

Survey Data Collection with Survey Solutions

Perugia, Italy

Survey Solutions: Advanced Designer

Sergiy Radyakin
sradyakin@worldbank.org

Development Data Group (DECDG),
The World Bank

May 28 - June 01, 2018



THE WORLD BANK



Food and Agriculture
Organization of the
United Nations

1 Novel Question Types

- Picture
- Audio recording
- Barcode
- Masked edit

2 Large number of options

- Combo box
- Cascading
- Filtering

3 Rosters

- Numeric
- List
- Multichoice
- Fixed

4 Nested Rosters

5 Linked Questions

Picture

- Picture type question can be used to capture the appearance of the respondent, house, neighborhood, product, etc.
- There are no special settings for the picture type question. The interviewer will capture the image by pressing the button in the questionnaire, pointing the camera at the object, and confirming the image is satisfactory.
- An autofocus lense capability may be essential for some images. Shooting conditions are not always favorable for simplest lenses of some tablets.

NB: adding images to questionnaire increases the amount of data circulated in the system and may considerably increase the wireless bill.

Audio Recording

Respondent's knowledge of foreign languages

Read the text from the card A1.

INSTRUCTION

INTERVIEWER: handle the card A1 with sample of text to the respondent. Start the recording when the respondent is ready to read. Stop when (s)he finishes or at the 60 seconds mark.

Audio recording X

Read the text from the card A1.

00:11:81

CANCEL DONE

INTERVIEWER: handle the card A2 with sample of text to the respondent. Start the recording when the respondent is ready to read. Stop when (s)he finishes or at the 60 seconds mark.

Click to record audio

- Audio type question can be used to record the voice of the respondent.
- Recording is not hidden, but button-activated.
- Recording is limited to 3 minutes.
- Quality varies with tablets and environment.

NB: adding audio recording to questionnaire increases the amount of data circulated in the system and may considerably increase the wireless bill.

Barcode

- Barcode question allows to capture barcodes and QRcodes:



- Note that there are many different standards of barcodes, with different capabilities and appearances. Survey Solutions will try to automatically detect the type of the barcode.
- Questionnaire developers need to test the system using an example of the actual barcodes to be scanned to ensure compatibility.
- Autofocus capability on the tablet is essential for barcode recognition to work properly.
- Not all items will have barcodes, and ones that had them may have them damaged and unrecognizable. Consider direct entry to override.

Text conforming to a pattern

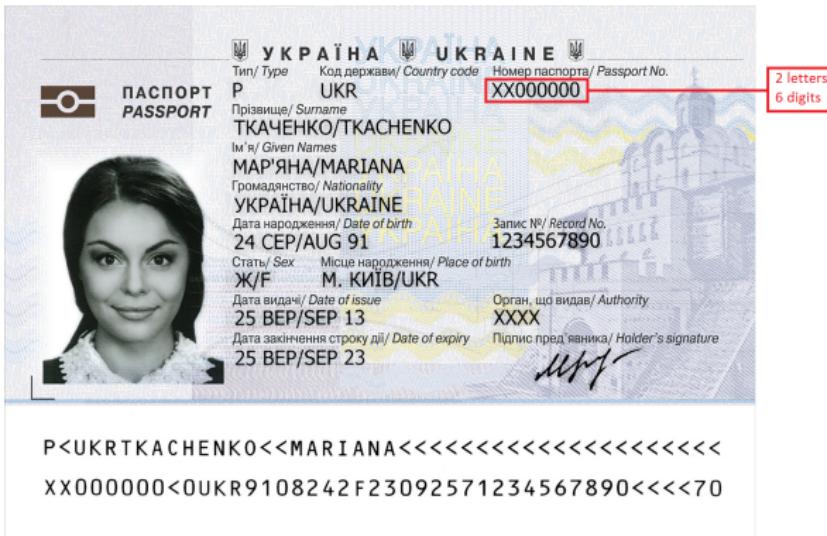
- Text questions may optionally include a pattern (aka mask) for content;
- The pattern reflects the format of the value to be entered;
- Typical use is entering various ID numbers, phone numbers, etc.

