

Table of Contents

Introduction	2
Basic Survey Authoring	2
1 - Your first question	2
2 - New Data Types	4
3 - Testing your work	5
4 - Notes and Dates	9
5 - Multiple Choice Questions	10
Intermediate Survey Authoring	14
1 - Multiple Language Support	14
2 - Hints and Metadata	14
3 - Conditional survey constructs.....	18
3.1 Required questions	18
3.2 Constraining Responses	19
3.3 Referencing a previous question	20
3.4 Asking Relevant Questions	21
3.5 The Specify Other Shortcut.....	23
Advanced Survey Authoring	24
1 - Grouping Questions.....	25
Skipping many questions at once	27
2 - Repeating a set of questions	27
3 - Special Formatting	29
4 - Making Dynamic Calculations.....	32
Appendix I - All Question Types.....	37
Appendix II - More Resources	37

Built by the [Modi Research Group](#) at Columbia University.

Code (BSD License) at github.com/modilabs/formhub

Copyright © 2012-2013, formhub.org

Introduction

This document will help you author surveys for ODK Collect using xlsform syntax. ODK Collect is a powerful tool to rapidly gather complex data types in the field using Android smartphones. The software is free, the source code is open, and the community is very active. Using ODK Collect with Formhub requires the creation of an Excel file that contains the questions, formatting instructions and data validation conditions that will allow enumerators collect data on smartphones. The following module will cover all the syntax elements you will need to author basic and advanced surveys.

Here are the steps involved in writing a survey with the XLSform format and getting it quickly deployed on formhub.

1. Write a survey using the XLSform syntax described below and save it as an .xls file.
2. Upload your .xls file into formhub.
3. Link ODK Collect to your formhub account, download forms, and start collecting data!

Basic Survey Authoring

1 - Your first question

Survey authoring in Excel is a simple process once you understand the basic XLSform syntax.

To make your learning process smoother a generic household survey example will be used across the training modules. This sample survey is a subset of the [CDC Household Water Use and Health Survey for a Water Safety Plan](#). You can download the full Excel survey file from [Formhub University](#), the name of the form is household_survey.

The screenshot shows a Microsoft Excel spreadsheet titled "household_survey.xls". The window title bar is at the top. Below it is a ribbon with tabs: Home, Layout, Tables, Charts, and a few others. The main area is a grid with columns labeled A, B, C, and D. Row 1 has three cells: A1 contains "type", B1 contains "name", and C1 contains "label". Row 2 has three cells: A2 contains "text", B2 contains "respondent_name", and C2 contains "What is your full name?". Rows 3 through 5 are empty. At the bottom of the grid, there is a toolbar with icons for navigating between sheets and saving. The current sheet is labeled "survey".

	A	B	C	D
1	type	name	label	
2	text	respondent_name	What is your full name?	
3				
4				
5				

1. We will save the survey with the file name household_survey.xls.
2. There is a single worksheet name 'survey' this is where all questions should be located (each row in this worksheet is a question in your survey).
3. There are three columns 'type', 'name', and 'label'. The 'type' column describes the question type (text, number, photo, etc.). The 'name' column assigns a unique variable name that will serve as a reference to the survey question (the name must be unique, must begin with a letter, and can only contain letters, numbers, dashes (-) and underscores (_)). The 'label' column contains the text that will actually be presented on the Android smartphone.
4. The survey pictured above has one basic question which asks "What is your full name?" (the 'label') and presents a text input box (as specified in the 'type' column) to the surveyor. Once data has been collected and sent to the server the data results Excel file that can be downloaded from Formhub will have a column called 'respondent_name' (the identifier for this question as specified in the 'name' column)
5. This file is saved in the 'Excel 97-2004 Workbook (.xls)' format.

This form would look like the following on your Android smartphone:

The screenshot shows a survey form titled "What is your full name?". The answer "John Doe" is entered into the field, which is highlighted with a red border. Below the input field is a virtual keyboard. The keyboard has three rows of letters: q w e r t y u i o p, a s d f g h j k l, and z x c v b n m. At the bottom of the keyboard are several function keys: XT9, a gear icon, 123, @!, a comma key, a backspace key, a period key, and a left arrow key.

Important rules to remember:

- Make sure your file is saved in the .xls format and contains no spaces or special characters ('-' and '_' are allowed).
- Make sure that your column headers are in lowercase (i.e. "label" or "name", not "Label" or "Name")
- Make sure that your sheet names are appropriately named (i.e. "survey" not "Sheet 1", "Survey" or "surveys")
- Make sure that the question names are unique and do not contain spaces or special characters ('-' and '_' are allowed).

Now let's start populating the survey questionnaire. The first section may include basic information such as the respondent's name, age of the respondent, household number, etc. We already know how to collect the name. For age, household number, and other fields, ODK lets us specify the "type" of the information to be entered (numbers, dates, etc). This will help us reduce data entry errors. To do this we can use a different question 'type', ODK Collect has defined a number question types to support all data types that surveys usually collect. Here we present some of the most basic types ODK Collect supports (you'll see the full list later).

Type	Question Input
text	Text input.
integer	Integer (ie, whole number) input.
decimal	Decimal input.

Let's add the household id and respondent's age as numeric integer fields and the respondent's name as a simple alphanumeric field. This will require the use of the 'type' 'text' and the 'type' 'integer'.

	A	B	C
1	type	name	label
2	text	interviewer	A.1.0 Interviewer name
3	integer	hh_id	A.1.1 Enter the household number
4	text	respondent_name	A.1.2 Enter the full name of the respondent
5	integer	respondent_age	A.1.3 Enter the age of the respondent
6			
7			
8			

An integer field would look like this on the Android smartphone.

The screenshot shows the ODK Collect interface on an Android device. At the top, it says "ODK Collect > household_survey". Below that is a question labeled "A.1.1 Enter the household number" with the value "6272904" entered into a text input field. A numeric keypad is overlaid on the screen, covering the bottom half of the question text. The keypad includes digits from 1 to 0, various symbols like @, £, €, \$, *, ., /, -, +, and special characters like up arrow, parentheses, percent, colon, exclamation, question, and a delete key. At the bottom, there are icons for KTS, settings, abc (switching to alphabetic keyboard), and other input controls.

Notice how the numeric keypad appears automatically and the alphabetic keyboard is deactivated.

2 - New Data Types

For those of you coming from paper surveys, you will appreciate that Android smartphones can help us collect “new” kinds of data, like GPS coordinates, images, audio, video, and information encoded in barcodes and QR codes. For now, we’ll introduce you to two additional questions: ‘types’ (‘geopoint’ and ‘image’).

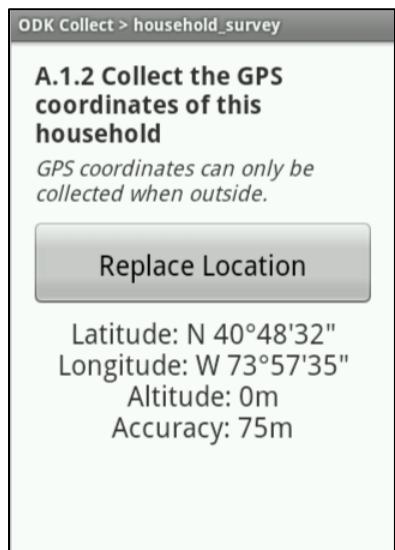
Type	Question Input
geopoint	Collect GPS coordinates.
image	Take a photograph.

