

QUESTIONNAIRE STRUCTURE DIAGRAMS

Griffith Feeney

July 1999

Abstract

Questionnaire structure is defined by skip instructions, which determine which questions are asked of which persons. Development and understanding of complex questionnaires is facilitated by representing them as flow diagrams, a device used in computer program design. A questionnaire may be thought of as a “program” that guides the enumerator through the interview. Questionnaires are better represented by “structured” flow diagrams, however, than by conventional unstructured diagrams. This note explains the concept of questionnaire structure diagrams, which apply structured flow diagrams to represent questionnaire structure.

Census and survey questionnaires have both a content and a structure. Content refers to what questions are included, structure to which persons are asked which questions and in which order. Structure arises whenever some questions are addressed to some persons and not to others. The simplest possible structure, in which all questions are addressed to all persons, rarely occurs in practice. Most questionnaires solicit information that is undefined for various groups of persons and include “skip” instructions to insure that no respondent is asked questions to which there is no answer. Never married persons, for example, will not be asked the age at which they first married.

Questionnaires

Questionnaires consist of a numbered series of items. Items may be **questions** to be addressed to the person interviewed or **enumerator instructions** indicating various actions to be taken by the **enumerator**. The person interviewed is called the **respondent**. The person to whom a given question refers is called the **subject** of this question. The subject is not necessarily the same as the respondent.

Questionnaire items are usually executed in sequence, but any item may include **skip** instructions, which direct the interviewer to some question other than the next in sequence. The structure of a questionnaire is determined by its skip instructions.

Questionnaire Structure Diagrams

The structure of simple questionnaires will be apparent from inspection. As questionnaires become more complicated, however, it advantageous to find a better representation of their structure than the questionnaire itself. **Questionnaire structure diagrams**, described in this section, provide such a representation.

Sequence The simplest questionnaire structure consists of a sequence of questions addressed to every person interviewed. It may be represented as shown in Figure 1(a). Each question or group of questions is represented by a rectangular box or block

Selection Suppose that a particular question is asked only of certain individuals. Age at first marriage, for example, applies only to ever married persons. The enumerator is thus instructed to skip this question if the subject has never been married.

This process of selection is represented as shown in Figure 2(b). The “Ever Married?” block shows two alternatives, “no” for never married persons, at left, and “yes” for ever married persons, at right. The two blocks immediately beneath indicate what the enumerator is to do in either case. On “no”, the enumerator exits the “Ever Married?” block into the empty block at left. Empty blocks signify no action, whence the enumerator to the next question in sequence. On “yes”, however, the enumerator exits the “Ever Married?” block into the “Age at first marriage” block at right and asks the respondent for the age at first marriage of the subject.

The progress of the enumerator through the questionnaire corresponds to a path through the diagram. Every block is entered only once, through the upper boundary, and exited only once, through the lower boundary. Movement across vertical block boundaries is not allowed. Where there are two or more blocks at the same horizontal level, the path must therefore pass through only one of them.

Repetition Suppose that the enumerator is to collect similar information for all entities of a given type, where the number of entities will vary from one case to another and cannot be known in advance. In enumerating a household, for example, the enumerator will collect certain information for each member of the household, or, in a fertility survey, certain information for each child a woman has given birth to. Repeated questions are represented in Figure 3(c). The slashes ‘//’ signify that the information indicated in the interior box is to be collected for all instances described by the heading above it.

Any questionnaire, no matter how long or complex, may be expressed in terms of the three basic structures illustrated in Figure 1(a-c). In practice, however, a third graphical device is important. Any questionnaire may be represented by a single diagram if we are willing to make the diagram large enough. Drawing large diagrams is problematic, however, and for practical application it must be possible to break a single large diagram into two or more smaller ones. “Callouts”, described below, provide a systematic way of doing this.

