Metadata Management
Questions, Measures and Questionnaires

Hayley Mills

12 September 2019

CLOSER, UCL Institute of Education

Overview

- Questions and measures

- Questionnaires

- Parameters and bindings

Focus on DDI-Lifecycle

- Codebook is variable centric – it only provides a partial description of question as the source of data for a variable

- Lifecycle supports stand-alone question specification and management

- Lifecycle can describe the use of a question within a questionnaire flow-logic

Focus on DDI-Lifecycle

Codebook- Example

IHSN - 5th Census of Population 1992 - IPUMS Subset
El Salvador

Lighting type (SV1992A_0041)

CSV JSON

Data file: SLV1992-H-H

Overview

vald **0**	Interval **discrete**
invd **0**	Decimal **0**
Range **1 - 9**	

Questions and instructions

LITERAL QUESTION
II. Household information

2. Characteristics of the dwelling

[Questions 4-15 were asked of occupied private households, per question 3 and Section 1.]

10. What type of lighting is used in the dwelling?

[] 1 Electricity
[] 2 Kerosene (gas)
[] 3 Other (specify) _____

CATEGORIES

Value	Category
1	Electricity
2	Kerosene
3	Other
9	NIU (not in universe)

Variables, questions and measurements

Variable

- Description of data

- A variable can come from a question or measurement

person_est
140
180
370

Question

- Describes a means of capturing data

- Specifies a text and the form of the expected response

- Questions can be organised in an instrument

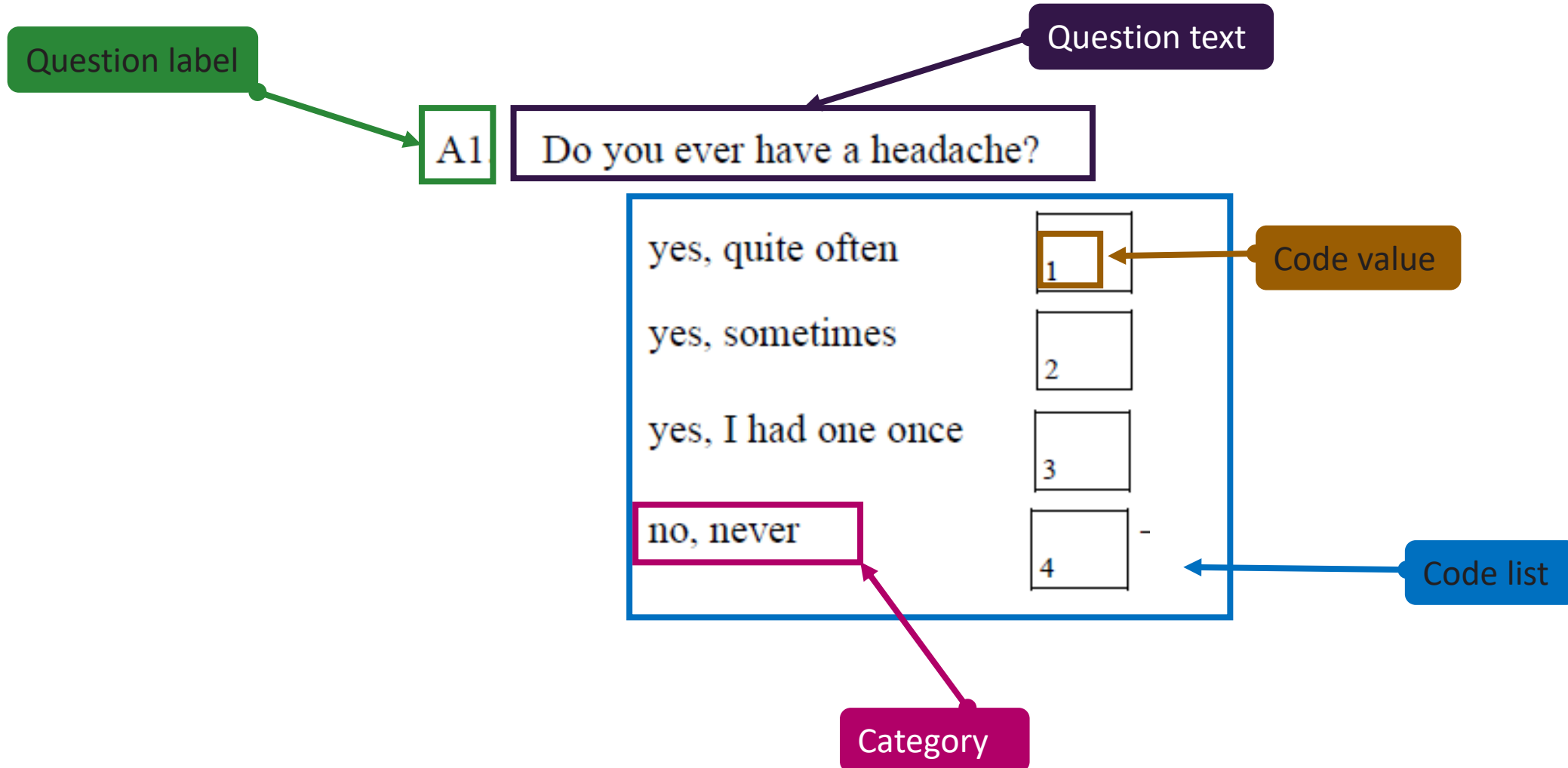Q1 How tall are you in inches?