Let’s collect the gps coordinates and a picture of the household using the following XLSform. To make this tutorial easier to read we will switch from the Excel file screenshots to a tabular notation.

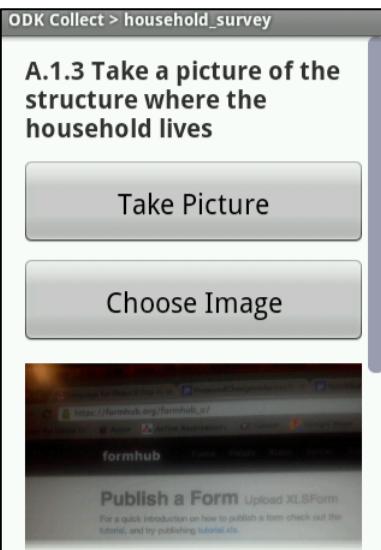
Excel 'survey' worksheet

type	name	label
text	interviewer	A.1.0 Interviewer name
integer	hh_id	A.1.1 Enter the household number
geopoint	hh_location	A.1.2 Collect the GPS coordinates of this household
image	hh_photo	A.1.3 Take a picture of the structure where the household lives
text	respondent_name	A.1.4 Enter the full name of the respondent
integer	respondent_age	A.1.5 Enter the age of the respondent

Notice how the Excel ‘a_worksheet_name’ worksheet preceding the table specifies on which worksheet of the workbook we are working on (we have been working on the Excel ‘survey’ worksheet for now but other worksheets will be added as we develop a more advanced survey).



A geopoint type



An image type

3 - Testing your work

Because the XLSform syntax requires a very specific set of instructions, it is quite easy to make errors in the Excel file. Let's say we misspelled the ‘image’ keyword and wrote ‘images’ instead.

	A	B	C
1	type	name	label
2	images	hh_photo	A.1.3 Take a picture of the structure where the household lives
3			
4			
5			
6			
7			
8			

When uploading to Formhub a form with errors, the system will show an error message that will give you an idea of what the error was. In this case Formhub is telling us it does not recognize ‘images’ as a valid question ‘type’.

formhub

Forms People About Syntax Support Blog

formhub_u

Unknown question type 'images'.

Publish a Form

Upload XLSForm

For a quick introduction on how to publish a form check out the [tutorial](#), and try publishing [tutorial.xls](#).

Learn about the XLSForm syntax [here](#), email [formhub users mailing list](#) if you have questions, and checkout examples at: [Formhub University](#).

Upload URL

XLS File: No file chosen

 **formhub_u** | Formhub University
View Profile Page

10 FORMS 9 SHARED FORMS 6 SUBMISSIONS

The process of uploading forms, and correcting errors is called debugging. Let's debug our form and change 'images' to 'image' in the 'type' column for question A.1.3. (Hint: The "Find" feature in the "Edit" menu of Excel is often helpful here). If we upload the form again we should see the following screen.

formhub

Forms People About Syntax Support Blog

formhub_u

Successfully published household_survey.

Publish a Form

Upload XLSForm

For a quick introduction on how to publish a form check out the [tutorial](#), and try publishing [tutorial.xls](#)

Learn about the XLSForm syntax [here](#), email [formhub users mailing list](#) if you have questions, and checkout examples at: [Formhub University](#).

Upload URL

XLS File: No file chosen

 **formhub_u** | Formhub University
View Profile Page

11 FORMS 9 SHARED FORMS 6 SUBMISSIONS

It is good practice to upload your XLSform (.xls file) to formhub on regular intervals. Formhub will show only an error message for the first problem it finds with a description of where the error is located. The fewer the changes since your last upload, the easier it is to find and solve the error formhub finds. You should get in the habit of testing your surveys, identifying, and fixing early on.

Sometimes, even if your form successfully uploads to formhub, it may not behave exactly as you intended on the Android smartphone. Therefore, although less frequently, it is also a good idea to test your form as a webform as well as directly on an Android smartphone.

At this stage, let's try to upload the current file into Formhub and test the form as a webform:

1. Sign in on the Formhub website (if you haven't created a user please refer to the quickstart section of the website).
2. Try to upload your 'household_survey.xls' XLSform and correct, if necessary, any errors that formhub reports. When the uploading is successful you will see the green message in the page "Successfully uploaded household_survey".
3. Now, when you scroll down, you will notice that your survey is listed in the "Published Forms". If no data has been submitted then on the right hand side is a small gear. Clicking it you access options to replace the form with a new version or delete the form.

Name	Submissions	Enter Data	View	Download	Last Submission
QR_test	3	<input checked="" type="checkbox"/> Web <input type="checkbox"/> Mobile			May 21, 2014
Shelter_ProgressUpdate	5	<input checked="" type="checkbox"/> Web <input type="checkbox"/> Mobile			May 22, 2014
Training_demoSurvey	0	<input checked="" type="checkbox"/> Web <input type="checkbox"/> Mobile			

Show inactive: Search:

Showing 1 to 3 of 3 entries

4. You can click the "Web" button under "Enter Data" to enter the webform page and be able to test certain aspects of your form.

A.1.0 Interviewer name (free-text)

A.1.1 Enter the household number (numeric)

A.1.2 Collect the GPS coordinates of this household
 ???.????? ????.?????

A.1.3 Take a picture of the structure where the household lives
 Sorry, web entry cannot support this type of question (binary)

A.1.4 Enter the full name of the respondent (free-text)

A.1.5 Enter the age of the respondent (numeric)

Be aware that the webform mode does not mimic the exact same format that will be displayed on the Android smartphone. You will still need to test the form on an Android smartphone but this intermediate step will let you test many aspects of the question flow and that data entry constraints are working correctly.

5. If you haven't downloaded ODK Collect on your Android smartphone and changed the settings to connect to the Formhub server and to your specific user account you can do it now. Scroll to the bottom of the web page to find instructions on how to do this.

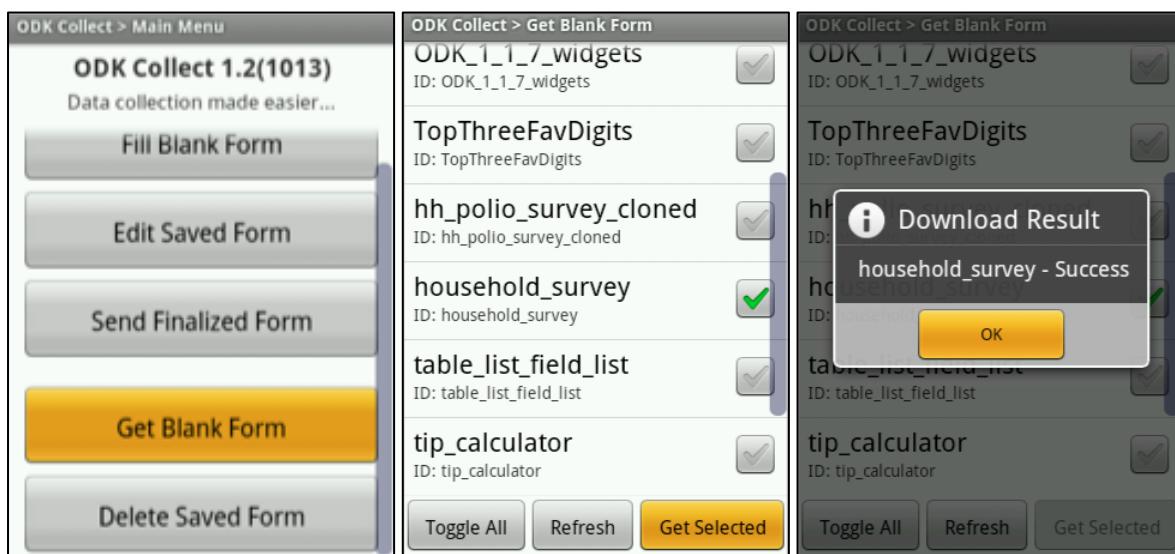
The url "of" this web application <https://formhub.org/danbjoseph> must be given to ODK Collect before it will get forms from and submit data to formhub. In Collect's Main Menu, press the Menu button. Select Server Preferences, then Server. Enter <https://formhub.org/danbjoseph> as the server.