Callouts If the diagram representing a questionnaire becomes too large, one or more series of consecutive and related questions may be grouped together, given a name, and moved to a separate diagram, as illustrated in Figure 1(d). Note that the overall questionnaire is still represented by a single diagram, though some of the blocks in this diagram point to other diagrams. Callouts may be nested, but this will not usually be necessary in practice.

Relation to Computer Programming Flowcharts The diagrams described here are computer programming design and documentation tools developed by Chapin (1974) as an improvement over conventional “unstructured” flow diagrams. Conventional flow diagrams impose little discipline on the diagrammatic representation. In particular, they provide no uniform convention for handling “subroutines” or “procedures” and they place no restriction on “go to” statements. Chapin’s diagrams effectively impose the structured programming ban on “go to” statements and provide a systematic way of representing subroutines/procedures.

The connection between questionnaires and programming is established by observing that a questionnaire may be regarded as a “program” that guides the enumerator's work during the interview. Flowcharts have been used to describe questionnaire structure (US Bureau of the Census 1979), but their use is less standard than one might expect. The book *Data Collection in Developing Countries* (Dasley and Lury 1981), for example, does not mention their use. The structured flow diagrams presented here have not to my knowledge been used before to represent questionnaire structure.

Use of Questionnaire Structure Diagrams

Questionnaire structure diagrams differ from the questionnaires they represent in several important ways. The abstraction of large groups of questions into a single block provides a useful tool for understanding and developing long and complicated questionnaires. No matter how large and complex the questionnaire, the over all content and structure may be represented by a reasonably simple diagram that fits on a single page. This is accomplished by forcing large areas of detail into separate diagrams. The quantity of detail with which one is faced at any given point is in this way strictly limited.

The diagrams use two dimensions rather than one to describe the questionnaire. The vertical dimension represents the progress from one question to the next as the interview proceeds. The horizontal dimension represents decisions made by the interviewer as to which question is to be asked next. Visually separating these two aspects of the work makes it easier to discern the content and structure of the questionnaire than the sequential listing of items given in the questionnaire.

The various subgroups of persons to whom particular groups of questions are addressed may be identified with horizontal position in the diagram by imagining lines drawn down through the body of the diagram at each point in the diagram at which a vertical line appears. This identification is useful for designing and specifying edit tests and tabulations.

Questionnaires have a “logical” and a “physical” aspect. The logical aspect refers to the content and structure of the questionnaire, to what questions are asked of which persons in what order. The physical aspect refers to the size of the paper on which the questionnaire is printed, whether loose pages or booklets are used, how the questions printed, in what language, and similar considerations.

While physical aspects can be very important in data collection operations and deserve careful attention, they should be considered only after the content and the structure of the questionnaire have been decided. Questionnaire structure diagrams

abstract content and structure from format and so provide a useful tool for questionnaire design.

Structure diagrams facilitate comparison of different questionnaires. Working from the questionnaires themselves, the only mode of comparison is item by item, and this will prove most unwieldy unless the questionnaires compared are either very short or very similar. Structure diagrams will reveal at once whether there are basic similarities in the overall structure or content. If there are, one can proceed to the detail charts for further comparisons. If there are not, time will be saved by recognizing this quickly rather than as a result of a continuing failure to draw item by item correspondences.

Comparison is useful both in international studies, which one is comparing, say, population census questionnaires for different countries, and in questionnaire development, where one is comparing alternative versions of a questionnaire under development. Diagramming should also prove useful in developing questionnaires in different languages.

The June 1980 Current Population Survey Supplemental Questionnaire on Marriage and Childbearing

Figure 2 shows the overall structure of a supplemental questionnaire used in the June 1980 Current Population Survey (CPS) of the United States, designed to obtain information on marriage and birth histories. A facsimile of the questionnaire may be found in U.S. Bureau of the Census (1982). Though it consists of only 31 items, 19 of these involve skip instructions, some to as many as four different places in the questionnaire, an unusually complex structure.