inches

Measurement

- Describes a means of capturing data

- Specifies the measurement and the form of the expected response

- Measurements can be organised in an instrument

Measured height

inches

Questions in Questionnaires

Question label

Question text

A1 Do you ever have a headache?

yes, quite often   1   Code value

yes, sometimes   2

yes, I had one once   3

no, never   4   Code list

Category

Questions in CLOSER Discovery

discovery.closer.ac.uk

Questions and Measures

Question Item	Measurement Item (DDI-L 3.3)	Description
Question Item Name	Question Item Name	Name
Label	Label	Label
Question Text	Type Of Measurement Item	Question, or description of measure
Question Intent	Measurement Intent	What are you trying to measure
Response Domain	Response Domain	How are you classifying the response
Response Cardinality	Response Cardinality	Number of allowed responses
Concept Reference	Concept Reference	Concept being captured
External Aid	External Aid	What might assist in response
Interviewer Instruction	Interviewer Instruction	Clarification, information to obtain consistent responses
Parameters / Bindings	Parameters / Bindings	Regulate the flow of information within an instrument
Represented Variable	Represented Variable	'Template' for the data produced

Response Domains

- A response domain defines response options to a question

- There may be more than one response domain for a question

- Response domains clarify the bounds of accepted valid and invalid values for the question

- Response domains capture a response that may then be recoded for entry in a data file

- They are not the same as variable value representations

Style of Response	Description
Code	Pick list (codes and associated categories)
Category	Pick list set of categories (no associated code)
Numeric	Number field can be defined by type, range, and precision
Text	Text field can be defined by length and regular expression
DateTime	Date and/or Time field can be defined by format, range, and regular expression
Geographic	Set of fields to capture GPS position
Geographic Structure	Pick list for geographic levels (State)
Geographic Location	Pick list for individual geographic locations (Alabama)
Scale	Layout of a scale (Likert or similar)
Distribution	Layout of a distribution (assigning %)
Ranking	Ordering items
Location	Specifying the location on an image, sound, video, etc
Nominal	Marked or unmarked

Question   Related Variable

Question   Related Variables

Na

La

Question

Multiple Choice Respons

Co

- 1   yes, quite often
- 2   yes, sometimes
- 3   yes, I had one once
- 4   no, never

ccaf100 - A1: Child ever had a headache. CCAF file

Dataset

Value	Label	Frequency	
-11	Triplet / quadruplet	13	
-10	Not completed	8087	███████████████
-1	Missing	22	
1	Yes, quite often	580	█
2	Yes, sometimes	4365	████████
3	Yes, I had one once	1325	█
4	No, never	1066	█

Valid	Invalid	Min	Max	Mean
7336	8122	1	4	2.39

Use of a Question

Managed within a Question Bank

- Made available for later use

- Managed in terms of version changes

Assembled into a Questionnaire

- Ordered

- Adding statements and information specific to the questionnaire

- Routing based on responses

Questionnaires

- A questionnaire mainly uses questions as a means of capturing data

- It describes the sequence of questions and measures, intervening text, and conditional routing

- DDI Lifecycle describes questionnaire starting with the Instrument item type

Things For You to Do

Statement

Thank you for filling this in. Children of the 90s loves to look at the things you draw!