Built by the [Modi Research Group](#) at Columbia University.

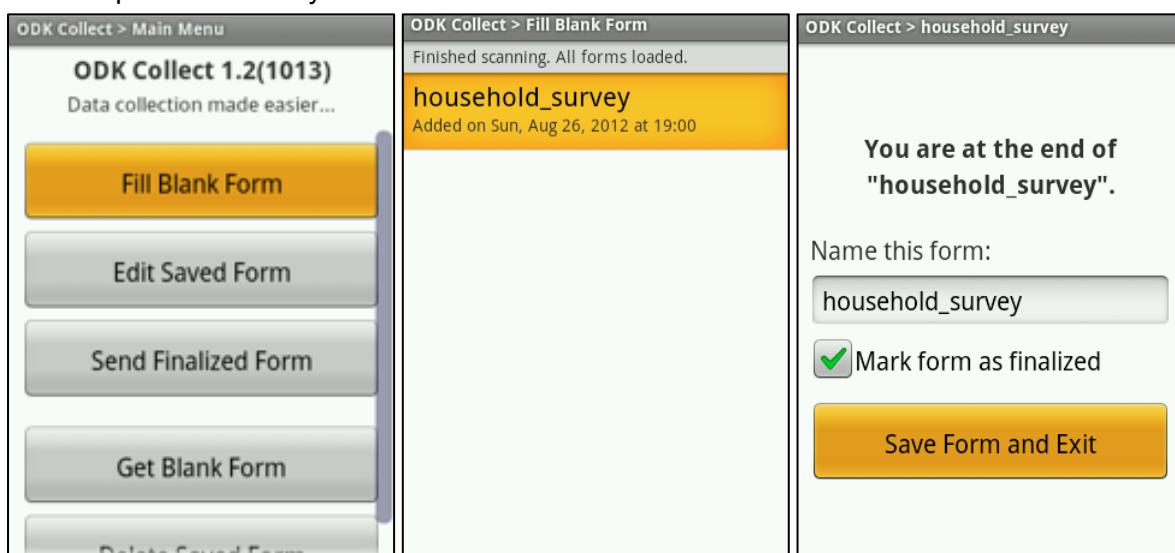
Code (BSD License) at github.com/modilabs/formhub

Copyright © 2012-2013, formhub.org

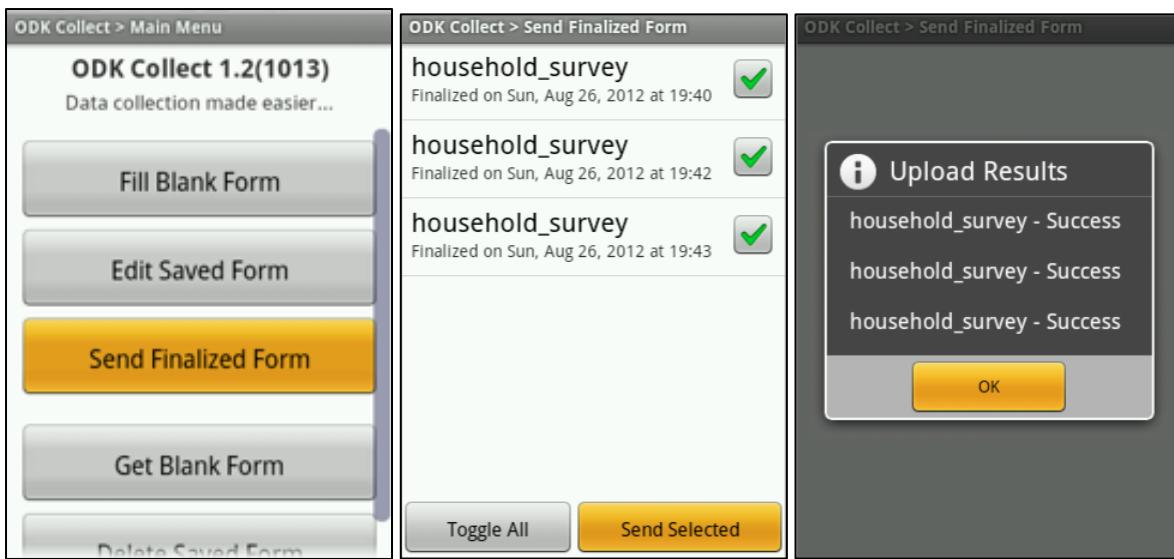
6. Download the form (Get Blank Form).



7. Fill in a couple of new blank forms (click the Fill Blank Form button and choose household_survey). Complete the survey and save the data.



8. Send your finalized submissions to the Formhub server.



9. Explore the different options that Formhub offers to use the data you just collected. Start by downloading your data as an xls file.

Name	Submissions	Enter Data	View	Download	Last Submission
QR_test	3	<input checked="" type="checkbox"/> Web <input type="checkbox"/> Mobile			May 21, 2014
Shelter_ProgressUpdate	5	<input checked="" type="checkbox"/> Web <input type="checkbox"/> Mobile			May 22, 2014
Training_demoSurvey	0	<input checked="" type="checkbox"/> Web <input type="checkbox"/> Mobile			

Show inactive: Search: _____

Showing 1 to 3 of 3 entries

Notice that our form has three submissions.

	A	B	C	D	E	F
1	interviewer	hh_id	hh_location	hh_photo	respondent_name	respondent_age
2	John Doe	6272904	40.80906813333334 -73.95988756666667 0.0 75.0	1346023535070.jpg	James wallace	32
3	John Doe	5432145	40.809177283333334 -73.95974123333333 0.0 75.0	1346024500936.jpg	Jim moore	45
4	John Doe	5167389	40.80916794285714 -73.95972777142858 0.0 75.0	1346024562785.jpg	Elvira jameson	15
5						
6						

The excel survey results are presented in tabular format, each column representing one question. Notice that the name of each column comes from our XLSform's 'name' column.

4 - Notes and Dates

Notes

Another question ‘type’ that can be very useful to provide information to your enumerator while conducting a survey are notes. Notes are no-input text prompts which can give the enumerator information to read to the respondent, remind her that she is now moving into a new section of the survey, or any other hints or messages to respond to.

Excel 'survey' worksheet

type	name	label
note	note_consent	Administer informed consent. If subject agrees to participate, proceed to questionnaire.

The screenshot shows the ODK Collect interface for a household survey. At the top, it says "ODK Collect > household_survey". Below that is a text area containing the note: "Administer informed consent. If subject agrees to participate, proceed to questionnaire."

Dates

Dates are fairly intuitive, they help collect a specific calendar date.