The questionnaire covers four broad areas, marriage history, birth history, birth expectations, and children of previous marriages. Four characteristics determine whether the questions in any given block will be addressed to any given person, marital status, sex, and age (whether under or over 18 years only), and an item that identifies whether the subject may have children of a previous marriage living elsewhere. Each of these items corresponds to a decision block in the diagram, and to a skip instruction in the questionnaire. Item numbers for each area are indicated in the diagram to facilitate cross-reference to the questionnaire.

The structure of the questionnaire is defined by the various possible paths that may be taken down through the diagram. Because the number of blocks is small, it is a simple matter to enumerate the possibilities. The marriage history questions are addressed to all ever-married persons, males as well as females. Ever-married females are also asked the birth history and birth expectations questions, ever-married males the questions on children of previous marriages. Never married females are asked the birth history and birth expectation questions only if they are 18 years old or over. Ever married males who are currently in their second or later marriage, or who are currently widowed, separated or divorced, are asked about children of previous marriages. Never married males and never married females under 18 years of age are asked no questions.

Structure of the Marriage History

Figure 3 details the marriage history indicated in Figure 2. Given the number and complexity of the questions involved, the advantage of separating the information in these two diagrams is obvious.

The marriage history questions secure information on the first, second, and most recent marriages. Information for a person married once only will be recorded in items 40-43, which refer to most recent marriage, whereas the same information for a woman married twice or more will be recorded in items 30-33, which refer to most recent marriage. The reasons for this apparently peculiar proceeding are discussed below.

If we forget the original questionnaire for a moment and write out marriage history questions in logical sequence, we easily arrive at something like Sequence A in Figure 4. This sequence is straightforward, despite the large number of skips, because these skips follow naturally from the contingent nature of the information. Observe however that this sequence requires repeatedly asking “Still in this marriage?” and “Another marriage?” It would be possible to avoid this if we began by asking (a) current marital status (which we will in fact usually have done) and (b) number of times married. The interviewer could then execute the necessary skips by inference from this information, as indicated in Sequence B in Figure 4.

From the second item on, the information secured from Sequence B corresponds item for item to that secured from Sequence A. The difference is that the “Still in ... marriage?” and “Another marriage?” questions have been replaced by interviewer items, the interviewer inferring the answers from the information on number of marriages and current marital status.

Sequence B contains a flaw, however. A person married once only and not currently married will be asked how his or her first marriage ended, and since we know this already from current marital status, the question is redundant. The question on how the second marriage ended is similarly redundant for persons married twice and not currently married. To remove this redundancy, we need extra skip instructions, as indicated in Sequence C in Figure 4.

Observe finally that we may reduce the number of items slightly by modifying the skip instruction in the first item. Since the second five items in Sequence D are addressed only to persons married two or more times, the first three skips contained in Sequence C may be omitted. Sequence D is essentially that contained in the questionnaire and diagramed in Figure 3, the only difference being that the complications introduced by separation have been suppressed here.

The four alternative sequences of marriage history questions provide a further illustration of the use of structure diagrams. The logic of Sequence A is obvious without any diagram, but as we proceed to Sequences B-D the structure becomes more intricate and less obvious. A useful exercise in the construction of structure diagrams is to create one for each of the sequences B, C and D.

Are the advantages of Sequence D worth the logical complexities it entails? Are they worth the intricate recodes that will be necessary to extract useful information from the survey results (consider for example the recoding necessary to obtain age at first marriage in Sequence D)?

The answer to these questions depends on the importance of minimizing the number of items in the questionnaire and, perhaps most importantly, on the reaction of the respondents to questions about marital dissolution and remarriage. This will of course vary with context.