Question Item

A1. Do you ever have a headache?

yes, quite often [1]

yes, sometimes [2]

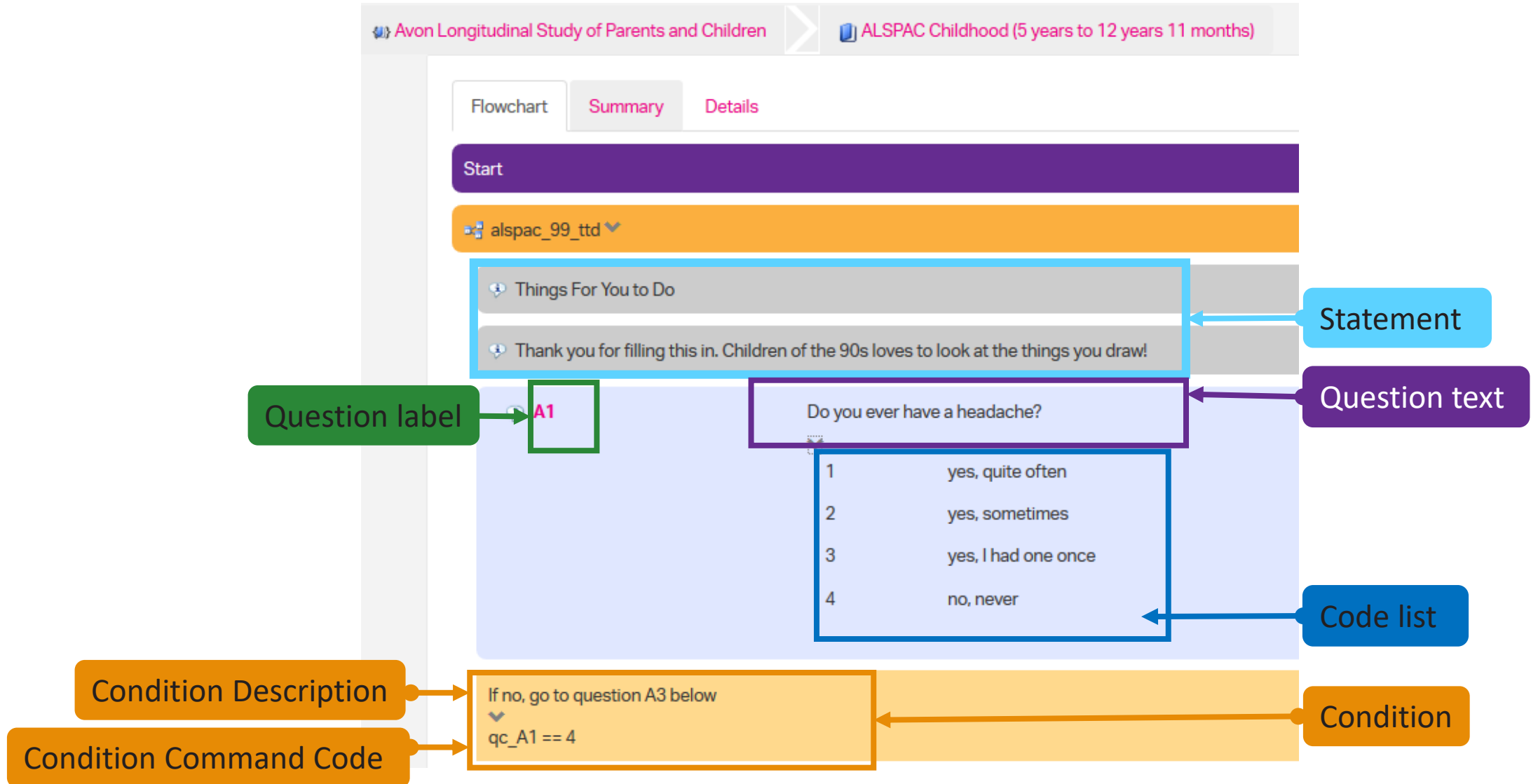yes, I had one once [3]

no, never [4] → **If no, go to question A3 below**

Condition

A2. For your last headache, please shade in where the pain was in these two pictures of a head.