Example



- These two identifier numbers may not be entered as “numeric” type.
- The first ID in the example is a composite ID comprised of the date code and unique code combined with a dash in the middle.
- Second ID requires leading zeroes. Numeric questions will not retain leading zeroes!

Example 2



- This passport number is not a number to be entered as “numeric” type.
 - It consists of two letters and six digits. Letters may not be entered into a numeric field.

Mask structure

- Rules for pattern specification:
 - a tilde (~) represents one alpha character (A..Za..z);
 - a hash sign (#) represents one numeric character (0..9);
 - a star (*) represents one alphanumeric character in that position;
 - any other characters entered in the pattern will be visible on the screen during data entry, while the placeholders for data entry will be rendered as underscores (_).
- So the patterns appropriate for the examples above are:
 - for the record number: #####-#####
 - for the document number: #########
 - for the passport number: ~~#####

Setting Pattern in Designer

MAIN /

question type

AB	Text	▼
----	------	---

variable name (?)

passnum

variable label (?)

--

Question text

INTERVIEWER: enter the respondent's passport number

Pattern (?)

~~#####

Example

- Note that:
 - pattern enforces fixed length of content;
 - pattern enforces “hard” validation.
- If using it for e.g. phone numbers, make sure all (potential) phone numbers are of the same length.
For example: “(###) ###-####” will allow to enter typical USA phone numbers (10 digits) but will not allow entering foreign numbers.

Large number of options

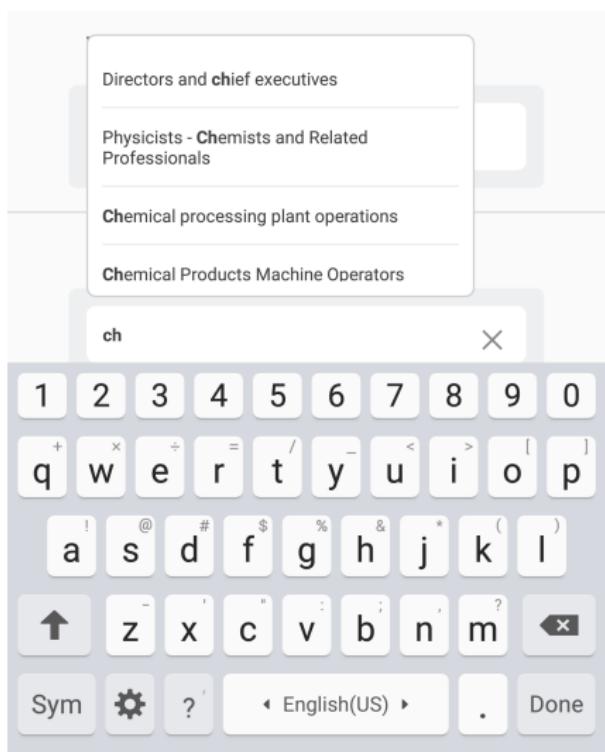
Selection among the large number of options is common when selecting a product code, occupation code, industry code, etc. Several strategies:

- select from shortlist (autocomplete);
- cascading selection (sequential, multistep);
- filtering based on known information.

Combo box (autocomplete)

- Combo box question is a particular kind of a single choice categorical question.
- It is convenient and even a necessity, when the number of choices is large and cannot be represented on a tablet screen.
- A user makes a selection from the autocomplete control by typing the beginning of the item, then selecting from a short list.