Conclusion

There is no better way to understand a questionnaire than to create a structure diagram for it, an exercise that may turn out to be more difficult than expected. Diagramming forces one to study the questionnaire systematically and completely. The results of this study are automatically recorded for future reference, and in a form that facilitates answering various questions about the data that arise in the course of any analysis. Diagramming helps avoid the errors that tend to creep into long and complicated questionnaires, and in developing new questionnaires from old ones. The “structured” diagrams described here are far more useful than conventional flow diagrams because they impose a far more stringent discipline.

References

Casley, D.J. and D.A. Lury. 1981. *Data Collection in Developing Countries*. Oxford: Clarendon Press. [HA31.2.C3] Chapter 7, The Questionnaire, pages 91-114, has nothing on flow charts, nor does rest of book.

Chapin, Ned. 1974. 'New format for flowcharts,' *Software—Practice and Experience*, Vol. 4, 341-357.

Sirken, Monroe G. 1972. *Designing forms for demographic surveys*. Laboratories for Population Statistics, Manual Series No. 3. Chapel Hill, N.C.: University of North Carolina. [HN29.S57]

U.S. Bureau of the Census. 1979. *Popstan: A Case Study for the 1980 Censuses of Population and Housing, Part B, Planning and Preparation for Population Censuses*. Statistical Training Document ISP-TR-4B, Washington, D.C. Chapter B-7, Design of Questionnaires and Administration Forms, pages. 117-143, has (old style) flow charts.

U.S. Bureau of the Census. 1982. *Fertility of American women: June 1980*, Current Population Reports, Series P-20, No. 375, U.S. GPO, Washington, D.C.

Copyright Notice

Copyright ©1999 by Griffith Feeney. All rights reserved with the following exception. You may download and print this file and give print or digital copies to other persons provided first, that the content, including this notice, is not changed in any way and, second, that any person to whom the file or print copy is given has the right to give copies to other persons on the same terms. The authoritative source for this file may be found on the world wide web at <http://www.gfeeney.com>.

Figure 1: Structured Flow Diagrams

(a) Sequence

Question 1
Question 2
Question 3

(b) Selection

Ever Married?	
No	Yes
	Age at first marriage?

(c) Repetition

//	For each birth ...
//	Date of birth
//	Sex of birth
//	Multiplicity of birth

(d) Callouts

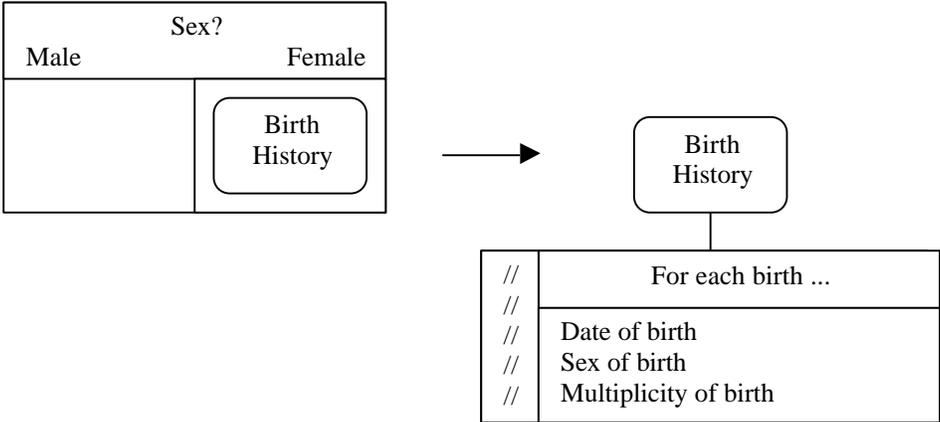


Figure 2: Supplemental Questionnaire for the June 1980 Current Population Survey (United States)

June 1980 Current Population Survey Supplemental Questionnaire Items 28-55 Civilians 15-75 Years of Age			
Ever Married? (Item 28A)			
Yes			No
Marriage History Items 29-43		Female 18+? Item 28B	
		Yes	No
Sex? Item 44			
Male		Female	
Male, currently married in 2 nd or later marriage or Male, currently widowed, separated or divorced? Item 44			
No		Yes	
Children of Previous Marriages (Items 45-46)		Birth History Items 47-52	
		Birth Expectations Items 53-55	