Questionnaires in CLOSER Discovery

Types of Control Constructs

Control Construct	Type	Description
Sequence	Structural	A section of a questionnaire which contains Control Constructs. All Control Constructs fall within a single Master Sequence
StatementItem	Structural	Allows for the insertion of a statement (text or object)
ComputationItem	Structural	Supports Conditional Constructs and can insert validation or other checks
QuestionConstruct	Structural	Inserts a Question
IfThenElse	Conditional	Supports splitting the path through a questionnaire based generally on the response to one or more questions
Loop	Conditional	Repeat the sequence of the loop until the condition is satisfied. Sets up initial value, step, and condition
RepeatWhile and RepeatUntil	Conditional	Repeat "while" a condition is true or "until" a condition is met

Creating questionnaire flows

- If you want to reuse a specific order of questions create a Sequence and reuse it (e.g. block of questions in a panel study)

- It is easier to manage bundles of sequences than a long string of questions; organise your questionnaire flow

- If you need to manage the flow of data through your questionnaire use Input / Output Parameters and Binding

Statements live in <StatementItem>

- URN – unique reference for an element

- Display Language

- Name

- LiteralText

```xml
<d:StatementItem>
  <r:URN>urn:ddi:uk.alspac:alspac_99_ttd-si-015950:1.0.0</r:URN>
  <d:ConstructName>
    <r:String xml:lang="en-GB">s_intro_ii</r:String>
  </d:ConstructName>
  <d:DisplayText audienceLanguage="en-GB">
    <d:LiteralText>
      <d:Text>Thank you for filling this in. Children of the 90s
      loves to look at the things you draw!</d:Text>
    </d:LiteralText>
  </d:DisplayText>
</d:StatementItem>
```

Thank you for filling this in. Children of the 90s loves to look at the things you draw!

Statement

Questions live in <QuestionItem>

- URN

- Display Language

- Question Name

- Question Label

- Question Text

- Response Domain (Code, Numeric, Text)

A1. Do you ever have a headache?

Question text

yes, quite often   1

yes, sometimes   2

```xml
<d:QuestionItem>
  <r:URN>urn:ddi:uk.alspac:alspac_99_ttd-qi-100853:1.0.0</r:URN>
  <r:UserAttributePair>
    <r:AttributeKey>extension:Label</r:AttributeKey>
    <r:AttributeValue>{"en-GB":"A1"}</r:AttributeValue>
  </r:UserAttributePair>
  <d:QuestionItemName>
    <r:String xml:lang="en-GB">qi_A1</r:String>
  </d:QuestionItemName>
  <d:QuestionText audienceLanguage="en-GB">
    <d:LiteralText>
      <d:Text>Do you ever have a headache?</d:Text>
    </d:LiteralText>
  </d:QuestionText>
  <d:CodeDomain>
    <r:CodeListReference>
      <r:URN>urn:ddi:uk.alspac:alspac_99_ttd-cl-040680:1.0.0</r:URN>
      <r:TypeOfObject>CodeList</r:TypeOfObject>
    </r:CodeListReference>
    <r:ResponseCardinality minimumResponses="1" maximumResponses="1"/>
  </d:CodeDomain>
</d:QuestionItem>
<d:QuestionItem>
```

Codelist - codes that reference categories

- URN
- Display Language
- Label
- Value
- Category reference

Code list

Code value

```
</l:CodeListSchemeName>
<l:CodeList>
  <r:URN>urn:ddi:uk.alspac:alspac_99_ttd-cl-040680:1.0.0</r:URN>
  <r:Label>
    <r:Content xml:lang="en-GB">cs_yqo_ys_yihoo_nn</r:Content>
  </r:Label>
  <l:Code>
    <r:URN>urn:ddi:uk.alspac:alspac_99_ttd-co-197436:1.0.0</r:URN>
    <r:CategoryReference>
      <r:URN>urn:ddi:uk.alspac:alspac_99_ttd-ca-156516:1.0.0</r:URN>
      <r:TypeOfObject>Category</r:TypeOfObject>
    </r:CategoryReference>
    <r:Value>1</r:Value>
  </l:Code>
  <l:Code>
    <r:URN>urn:ddi:uk.alspac:alspac_99_ttd-co-197437:1.0.0</r:URN>
    <r:CategoryReference>
      <r:URN>urn:ddi:uk.alspac:alspac_99_ttd-ca-156517:1.0.0</r:URN>
      <r:TypeOfObject>Category</r:TypeOfObject>
    </r:CategoryReference>
    <r:Value>2</r:Value>
  </l:Code>
</l:Code>
```