Excel 'survey' worksheet

type	name	label
date	respondent_dob	A.1.5 Enter the date of birth of the respondent

The screenshot shows the ODK Collect interface for a household survey. At the top, it says "ODK Collect > household_survey". Below that is a label "A.1.5 Enter the date of birth of the respondent". Below the label is a date input field consisting of three dropdown menus. The first menu shows the day "26", the second shows the month "Oct" (which is highlighted with an orange border), and the third shows the year "1998". Each menu has a plus sign at the top and a minus sign at the bottom for navigating through the options.

5 - Multiple Choice Questions

As we continue with the household survey, we are interested in the gender of the respondent, the number of people living in the household and the different sources of water used in the household. For these

questions, we could simply ask text and integer questions, but often it is useful to standardize data by asking enumerators multiple-choice questions. By using a list of predefined choices in survey questions, we can increase the value of data for quantitative analysis later. This technique removes the need for “coding” data in many cases, and reduces errors introduced by typos and the need to interpret data.

Multiple Choice Questions

ODK Collect has support for both ‘select one option’ (‘select_one’) and ‘select all options that apply’ (‘select_multiple’) questions. Writing a multiple choice question requires some additional steps to the other types we have used up to now. We need to specify what the predefined list of choices will be!

Type	Question Input
select_one [choices]	Multiple choice question; only one answer can be selected (the [] notation indicates that you will need to replace [choices] by a name in the choices sheet).
select_multiple [choices]	Multiple choice question; multiple answers can be selected (the [] notation indicates that you will need to replace [choices] by a name in the choices sheet).

1. Add a second worksheet called ‘choices’ (the name of this new worksheet has to be exactly ‘choices’ in lowercase) to our Excel workbook.
2. In this new worksheet create three columns ‘list name’ (the name of the list of alternative choices for this multiple option question and will need to be referenced from the Excel ‘survey’ worksheet), ‘name’ (a unique identifier of each choice alternative that contains no spaces or special characters) and ‘label’ (the text that will be displayed on the phone).
3. In the Excel ‘survey’ worksheet create a new row with the ‘type’ set to ‘select_one’ or ‘select_multiple’ followed by a space and replace [choices] by the name of the list of alternative choices as defined in the Excel ‘choices’ worksheet under the column ‘list name’.
4. Populate the Excel ‘choices’ worksheet with all the options that are anticipated.

Here is an example of a ‘select one’ question to restrict the gender to ‘Male’ and ‘Female’:

A	B	C
1 type	name	label
2 select_one gender_list	respondent_gender	A.1.7 Enter the gender of the respondent
3		
4		
5		

A	B	C	D	E
1 list name	name	label		
2 gender_list	male	Male		
3 gender_list	female	Female		
4				
5				

Note that gender_list appears in both worksheets, this is how the elements are linked between the choice list and the actual survey question. A choice list can be re-used in multiple survey questions.

ODK Collect > household_survey

A.1.7 Enter the gender of the respondent

Male

Female

A more elaborate multiple question that only accepts one selection would be the question on ownership of the house.

Excel 'survey' worksheet

type	name	label
select_one ownership_list	hh_ownership	A.1.9 This house is ...

Excel 'choices' worksheet

list name	name	label
ownership_list	own	Owned
ownership_list	rent	Rented/Leased
ownership_list	rent_free	Rent-Free
ownership_list	informal	Informal Settlement
ownership_list	other	Other

ODK Collect > household_survey

A.1.11 This house is ...

Owned

Rented/Leased

Rent-Free

Informal Settlement

Other

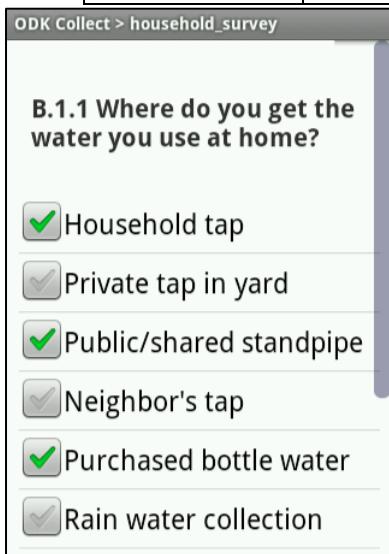
We can also add multiple choices questions that allow multiple answers to be selected. For example in our survey it could be that households obtain water from different sources for different purposes or in different seasons (dry vs wet season).

Excel 'survey' worksheet

type	name	label
select_multiple water_sources	hh_water_source	B.1.1 Where do you get the water you use at home? (select all that apply)

Excel 'choices' worksheet

list name	name	label
water_sources	house_tap	Household tap
water_sources	private_tap	Private tap in yard
water_sources	public_tap	Public/shared standpipe
water_sources	neighbor_tap	Neighbor's tap
water_sources	bottled	Purchased bottle water
water_sources	rain_collect	Rain water collection
water_sources	river	River/Stream/Creek
water_sources	spring	Spring
water_sources	ref_station	Refilling station
water_sources	other	Other



Things to remember for multiple choice questions:

- The second worksheet should be called 'choices' and not 'Choices' 'choicelist' or any other alternative.
- Make sure you don't forget the first row and the headings 'list name', 'name' and 'label'.
- The name of your choices list in the Excel 'choices' worksheet (under 'list name' in the choices survey) has to be used again right after the 'select_one' or 'select_multiple' question type in under the column 'type' in the Excel 'survey' worksheet.
- The name of each choice, the content of the 'name' column of the Excel 'choices' worksheet cannot contain spaces or special characters and will appear in the survey data results file.

This is how the survey result data for multiple questions will look like.

	A	B	C	D
1	hh_ownership	respondent_questions/respondent_gender	hh_water_source/house_tap	hh_water_source/private_tap
2	informal	female	FALSE	FALSE
3	rent	male	FALSE	FALSE
4	own	male	TRUE	TRUE
5				
6				

Notice that for the select one option type of questions the responses are in one column named as defined under the 'name' column in the Excel 'survey' worksheet. The content of the cells is the name of each alternative as defined in the Excel 'choices' worksheet under the 'name' column.

For select-all-that-apply type of questions, one column is created for each option defined in the Excel 'choices' worksheet and the name is the name of the question as defined under the 'name' column in the Excel 'survey' worksheet followed by '/' and the name of each choice as defined in the 'name' column of the Excel 'choices' worksheet. The content of the cells is TRUE or FALSE depending on if that particular option was selected or not.

Intermediate Survey Authoring

We have reviewed most data types and you should be comfortable with the creation of basic surveys. We will now focus on strengthening the logic flow of the survey questions and we will introduce conditional elements to increase data quality and validation.

1 - Multiple Language Support

It is very easy in xlsform to add multiple languages to a survey. To do this, you have to essentially just create a difference label column for each of your languages. For example a form with the columns label::English and label::French will allow the surveyors to choose between English and French. On ODK, to select a different language, press the Menu key, and the "Change Language" option. For the form below, "English" and "French" will show up as the possible options.

Excel 'survey' worksheet

type	name	label::English	label::French
select_one gender_list	respondent_gender	A.1.7 Enter the gender of the respondent	A.1.7 Indiquez le sexe du sondé

Make sure that your column headers in the "choices" sheet match those in the "survey" sheet (i.e. there will be errors if you columns are "label::English" on the survey and "label" on the choices worksheet)

Excel 'choices' worksheet

list name	name	label::English	label::French
gender_list	male	Male	Masculin
gender_list	female	Female	Feminin

[Advanced: See the [xlsform standard document](#) to see exactly what kinds of column headers can accept a language modification.]