Combo box (autocomplete)



Combo box 2

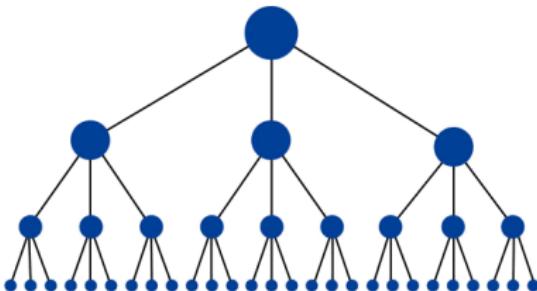
To create an autocomplete question:

- Prepare a file of choices: tab delimited file with two columns numeric code and text label.
- In designer add a categorical single select question.
- Select the checkbox: is filtered combobox.
- Upload the prepared file of options.
- Save your changes.

Cascading selection



Cascading selection



- Cascading selection is used when the number of items is large and there is a certain hierarchy among them.
- For example, administrative divisions:
 - State
 - County
 - Town
- Other examples: industries, occupations, products and goods classifications, etc.

Cascading selection 2

- Cascading selection is implemented as a number of single choice questions, one for each level of selection.
- The first question is a simple single choice question.
- Subsequent single choice questions must have “*is cascading*” flag set and parent question specified.
- Items of these questions must be prepared in a special way. For a normal single choice question the list of items contains only two columns: code and label, but for the cascading choice questions the list of items must contain three columns: code, label, and parent code.
- Items in the list must have unique codes (values), but may have repetitive parent codes (say, all of the districts within one state).

Filtering

- Cascading selection is a special form of filtering, where each item has only one parent group;
- In general an item may belong to multiple groups;
- Filtering allows reducing the number of options based on information captured earlier in the interview;
- For example: reduce number of occupational codes based on education category or reduce options for relation to the household head.
- A filter is a logical expression evaluated for every option of the categorical selection question. Refer to @optioncode in the expression to distinguish between the options.
- Here options 4 and above will only be available for respondents aged more than 15:

(age>15) || (@optioncode<4)

Rosters

- Rosters are not a question type, they are a special kind of groups.
- The purpose of the roster is to put structure on the questionnaire and to allow replication of questions and logic.
- Typical roster: demographic section of a household questionnaire, which asks questions about age, gender, relationship, education, etc of household members
- Every time you see a table in the paper style questionnaire, it is likely a rudimentary presentation of a roster.
- On paper the roster is usually a fixed table, with a certain number of rows reserved for the items. Since there is no knowledge about the number of e.g. the household members in a particular household an excessive amount is usually reserved.
- Rosters may sound complicated, but they are very helpful in questionnaire development.

Roster triggers

- In Survey Solutions the rosters can grow or shrink to accommodate the changing number of items to be surveyed.
- This change is triggered by a special question, the **trigger question**, which can be one of several types:
 - Numeric question: when the number of items to be described is determined by a numeric question somewhere earlier in the questionnaire.
 - Text list question: when there is a list of items already collected earlier in the questionnaire, and additional information about these items is to be acquired.
 - Multiple choice question: when a list of choices is presented first, and then detailed information is acquired about each item that was selected.
- One may use a fixed roster, when the number of items is known at design time.

Numeric roster

- Triggered by a numeric question. The numeric question should precede the roster in the order of the questions in the questionnaire.
- Typical use: *How many people live in this household? Now let me ask about name, age, gender, education, and employment of each of them.*

Example of a numeric roster

PAGE 60

A2.2 Non-agricultural

During the period JULY to DECEMBER 2015, did you or any member of your family receive salaries and wages from non-agricultural sectors as seasonal/occasional workers (e.g. building construction laborer) in cash (allowances, honoraria, tips, bonus, and commissions) and in kind (food, etc.)?