Categories

- URN

- Display Language

- Category Name

- Label

```
<l:Category>
  <r:URN>urn:ddi:uk.alspac:alspac_99_ttd-ca-156517:1.0.0</r:URN>
  <l:CategoryName>
    <r:String xml:lang="en-GB">156517</r:String>
  </l:CategoryName>
  <r:Label>
    <r:Content xml:lang="en-GB">yes, sometimes</r:Content>
  </r:Label>
</l:Category>
```
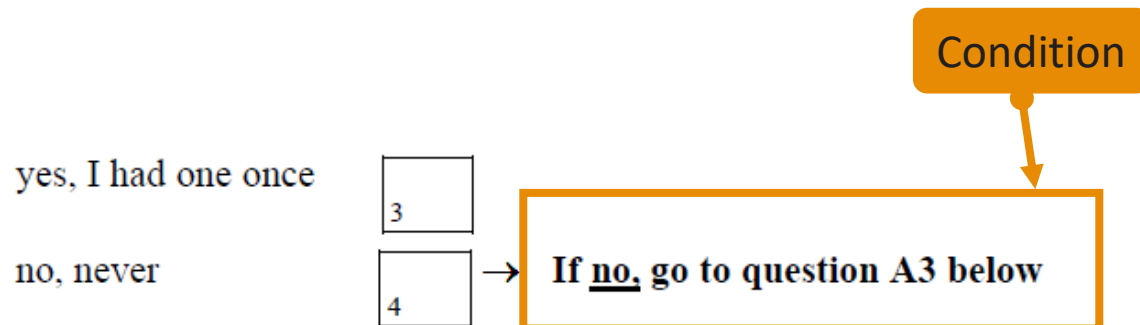
yes, quite often | 1 |

yes, sometimes | 2 | ← Category

yes, I had one once | 3 |

no, never | 4 |

Conditions - IfThenElse

- URN
- Display Language
- Construct Name
- Description
- Command

```xml
<d:IfThenElse>
  <r:URN>urn:ddi:uk.alspac:alspac_99_ttd-if-018736:1.0.0</r:URN>
  <d:ConstructName>
    <r:String xml:lang="en-GB">c_qA1</r:String>
  </d:ConstructName>
  <d:IfCondition>
    <r:Description>
      <r:Content xml:lang="en-GB">If no, go to question A3 below</r:Content>
    </r:Description>
    <r:Command>
      <r:ProgramLanguage>pseudo-code</r:ProgramLanguage>
      <r:CommandContent>qc_A1 == 4</r:CommandContent>
    </r:Command>
  </d:IfCondition>
```

Condition

yes, I had one once [3]