2 - Hints and Metadata

Hints

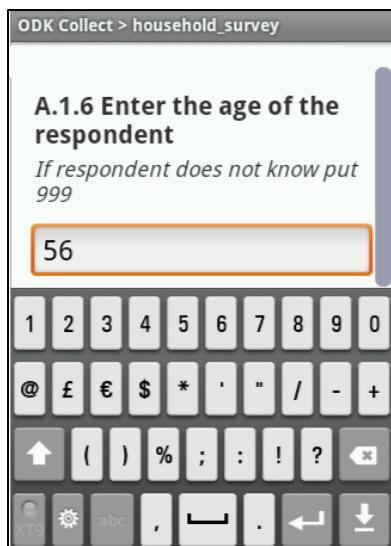
When asking questions, it often makes sense to give enumerators some extra guidance. You can already do this with notes, but what about something simpler that doesn't require an extra screen? The answer is

"hints," which are listed along-side the question in italicized font (questions are in bold font). To use them, just create new column in your survey worksheet named 'hint', and add your hints.

Excel 'survey' worksheet

type	name	label	hint
note	note_consent	Administer informed consent. If subject agrees to participate, proceed to questionnaire.	
text	interviewer	A.1.0 Interviewer name	
integer	hh_id	A.1.1 Enter the household number	
geopoint	hh_location	A.1.2 Collect the GPS coordinates of this household	Make sure the precision is less than 25 meters
image	hh_photo	A.1.3 Take a picture of the structure where the household lives	Take a picture of the front of the house
text	respondent_name	A.1.4 Enter the full name of the respondent	First, Middle and Last name
date	respondent_dob	A.1.5 Enter the date of birth of the respondent	
integer	respondent_age	A.1.6 Enter the age of the respondent	If respondent does not know put 999
select_one gender_list	respondent_gender	A.1.7 Enter the gender of the respondent	(select one)
integer	hh_people_nb	A.1.8 How many people live in this household?	Including the respondent
select_one owner_list	hh_ownership	A.1.9 This house is ...	(select one)
note	note_section_b	Section B - Water use practices. This section is composed of 23 questions related to the household usage of water.	
select_multiple water_sources	hh_water_source	B.1.1 Where do you get the water you use at home?	(select all that apply)

Question A.1.6 would now look like this on the Android smartphone.



ODK Collect > household_survey

A.1.6 Enter the age of the respondent

If respondent does not know put 999

56

1 2 3 4 5 6 7 8 9 0
@ £ € \$ * . / - +
↑ () % ; : ! ? ⌘
KTP abc , — . ⌘ ⌘

You can also have your hints in different languages. For example, the above form could use two additional columns “hint::English” and “hint::French” to provide hints in both of those languages.

Metadata

In surveying efforts, it is often useful to collect additional data for administrative purposes. For example, you might want to know the time it is taking enumerators to conduct the full survey. During the pilot stage, identifying which phone was used to collect certain GPS coordinates might help the detection of faulty GPS systems (important when using a large number of devices). Other similar information can be used for other administrative and review processes; the “start” and “end” date-time fields are particularly helpful when checking for duplicate information entry.

ODK Collect makes a number of metadata fields available for collection:

start	Start date and time of the survey.
end	End date and time of the survey.
today	Day of the survey.
deviceid	IMEI (International Mobile Equipment Identity)
subscriberid	IMSI (International Mobile Subscriber Identity)
imei	SIM serial number (if available).
phonenumbers	Phone number (if available).

If we wanted our survey to collect all of this metadata we would add the following rows at the top of the survey (they can be anywhere but it is a best practice to put metadata fields at the beginning of the file):

Excel ‘survey’ worksheet

type	name	label
start	start	Start Time
end	end	End Time
today	today	Date of Survey
deviceid	deviceid	Phone Serial Number
subscriberid	subscriberid	Subscriber Identifier
imei	simserial	SIM serial
phonenumbers	phonenumbers	Phone number

These questions do not appear on the phone at all, but will be automatically included in the data ODK Collect writes to store and transmit the survey data. To see these fields at work, fill out a new blank form on your Android smartphone, save it and send it to the server. When you download the data you will see that new columns have been added to the survey data.

	A	B	C
1	start	end	today
2	2012-08-28T00:37:56.978Z	2012-08-28T00:39:11.696Z	2012-08-28
3	2012-08-28T00:39:15.843Z	2012-08-28T00:40:30.833Z	2012-08-28
4	2012-08-28T00:34:23.590Z	2012-08-28T00:37:50.508Z	2012-08-28
5			
6			
7			

3 - Conditional survey constructs

In order to reduce input errors ODK Collect has some options to control and restrict the way data is entered by enumerators. Three types of constructs are available:

1. Requiring that certain questions are filled with data before moving forward.
2. Constraining the values of data entered in certain questions.
3. Skipping non-relevant questions depending on the answers to previous questions (avoid asking what type of rainwater catchment system is used if the household answered they do not collect rainwater).

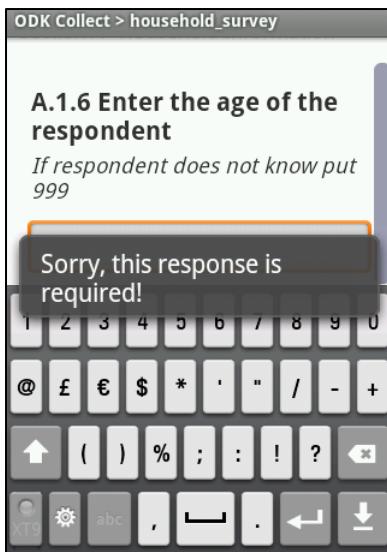
3.1 Required questions

ODK allows the survey author to designate certain questions as “required” questions. Normally, enumerators are able to “skip” answering any question by simply swiping forward on a page. However, many surveyors prefer that enumerators enter a specific value (such as “-999” for “Not available”) rather than simply skip question. ODK allows you to enforce this; simply column labeled ‘required’ and filling the cells with the reserved word ‘yes’ to the questions that you consider critical to have in your data.

Excel 'survey' worksheet

type	name	label	required
note	note_consent	Administer informed consent. If subject agrees to participate, proceed to questionnaire.	
text	interviewer	A.1.0 Interviewer name	
integer	hh_id	A.1.1 Enter the household number	yes
geopoint	hh_location	A.1.2 Collect the GPS coordinates of this household	
image	hh_photo	A.1.3 Take a picture of the structure where the household lives	
text	respondent_name	A.1.4 Enter the full name of the respondent	yes
date	respondent_dob	A.1.5 Enter the date of birth of the respondent	yes
integer	respondent_age	A.1.6 Enter the age of the respondent	yes
select_one gender_list	respondent_gender	A.1.7 Enter the gender of the respondent	
integer	hh_people_nb	A.1.8 How many people live in this household?	
select_one owner_list	hh_ownership	A.1.9 This house is ...	
note	note_section_b	Section B - Water use practices. This section is composed of 23 questions related to the household usage of water.	
select_multiple water_sources	hh_water_source	B.1.1 Where do you get the water you use at home?	yes

If you test the form, and try to swipe ahead on (for example) question A.1.6, you will not be able to move forward, and instead, get a prompt saying “Sorry, this response is required!”