1 - YES

2 - NO, GO TO NEXT PAGE (B)

Line No.	First Name of Family Member	Occupation	[Occupation Code]	Kind of Industry	[Industry Code]	CASH EARNINGS		EARNINGS IN KIND			
						Basic Salaries and Wages	Allowances, Honoraria, Tips, etc.	Total	Basic Salaries and Wages	Housing, Clothing, Food, etc.	Total
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
—	—	—	—	—	—	—	—	—	—	—	—
—	—	—	—	—	—	—	—	—	—	—	—
—	—	—	—	—	—	—	—	—	—	—	—
—	—	—	—	—	—	—	—	—	—	—	—
—	—	—	—	—	—	—	—	—	—	—	—

Text list roster

- Triggered by a text list question. This question should precede the roster in the order of the questions in the questionnaire.
- Typical use: *Tell me the names of the people who live in this household. Now let me ask about the age, gender, education, and employment of each of them.*
- In contrast to the numeric version you already have the names, which can be displayed in the text of the question with an inline substitution `%rostertitle%`.

Multiselect roster

- Triggered by a multiple select categorical question. This question should precede the roster in the order of the questions in the questionnaire.
- Typical use: *Which of the following appliances do you have in this household? (fridge, freezer, electric stove, gaz stove, microwave oven, TV, radio, phone). Now let me ask about the original price and age of each of the items you mentioned.*
- Item name can be displayed in the text of the question with an inline substitution `%rostertitle%`.

Fixed roster

- This roster does not change its size, it is not triggered by anything.
- The size of a fixed roster is determined by an internally stored list of items determined by the questionnaire designer. Each item should have a code and a label, of which the label is shown to the interviewer, and the code is stored as a value.
- Typical use: *Tell me about the amount of income that this household derives from the following sources: labor income, rent, domestic transfers, international transfers, other.*
- Item name can be displayed in the text of the question with an inline substitution `%rostertitle%`.
- One can't delete or add items to fixed rosters during an interview. But one can turn on and off questions depending on the situation.

Example of a fixed roster

PAGE 70

A2. LIVESTOCK AND POULTRY RAISING

- (c) During the period **JULY to DECEMBER 2015**, did you or any member of your family incur expenses, in cash or in kind, in raising the livestock and poultry disposed of, or in producing the livestock/poultry products?

1 - YES 2 - NO, GO TO NEXT PAGE (A3)

Item	Code	In Cash	In Kind	Total
1. Acquisition cost of stock	002310	_____	_____	_____
2. Feeds	002320	_____	_____	_____
3. Medicine	002330	_____	_____	_____
4. Labor	002340	_____	_____	_____
5. Fuel and oil	002350	_____	_____	_____
6. Electricity	002360	_____	_____	_____
7. Other expenses	002370	_____	_____	_____

Setting up a roster

- Click the *Add roster* button in the sections tree or select *Add roster after* item from the context menu.
- Give roster a title that the interviewer will see, and an identifier, by which you will refer to it at data processing.
- Select roster type (one of the four we discussed).
- Select the trigger question if applicable, or specify the items of the fixed roster.
- Add questions to the roster. An empty roster without any questions makes no sense!

Setting up a roster in designer

Demographics

12 Please, tell me the names of all members of this household. memberlist

Roster: **HOUSEHOLD MEMBERS** hhmembers

12 Relationship relhead

12 What is the gender of %rostername%? gender

12 What is the age of %rostername% in completed years? age

12 Does %rostername% attend any formal school? attends_school

12 Did %rostername% work for pay outside of the household on an works

[ADD QUESTION](#) [ADD SUB-SECTION](#) [ADD ROSTER](#) [ADD STATIC TEXT](#)

DEMOGRAPHICS /

Roster source

List question

Roster variable name (?)

hhmembers

Roster name

Household members

Source question (answer determines both the number and the name of roster rows):

DEMOCRAPHICS

12 Please, tell me the names of all members of this household.

Enabling condition (?)

SAVE CANCEL DELETE CLONE MOVE TO ▲

Trigger question changes

- When the value of a numeric trigger question increases, the number of lines corresponding to the increase is added by Survey Solutions to the end of the table. If any data was entered regarding any items before the increase it survives.
- When the value of a numeric trigger question decreases, the number of lines corresponding to the decrease is deleted from the end of the table. If any data was entered regarding the items being deleted, it will be forgotten. The user will get a warning.
- Since the text list type question allows deletion of items from the middle, the user can remove exactly the one item she wants, not necessarily the last one.
- Since the multiple select type question allows unchecking of any items, including items in the middle, the user can remove exactly the one item she wants, not necessarily the last one.