no, never [4] → **If <u>no,</u> go to question A3 below**

Constructs combine to make a questionnaire

```xml
<d:ControlConstructReference>
  <r:URN>urn:ddi:uk.alspac:alspac_99_ttd-si-015950:1.0.0</r:URN>
  <r:TypeOfObject>StatementItem</r:TypeOfObject>
</d:ControlConstructReference>
<d:ControlConstructReference>
  <r:URN>urn:ddi:uk.alspac:alspac_99_ttd-qc-112120:1.0.0</r:URN>
  <r:TypeOfObject>QuestionConstruct</r:TypeOfObject>
</d:ControlConstructReference>
<d:ControlConstructReference>
  <r:URN>urn:ddi:uk.alspac:alspac_99_ttd-if-018736:1.0.0</r:URN>
  <r:TypeOfObject>IfThenElse</r:TypeOfObject>
</d:ControlConstructReference>
<d:ControlConstructReference>
```

```xml
<d:StatementItem>
  <r:URN>urn:ddi:uk.alspac:alspac_99_ttd-si-015950:1.0.0</r:URN>
  <d:ConstructName>
    <r:String xml:lang="en-GB">s_intro_ii</r:String>
  </d:ConstructName>
  <d:DisplayText audienceLanguage="en-GB">
    <d:LiteralText>
      <d:Text>Thank you for filling this in. Children of the 90s
      loves to look at the things you draw!</d:Text>
    </d:LiteralText>
  </d:DisplayText>
</d:StatementItem>
```

```xml
<d:IfThenElse>
  <r:URN>urn:ddi:uk.alspac:alspac_99_ttd-if-018736:1.0.0</r:URN>
  <d:ConstructName>
    <r:String xml:lang="en-GB">c_A1</r:String>
  </d:ConstructName>
  <d:IfCondition>
    <r:Description>
      <r:Content xml:lang="en-GB">If no, go to question A3 below</r:Content>