Warning: Be careful when adding conditional constructs to a survey or the enumerator might find herself incapable of finishing the survey (if the GPS on the phone is not working and you have required the collection of geographic coordinates, enumerators might not be able to move forward to the next sections of the survey). It is also important to test every survey with conditional constructs as much as possible before deployment. For text and integer fields, be sure to specify what should be entered in case the respondent is not able to answer the enumerator.

3.2 Constraining Responses

Certain questions only have a limited range of possible answers that can be considered as valid. If, for example, you ask a household how many days of the week they have to fetch water, the response can only be a number between 0 and 7. ODK Collect handles this type of data validation with the use of the 'constraint' column.

Excel 'survey' worksheet

type	name	label	constraint
integer	respondent_age	A.1.6 Enter the age of the respondent	. >= 0

Here we've put a constraint on age so that negative values are not accepted. ODK Collect will not allow the enumerator to swipe to the next question if she enters -5. These constraints can help protect from typos and other unintentional errors.

The constraint formulas can be fairly complex, accommodating the use of and, or and not operators. Lets say that our age limit is 150 (much lower than the oldest person every known to exist), and that we want to allow for the enumerator to enter 999 in case the respondent cannot answer the question. The constraint can be written as follows.

Excel 'survey' worksheet

type	name	label	constraint	constraint_message
integer	respondent_age	A.1.6 Enter the age of the respondent	(. >= 0 and .<=150) or .=999	age has to b between 0 and 150, enter 999 if unknown

In fact, we have even added a new column labeled 'constraint_message' that specified the message displayed on the screen (the usual popup message is "Sorry, this response is invalid"):

ODK Collect > household_survey

A.1.6 Enter the age of the respondent

If respondent does not know put 999

age has to be between 0 and 150, enter 999 if unknown

3.3 Referencing a previous question

Let's move on, and add a question on the number of household occupants that are more than 18 years old. Here, the number of adults in the household has to be smaller than or equal to the total number of people living in the household.

XLSform can let you reference the results of a previous question by using a very specific syntax. \${name_of_previous_question} will actually fetch the value entered by the enumerator in the question that has a name (as defined in the 'name' column of the Excel 'survey' worksheet). Notice that the special character '.' will be interpreted by ODK Collect as a reference to the current question (as if we had typed \${name_of_current_question}).

Excel 'survey' worksheet

type	name	label	constraint	constraint_message
integer	hh_people_nb	A.1.8 How many people live in this household?		
integer	hh_adult_nb	A.1.9 How many adults live in this household?	. <= \${hh_people_nb}	has to be smaller than or equal to the number of people in household

So, here, the answer to question A.1.9 is constrained to be less than or equal to the answer given to question A.1.8.

A.1.8 How many people live in this household?
Including the respondent

4

1 2 3 4 5 6 7 8 9 0
@ £ € \$ * ' " / - +
↑ () % ; : ! ? ← ↓
KTS abc , ← .

A.1.9 How many adults live in this household?

5

has to be smaller than or equal to the number of people in the household

[Advanced: The column ‘constraint’ takes an XPath expression. The \${QUESTION_NAME} notation generates a valid XPath formula. A question is accepted if the ‘constraint’ column evaluates to true.]

3.4 Asking Relevant Questions

Many surveys have questions that should be asked depending on the answer given to previous questions. One great feature of ODK Collect is the ability to only present a question to the enumerator if a condition, based on the response to a previous question, is true.

For example, if you presented a multiple choice question and the answer by the interviewee does not correspond to any of the pre-programmed choices, you might want to offer ‘Other’ as an alternative. If this is the case you would need to ask the interviewee to specify what other means. To do this we add the ‘relevant’ column and enter a condition that specifies when the question is relevant and should be presented to the enumerator.

You can interpret this conditions as: “Only ask this question if the answer to the previous question: name_of_question (as defined in the ‘name’ column of the Excel ‘survey’ worksheet) was equal to ‘option_selected’ (as defined in the ‘name’ column of the Excel ‘choices’ worksheet)”.

Notice that in the case of ‘select_multiple’ questions the syntax is a bit different and you can actually specify relevant conditions for each one of the options selected. Below is an example of how to do this using XLSform.

Excel ‘survey’ worksheet

type	name	label	relevant
select_one owner_list	hh_ownership	A.1.10 This house is ...	
text	hh_ownership_other	A.1.10.0 Please specify other	\${hh_ownership} = 'other'
select_multiple water_sources	hh_water_source	B.1.1 Where do you get the water you use at home?	
text	hh_water_source_other	B.1.1.0 Please specify other	selected(\${hh_water_source}, 'other')
decimal	hh_bottle_expenditure	B.1.2 How much do you	selected(\${hh_water_source}, 'bottled')

		spend on bottled water per month	
--	--	--	--

Excel 'choices' worksheet

list name	name	label
ownership_list	own	Owned
ownership_list	rent	Rented/Leased
ownership_list	rent_free	Rent-Free
ownership_list	informal	Informal Settlement
ownership_list	other	Other
water_sources	house_tap	Household tap
water_sources	private_tap	Private tap in yard
water_sources	public_tap	Public/shared standpipe
water_sources	neighbor_tap	Neighbor's tap
water_sources	bottled	Purchased bottle water
water_sources	rain_collect	Rain water collection
water_sources	river	River/Stream/Creek
water_sources	spring	Spring
water_sources	ref_station	Refilling station
water_sources	other	Other

If in question A.1.10 the interviewee answers 'Other' then the phone will not move to question B.1.1 but will present question A.1.10.0 Please specify other. The content in the 'relevant' column: \${hh_ownership} = 'other' actually means: "only ask this question if the answer to the question with 'name' equal to hh_ownership was the choice alternative with a 'name' equal to other."

ODK Collect > household_survey

A.1.11 This house is ...
(select one)

Owned

Rented/Leased

Rent-Free

Informal Settlement

Other

ODK Collect > household_survey

A.1.11.0 Please specify other

Special arrangement

If in question "B.1.1 Where do you get the water you use at home?" the interviewee selects the option "Bottled Water" regardless of what else was selected, then question B.1.2 should be presented on the phone, if not then it should be skipped. The content in the 'relevant' column: selected(\${hh_water_source}, 'bottled') actually means: "only ask this question if the choice alternative with

a ‘name’ equal to ‘bottled’ was selected amongst any other options in the question with ‘name’ equal to hh_water_source was the choice alternative with a ‘name’ equal to other.

Ask if relevant:

ODK Collect > household_survey

B.1.1 Where do you get the water you use at home?
(select all that apply)

- Household tap
- Private tap in yard
- Public/shared standpipe
- Neighbor's tap
- Purchased bottle water
- Rain water collection

ODK Collect > household_survey

B.1.2.1 How much do you spend on bottled water per month

34.22

1 2 3 4 5 6 7 8 9 0
@ £ € \$ * . - / +
↑ () % ; : ! ? ← →
XT9 abc , — . ← →

Skip if not relevant:

ODK Collect > household_survey

B.1.1 Where do you get the water you use at home?
(select all that apply)

- Household tap
- Private tap in yard
- Public/shared standpipe
- Neighbor's tap
- Purchased bottle water
- Rain water collection

ODK Collect > household_survey

You are at the end of "household_survey".

Name this form:

Mark form as finalized

You will learn how to skip more than one question at a time later on.

Important to remember: For skip logic based on ‘select_one’ questions, you can use the = (equals) construct (see the living arrangement question above). For skip logic based on ‘select_multiple’ questions, you can use the selected() function, as seen in the water sources question.

