core of the project is the development of the Survey Specifications Manager (SSM) application itself. This involves building a central repository to house survey specifications, instrument source code, and related artifacts with a web-based user interface. There are three key components to the SSM:

- Manage Blocks
- Manage Surveys
- Search

4.1 Manage Blocks

Instrument specifications and code are organized as components or collections of “blocks” with each block consisting of one or more questions with conditional flow logic. One of the goals of the SSM is to facilitate the re-use of blocks across multiple surveys.

The ‘Manage Blocks’ feature is being designed to allow users to create and manipulate individual blocks, outside the context of a survey. This is particularly important for a team of subject matter users who are currently tasked with developing a series of harmonized questions to be used by all social and household surveys.

4.2 Manage Surveys

The 'Manage Surveys' feature will essentially provides users with a means of managing all survey specifications associated with collection. For each survey users will be able to:

- define a mode-independent survey specification
- define survey instrument specifications, making use of existing specifications when available
- define all input associated with the collection of the survey
- produce output in specific formats such as a technical specification for Blaise authors or a DDI-based XML file for survey analysts

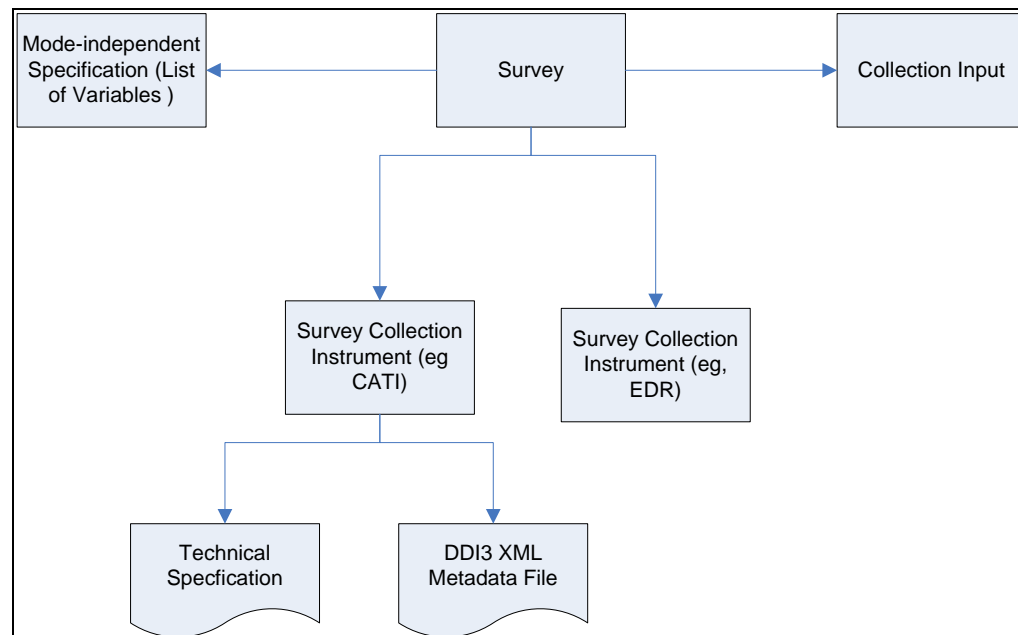


Figure 3 Survey Specification

4.3 Search

The 'Search' feature will allow users to search for questionnaire blocks by various attributes including:

- themes (e.g. Demographics, Health, Education, Employment)
- concepts (e.g. Age, Smoking,)
- variables (e.g. Age of person, Age when person first smoked)
- approval-level (Corporate Standard, Program Standard, Survey-specific)
- survey (e.g. Labour Force Survey, Canadian Community Health Survey)
- question techniques (lookups, mark-all-that-apply etc)

5. Challenges

5.1 Questionnaire Flow Logic

One of the biggest challenges in creating a survey specifications development tool is in facilitating the specification of questionnaire flow logic. CAI has resulted in complex specifications that go far beyond the simple answer-based 'go to' statement. A number of options will be investigated in order to meet this challenge, including allowing for users to specify flow using a graphical interface.

5.2 Historical Metadata

To maximize the benefits of the Survey Specifications Manager, it will be necessary to load historical metadata into the central repository. To facilitate this, plans are being made to create XML versions of historical block and questionnaire specifications which can then be easily imported into the structured SSM database. The Michigan Questionnaire Documentation System (MQDS) is being considered as a means of producing this file until such a time as it is possible to generate XML files from Blaise.

6. Conclusion

The SSM is intending to provide a one-stop shop for the management of all specifications related to survey collection. The expectation is that using this tool within a formal specification process will result in making the overall survey collection process more effective.

7. References

User Guide – Michigan Questionnaire Documentation System (April 2007), Survey Research Center, University of Michigan.

8. Contact Info

Pamela.Ford@statcan.ca