    </r:Description>
    <r:Command>
      <r:ProgramLanguage>pseudo-code</r:ProgramLanguage>
      <r:CommandContent>qc_A1 == 4</r:CommandContent>
    </r:Command>
  </d:IfCondition>
  <d:ElseConstructReference>
    <r:URN>urn:ddi:uk.alspac:alspac_99_ttd-seel-018736:1.0.0</r:URN>
    <r:TypeOfObject>Sequence</r:TypeOfObject>
  </d:ElseConstructReference>
</d:IfThenElse>
```

```xml
<d:QuestionItem>
  <r:URN>urn:ddi:uk.alspac:alspac_99_ttd-qi-100853:1.0.0</r:URN>
  <r:UserAttributePair>
    <r:AttributeKey>extension:Label</r:AttributeKey>
    <r:AttributeValue>{"en-GB":"A1"}</r:AttributeValue>
  </r:UserAttributePair>
  <d:QuestionItemName>
    <r:String xml:lang="en-GB">qi_A1</r:String>
  </d:QuestionItemName>
  <d:QuestionText audienceLanguage="en-GB">
    <d:LiteralText>
      <d:Text>Do you ever have a headache?</d:Text>
    </d:LiteralText>
  </d:QuestionText>
  <d:CodeDomain>
    <r:CodeListReference>
      <r:URN>urn:ddi:uk.alspac:alspac_99_ttd-cl-040680:1.0.0</r:URN>
      <r:TypeOfObject>CodeList</r:TypeOfObject>
    </r:CodeListReference>
    <r:ResponseCardinality minimumResponses="1" maximumResponses="1"/>
  </d:CodeDomain>
</d:QuestionItem>
```
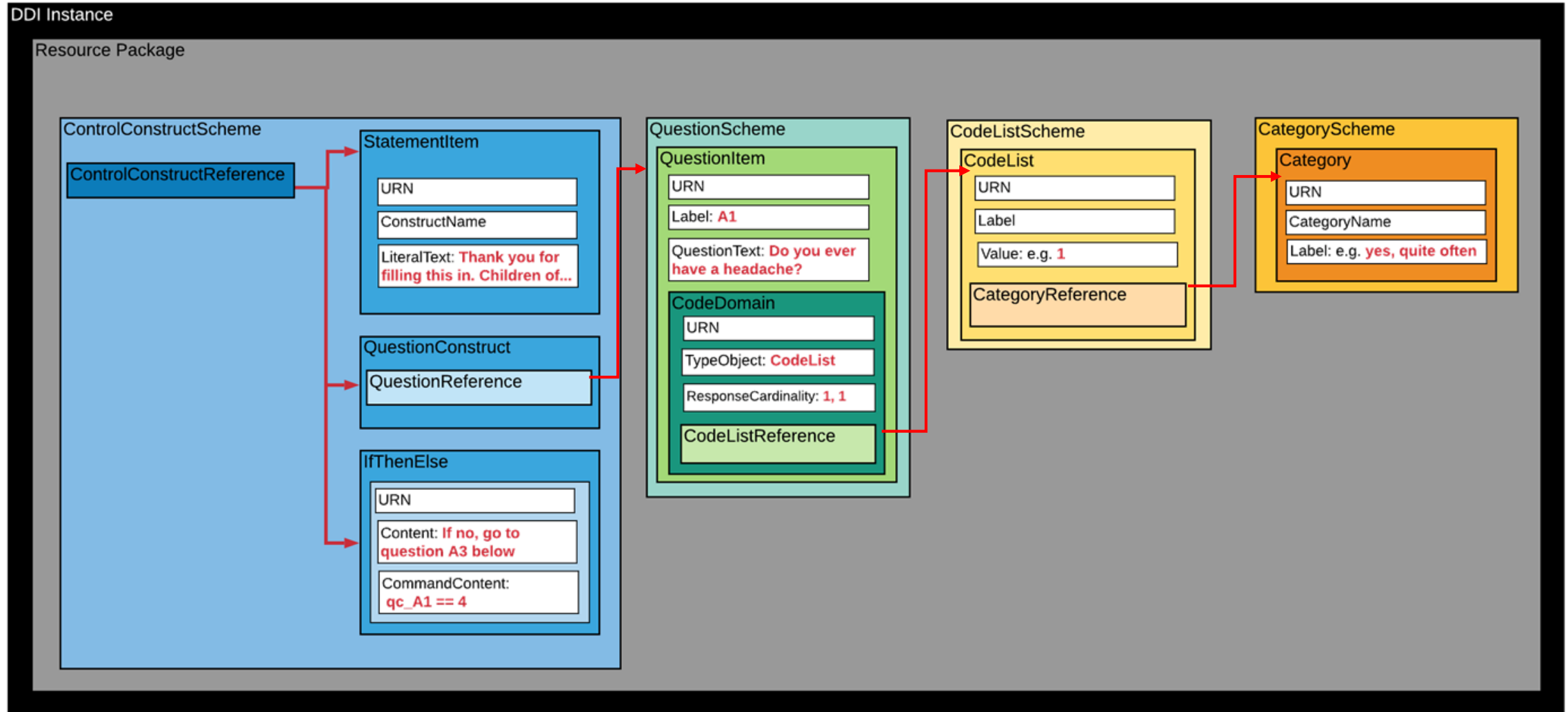
Relationships in a questionnaire - Example
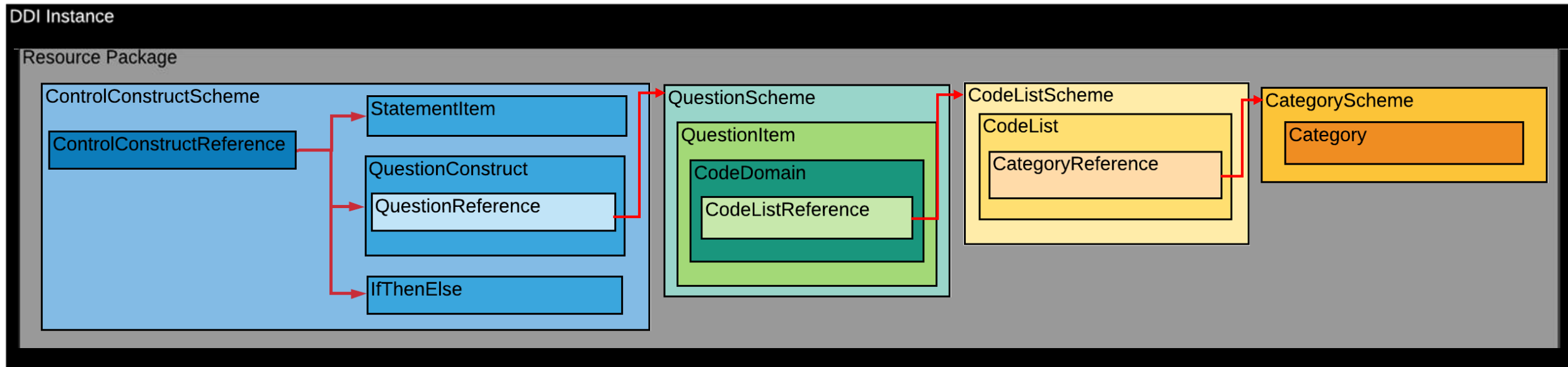
# Relationships in a questionnaire - Example