For more on proper XPath expressions, see this page:

<https://bitbucket.org/javarosa/javarosa/wiki/buildxforms>

3.5 The Specify Other Shortcut

For multiple choice options, frequently, especially in pilot phases, survey authors include an “other” option that has to be specified if selected. XLSForm has a shortcut to make this easier (for English Language surveys; unfortunately, this feature is not available for multiple languages):

Excel ‘survey’ worksheet

type	name	label	relevant
------	------	-------	----------

select_one owner_list or_other	hh_ownership	A.1.10 This house is ...	
--------------------------------	--------------	--------------------------	--

Excel 'choices' worksheet

list name	name	label
ownership_list	own	Owned
ownership_list	rent	Rented/Leased
ownership_list	rent_free	Rent-Free
ownership_list	informal	Informal Settlement

instead of...

Excel 'survey' worksheet

type	name	label	relevant
select_one owner_list	hh_ownership	A.1.10 This house is ...	
text	hh_ownership_other	A.1.10.0 Please specify other	\${hh_ownership} = 'other'

Excel 'choices' worksheet

list name	name	label
ownership_list	own	Owned
ownership_list	rent	Rented/Leased
ownership_list	rent_free	Rent-Free
ownership_list	informal	Informal Settlement
ownership_list	other	Other

Important to remember: If you have data like this, when you download the survey results data, in the hh_ownership column, you will see a value 'other'. A second column will have the answer for the questions which the user selected 'other.' This can make doing data analysis more cumbersome, so we don't recommend the 'or_other' construct for large-scale data collection efforts. The option is perfectly appropriate for pilots, however.

Advanced Survey Authoring

You should be comfortable now with the use of required, constraint and relevant. By now you should be able to author an intermediate survey that validates the data entered by the enumerator improving data quality.

The use of relevant conditions is key to a successful survey so make sure you understand it's use before moving forward. Advanced survey authoring presents some miscellaneous syntax elements that will improve your survey.

We will cover the use of groups for skipping several questions and present groups of questions on the same page. We will also improve the formatting of our survey with the use of the 'appearance' column and finally we will present the cascading select functionality that will allow to present cascading multiple choice selections with dependencies between them.

1 - Grouping Questions

Grouping questions has several advantages, the most immediate one is that the group name will be presented to the enumerator as small text before each question. This can be helpful to remember in which section of the survey we are. Other advantages of groups will be explored in the next sections. To create a group of questions try the following:

Excel 'survey' worksheet

type	name	label
begin group	sectionB	Section B - Water use practices
note	note_section_b	Section B - Water use practices. This section is composed of 23 questions related to the household usage of water.
select_multiple_water_sources	hh_water_source	B.1.1 Where do you get the water you use at home?
text	hh_water_source_other	B.1.1.0 Please specify other
decimal	hh_bottle_expenditure	B.1.2 How much do you spend on bottled water per month
end group		

Using groups adds the 'label' content of the group construct to all questions pertaining to that group.

Using groups also changes how the survey data file columns are named. The group 'name' now precedes the question 'name' with a '/' in between. Like "sectionB/hh_water_source".

	A	B	C	D	E
1	sectionA/respondent_questions/respondent_name	sectionA/respondent_questions/respondent_age	sectionA/respondent_questions/respondent_gender	sectionB/hh_water_source/house_tap	sectionB/hh_water_source/private_tap
2	Simon Harper	17	female	TRUE	FALSE
3	Emily Watson	32	male	FALSE	FALSE
4					
5					
6					

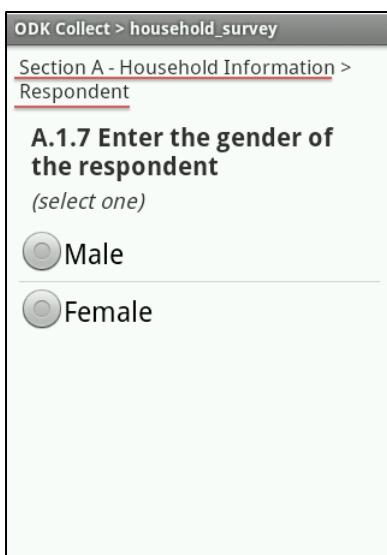
If you have subsections within your sections, groups can be nested very easily. Here is an example, where we have grouped all questions asked to the respondent within sectionA:

Excel 'survey' worksheet

type	name	label
begin group	sectionA	Section A - Household Information

note	note_consent	Administer informed consent. If subject agrees to participate, proceed to questionnaire.
text	interviewer	A.1.0 Interviewer name
integer	hh_id	A.1.1 Enter the household number
geopoint	hh_location	A.1.2 Collect the GPS coordinates of this household
begin group	respondent_questions	Respondent
image	hh_photo	A.1.3 Take a picture of the structure where the household lives
text	respondent_name	A.1.4 Enter the full name of the respondent
date	respondent_dob	A.1.5 Enter the date of birth of the respondent
integer	respondent_age	A.1.6 Enter the age of the respondent
select_one gender_list	respondent_gender	A.1.7 Enter the gender of the respondent
end group		
integer	hh_people_nb	A.1.8 How many people live in this household?
integer	hh_adult_nb	A.1.9 How many adults live in this household?
select_one owner_list	hh_ownership	A.1.10 This house is ...
text	hh_ownership_other	A.1.10.0 Please specify other
end group		

Notice how the nested group is shown on the phone with the parent group 'label' and the '>' character in the group text area now. This can be useful for subsections, just remember not to put too many nested groups or your phone screen will become illegible.



In the survey data result file notice how nested groups add the 'label' of the parent group to the column name for that variable. The answers for question "respondent_age" are now under "sectionA/respondent_questions/respondent_age".

	A	B	C	D	E
1	sectionA/respondent_questions/respondent_name	sectionA/respondent_questions/respondent_age	sectionA/respondent_questions/respondent_gender	sectionB/hh_water_source/house_tap	sectionB/hh_water_source/private_tap
2	Simon Harper	17	female	TRUE	FALSE
3	Emily Watson	32	male	FALSE	FALSE
4					
5					
6					

Skipping many questions at once

One of the most useful uses of groups is to skip a whole bunch of questions at once. In our survey, we only want to ask questions in section B if the respondent is over 18 years of age.

We will need a relevant condition telling ODK the following: “Only present this question if the answer to the question ‘respondent_age’ was larger than or equal to 18” ($\{\text{respondent_age}\} \geq 18$). The following example would skip questions B.1.1, B.1.1.0 and B.1.2 if the respondent is less than 18 years of age.

Excel 'survey' worksheet

type	name	label	relevant
integer	respondent_age	A.1.6 Enter the age of the respondent	
begin group	sectionB	Section B - Water use practices	$\{\text{respondent_age}\} \geq 18$
note	note_section_b	Section B - Water use practices. This section is composed of 23 questions related to the household usage of water.	
select_multiple water_sources	hh_water_source	B.1.1 Where do you get the water you use at home?	
text	hh_water_source_other	B.1.1.0 Please specify other	selected($\{\text{hh_water_source}\}$, 'other')
decimal	hh_bottle_expenditure	B.1.2 How much do you spend on bottled water per month	selected($\{\text{hh_water_source}\}$, 'bottled')
end group			

Without the use of groups you would need to add the relevant condition in each row. This would be cumbersome and difficult to maintain for the survey author. Using groups is a much more efficient way of performing this task.