Roster conditions

- One can add enabling conditions to any questions in the roster.
- These conditions can refer to all the variables of the questionnaire-level (typically household).
- If a roster variable is mentioned, it is interpreted as the variable referring to the same item. For example, if employment status has a condition $age > 15$ (individual-level variable age is mentioned), the value of age will be taken from the same person.
- One can add a condition on the roster itself. This will determine which items will be included into the roster. For example, one can have a numeric question “Household size”, followed by a roster “Household members”, then another roster “Adults” triggered by the same question, but with a condition $age \geq 14$, referring to the variable age from the first roster.

Selection of members

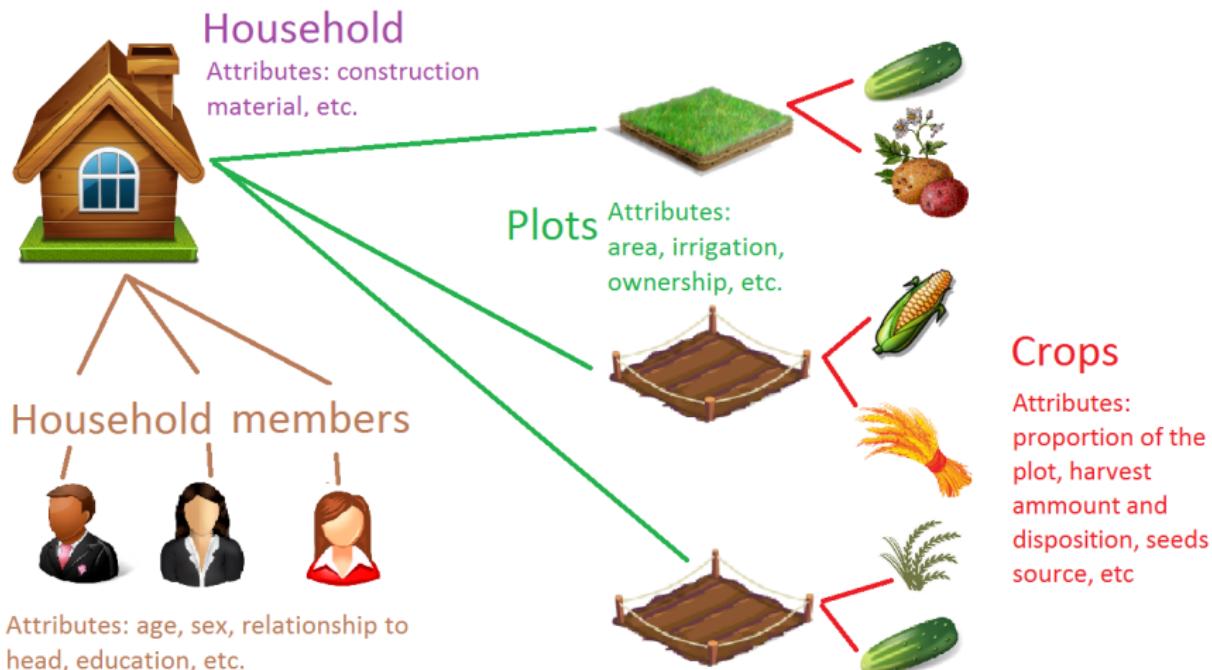
Selection of members

@rowcode	name	age	sex	@rowcode	name	age_married	children_planned
1	John	40	1	1	John		
2	Jessica	38	2	2	Jessica	22	3
3	Joseph	18	1	3	Joseph		
4	Joel	18	1	4	Joel		
5	Julia	18	2	5	Julia	18	1
6	Jeffrey	14	1	6	Jeffrey		
7	Jack	12	1	7	Jack		
8	Juan	10	1	8	Juan		
9	Justina	1	2	9	Justina		
10	James	0	1	10	James		
(sex==2) && (age>=14)							