Using parameters and binding

- If you need to manage the flow of data through your questionnaire use Input / Output Parameters and Binding

- In and out parameters assign an ID to the object entering a particular 'box' or exiting that 'box'.

- Regulate the flow of information within an instrument

- It is recommended to track the parameter value at each step of the DDI questionnaire flow.

Parameters and binding - Schematic
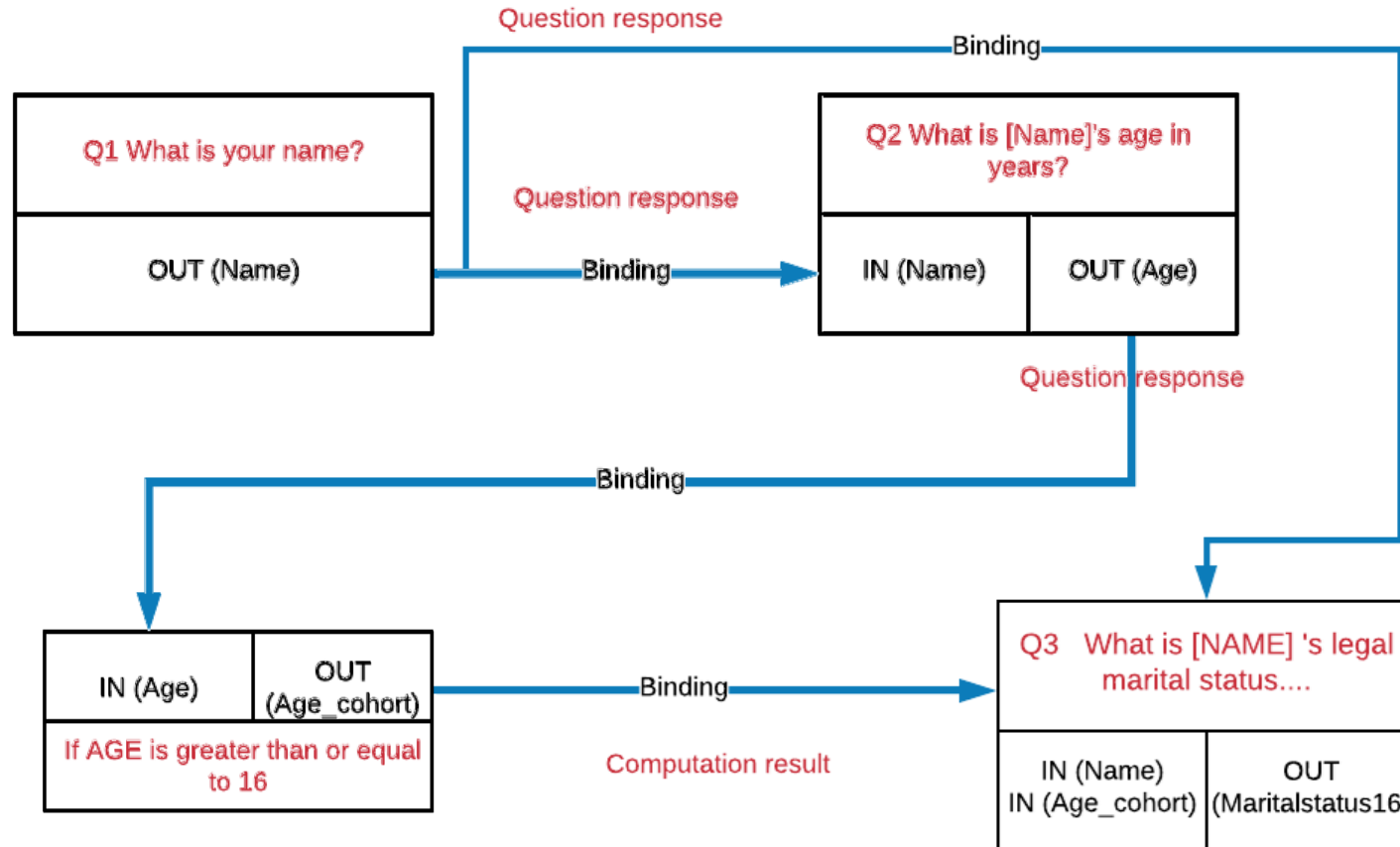
Summary

- Documenting questionnaires and other data collection instruments provides rich information about both what was asked, the responses, and of whom and by who the questions were asked

- This can be linked to variables, concepts, universes and other information that makes the data collection process more transparent, and provides a level of provenance for reproducibility

- Structures in DDI-Lifecycle can assist better management of both question development and questionnaire specification

- Reuse of questions both where they have been reused, and as a resource for future reuse creates a rich resource for harmonisation