Figure 3: Marriage History Section of the June 1980 CPS

Marriage History – Items 29-43 – All every-married persons

1	29 How many times married?		2+
	First Marriage – Items 30-33		
	30 When did ... get married for the first time?		
	31 Did ...'s first marriage end in widowhood or divorce?		
	32 When did ...'s first marriage end in widowhood or divorce?		
	Divorced?		
	No		Yes
		33 When did ... actually stop living with his/her first spouse?	
	Second Marriage – Items 34-38		
	34 When did ... get married for the second time?		
	35 Married 3+ times?		
	Yes		No
	36 Did ...'s second marriage end in widowhood or divorce?	Currently widowed or divorced?	
		Yes	No
	37 When did ...'s second marriage end in widowhood or divorce?	Currently separated?	
		Yes	No
	38 When did ... actually stop living with his/her second spouse?		
	39 How many times married?		
	3+		2
	Most Recent Marriage – Items 40-43		
	40 What was the date of ...'s (most recent) marriage?		
	41 Currently married?		
	No		Yes
	41 Currently separated?		
	Yes		No
		42 When did ...'s most recent marriage end in widowhood or divorce?	
	Currently divorced?		
	Yes		No
	43 When did ... actually stop living with his/her (most recent) spouse?		

Figure 4: Alternative Sequences of Marriage History Questions

SEQUENCE A

Ever Married? *If no, end.*

1ST MARRIAGE
 When was first marriage?
 Still in first marriage? *If yes, end.*
 How did first marriage end?
 When did first marriage end?
 Another marriage? *If no, end.*

2ND MARRIAGE
 When was second marriage?
 Still in second marriage? *If yes, end.*
 How did second marriage end?
 When did second marriage end?
 Another marriage? *If no, end.*

MOST RECENT MARRIAGE
 When was most recent marriage?
 Still in most recent marriage? *If yes, end.*
 How did most recent marriage end?
 When did most recent marriage end?

SEQUENCE B

Number of marriages? *If none, end.*
 Marital Status? (NM, M, W, D)

1ST MARRIAGE
 When was first marriage?
If married once and currently married, end.
 How did first marriage end?
 When did first marriage end?
If married once only, end.

2ND MARRIAGE
 When was second marriage?
If married twice and currently married, end.
 How did second marriage end?
 When did second marriage end?
If married twice only, end.

MOST RECENT MARRIAGE
 When was most recent marriage?
If currently married, end.
 How did most recent marriage end?
 When did most recent marriage end?

SEQUENCE C

Number of marriages? *If none, end.*
 Current marital status? (NM, M, W, D)

1ST MARRIAGE
 When was first marriage?
If married once and currently married, end.
*If married once and currently widowed or divorced, go to **
 How did first marriage end?
 *When did first marriage end? *End.*

2ND MARRIAGE
 When was second marriage?
If married twice and currently married, end
*If married twice and currently widowed or divorced, go to ***
 How did second marriage end?
 **When did second marriage end? *End.*

MOST RECENT MARRIAGE
 *When was most recent marriage?
If currently married, end
 When did most recent marriage end?

SEQUENCE D

Number of marriages?
*If none, end; if one only, go to *.*
 Current marital status? (NM, M, W, D)

1ST MARRIAGE
 When was first marriage?
 How did first marriage end?
 When did first marriage end?

2ND MARRIAGE
 When was second marriage?
If married twice and currently married, end.
*If married twice & currently widowed/divorced, go to **.*
 How did second marriage end?
 **When did second marriage end? *End.*

MOST RECENT MARRIAGE
 When was most recent marriage?
If currently married, end.
 When did most recent marriage end?