Nested rosters

- Rosters can be nested (included one into another).
- Typical use situations:
 - for every plot that the household owns, ask about which crops are grown on them;
 - for every household member ask about every job the person ever had before;
 - for every production item ask about components, and suppliers of these components, and characteristics of these suppliers;
- One can include rosters of different types one into another, for example, first level roster: list-triggered household members information, second level roster triggered by numeric question *how many university degrees does this person have?* collecting information on each degree (year, country, university name, field, etc);

Nested rosters illustration



Linked questions

- Linked categorical questions are a special case of categorical questions.
- Regular categorical questions have a fixed set of choices specified at the time when the questionnaire is designed.
- Linked categorical questions draw their choices from the entries in the roster. This allows to form the list of choices specific to the specific respondent.
- Typical use: selecting one of the household members as an answer to a question, e.g. "*Who of the household members is the household head?*"

Linked questions illustration

3. Household roster												
3.1	3.2	3.3	3.4	3.5	3.6	3.7	3.8	3.9	3.10	3.11		
Line no.	Write the name of each household member	What is the relationship of <name> to the head of household?	How old is <name>? For children less than one year, write '00'	Is <name> male or female?	What is <name's> marital status? For codes, see at right 1=Male 2=Female	Line number of (first) spouse If not married, go to 3.8 If not in this household, write '98'	Does <name's> father live in this household? 1=Yes 2>No If No, go to 3.10	Line number of this father	Does <name's> mother live in this household? 1=Yes 2>No If No, go to next line	Line number of this mother		
01		<input type="checkbox"/> 0 <input type="checkbox"/> 1	<input type="checkbox"/> <input type="checkbox"/>		1 2	<input type="checkbox"/>	<input type="checkbox"/> <input type="checkbox"/>	1 2	<input type="checkbox"/> <input type="checkbox"/>	1 2	<input type="checkbox"/> <input type="checkbox"/>	
02		<input type="checkbox"/> <input type="checkbox"/>	<input type="checkbox"/> <input type="checkbox"/>		1 2	<input type="checkbox"/>	<input type="checkbox"/> <input type="checkbox"/>	1 2	<input type="checkbox"/> <input type="checkbox"/>	1 2	<input type="checkbox"/> <input type="checkbox"/>	
03		<input type="checkbox"/> <input type="checkbox"/>	<input type="checkbox"/> <input type="checkbox"/>		1 2	<input type="checkbox"/>	<input type="checkbox"/> <input type="checkbox"/>	1 2	<input type="checkbox"/> <input type="checkbox"/>	1 2	<input type="checkbox"/> <input type="checkbox"/>	
04		<input type="checkbox"/> <input type="checkbox"/>	<input type="checkbox"/> <input type="checkbox"/>		1 2	<input type="checkbox"/>	<input type="checkbox"/> <input type="checkbox"/>	1 2	<input type="checkbox"/> <input type="checkbox"/>	1 2	<input type="checkbox"/> <input type="checkbox"/>	
05		<input type="checkbox"/> <input type="checkbox"/>	<input type="checkbox"/> <input type="checkbox"/>		1 2	<input type="checkbox"/>	<input type="checkbox"/> <input type="checkbox"/>	1 2	<input type="checkbox"/> <input type="checkbox"/>	1 2	<input type="checkbox"/> <input type="checkbox"/>	

Setting up a linked question

- Your questionnaire must already contain a roster with a question that can be used for item names. For example, this may be a roster (driven by a numeric question) containing questions on the name, age and gender of each household member.
- Add new question after the roster. Select for question type: categorical single select.
- Fill out the common properties: variable name, label, and question text.
- Check the “*is linked*” checkbox.
- Select the *name* variable from the household members roster in the “*Bind to question from roster group*” setting.