2 - Repeating a set of questions

You might also want to create a repeating set of questions. In our case we might be interested in the age and gender of every member of the household. We cannot do this with existing tools, because we don't know the number of household members in advance. To do this, we use the following construct:

Excel 'survey' worksheet

type	name	label
integer	hh_people_nb	A.1.8 How many people live in this household?
integer	hh_adult_nb	A.1.9 How many adults live in this household?
begin repeat	household_member	A.1.10 Household Member
integer	hh_member_age	A.1.10.1 Enter the age of this member
select_one gender_list	hh_member_gender	A.1.10.2 Enter the gender of this member
end repeat		
select_one owner_list	hh_ownership	A.1.11 This house is ...

text	hh_ownership_other	A.1.11.0 Please specify other
------	--------------------	-------------------------------

This is a list of household members, with the age and gender of each household member. The phone will ask the name and age of the first household member, and then ask if the enumerator wants to add a new 'Household Member' group. If the enumerator responds with a 'yes' then the name and age of the second household member will be asked, and so on.

The first screenshot shows the initial screen for adding a new household member. It asks if the user wants to "Add New Group?" and provides "Add a new 'A.1.10 Household Member' group?" options. The second screenshot shows the gender selection screen for the first member, with "Male" selected. The third screenshot shows the option to "Add One More Group?" after the first member has been added.

Notice that in the group text area there is a counter (a number between parentheses) that tells us how many household members we have added.

Note that the exported data for repeats comes in a variety of formats depending on the export you perform, and data within repeats can be difficult to perform analysis with. We recommend using repeats with care, and constructing a methodology around the use of that repeat data before doing large scale data collection with repeats.

We will walk you through a short example of repeat exports, but please use repeats only after you understand them fully. Let's say that one of our households had three inhabitants and the other only one. The data would look like this:

	B	C	D	E
1	sectionA/respondent_questions/respondent_age	sectionA/respondent_questions/respondent_gender	index	
2	17	female	1	
3	32	male	2	
4				
5				
6				
7				
8				
9				
10				

household_survey household_member +

A	B	C	D	E	F	G	H	I
1	sectionA/household_member/hh_member_age	sectionA/household_member/hh_member_gender	index	parent_table_name	parent_index			
2	17	female	1	household_survey	1			
3	21	male	2	household_survey	2			
4	32	female	3	household_survey	2			
5	43	female	4	household_survey	2			
6								
7								
8								
9								

household_survey household_member +

Normal View Ready Sum

The data results for the repeating group are stored in a different sheet in the survey results file (see the second picture; ‘household_member’ sheet). The name of the sheet was derived by the name of this “begin repeat” question in your XLSform.

Explaining in detail what is happening is actually a relational databases 101 course which is not what we want you to learn. Just know that there is a new results sheet and that each row corresponds to one repeating group element. How do we know to which household is each household member linked to? Look at the _parent_table_name and the _parent_index column, these columns were automatically generated by Formhub. They basically are telling us that this household member belongs to a household that has index number 1 or 2 in the household_survey worksheet. In this case, the _index column in the household_survey worksheet corresponds to _parent_index in the household_member file.

A data analyst with database abilities will find no problem in using this information. If you have an experienced database manager then go ahead and use begin repeat (with some practice beforehand). If not, we recommend you avoid using this construct and just create multiple questions.

Note: Using nested groups with a field list appearance will crash the ODK Collect software.

3 - Special Formatting

ODK Collect allows for certain changes in the formatting of smartphone screens, for example you might want to show multiple choice select_one question as a list box instead of a long list of checkboxes. You might also want to present two questions on the same screen on the phone. Let’s say question B.1.2 is unit-less and you want the respondent to specify the currency of expenditure on bottled water. Up to now we would have done this as follows.

Excel 'survey' worksheet

type	name	label
decimal	hh_bottle_expenditure	B.1.2 How much do you spend on bottled water per month
select_one currencies	hh_bottle_currency	Currency (specify one)

Excel 'choices' worksheet

list name	name	label
currencies	dollars	USD
currencies	local	Local currency

which would have shown:

<p>ODK Collect > household_survey</p> <p>Section B - Water use practices > B.1.2 Bottled water expenditure</p> <p>B.1.2.1 How much do you spend on bottled water per month</p> <p>34.22</p>	<p>ODK Collect > household_survey</p> <p>Section B - Water use practices > B.1.2 Bottled water expenditure</p> <p>Currency (specify one)</p> <p><input checked="" type="radio"/> USD</p> <hr/> <p><input type="radio"/> Local currency</p>
--	---

Here is how you would get them to show up on the same screen:

Excel 'survey' worksheet

type	name	label	appearance
begin group	hh_bottle_group	B.1.2 Bottled water expenditure	field-list
decimal	hh_bottle_expenditure	B.1.2.1 How much do you spend on bottled water per month	
select_one currencies	hh_bottle_currency	Currency (specify one)	
end group			

Now, the questions are shown one on top of the other separated by a gray line.

ODK Collect > household_survey

Section B - Water use practices > B.1.2
Bottled water expenditure

B.1.2.1 How much do you spend on bottled water per month

34.22

Currency (specify one)

USD

Local currency

Now let's use the table-list appearance widget to present the currency choices as a horizontal, more compact, list.

Excel 'survey' worksheet

type	name	label	appearance
begin group	hh_bottle_group	B.1.2 Bottled water expenditure	table-list
decimal	hh_bottle_expenditure	B.1.2.1 How much do you spend on bottled water per month	
select_one currencies	hh_bottle_currency	Currency (specify one)	
end group			

ODK Collect > household_survey

Section B - Water use practices > B.1.2
Bottled water expenditure

B.1.2.1 How much do you spend on bottled water per month

53.66

USD Local currency

Currency (specify one)

Formatting can also be done at the question level. For example, applying the minimal “widget” to the select_one currency question will make it show up as a dropdown list:

Excel 'survey' worksheet

type	name	label	appearance
begin group	hh_bottle_group	B.1.2 Bottled water expenditure	field-list
decimal	hh_bottle_expenditure	B.1.2.1 How much do you spend on bottled water per month	
select_one currencies	hh_bottle_currency	Currency (specify one)	minimal
end group			

ODK Collect > household_survey

Section B - Water use practices > B.1.2
Bottled water expenditure

B.1.2.1 How much do you spend on bottled water per month

65.32

Currency (specify one)

Select One Answer

ODK Collect > household_survey

Section B - Water use practices > B.1.2
Bottled water expenditure

B.1.2.1 How much do you spend on bottled water per month

Currency (specify one)

USD

Local currency

Remove response

ODK Collect > household_survey

Section B - Water use practices > B.1.2
Bottled water expenditure

B.1.2.1 How much do you spend on bottled water per month

65.32

Currency (specify one)

USD

ODK Collect offers many additional widgets that can change the survey formatting. Please look at the [ODK_1_1_7_widgets.xlsForm](#) shared on the [Formhub University account](#). The Formhub University is a public account on the site that presents advanced features in the form of examples so you can start making your surveys fit your exact needs.

4 - Making Dynamic Calculations

Your survey can perform dynamic calculations using the values of preceding questions. ODK Collect allows the use of a 'calculate' type, similar to any other question type but acts as a hidden field that will not show on the Android smartphone screen. The calculation can be referenced like any other question using the \${name_of_calculated_field} expression.

For example, in our survey, we want to calculate the number of children in the household and show it to the surveyor. This requires using the 'calculation' column. Once a value has been calculated it is not presented to the enumerator directly but can be presented in a note, as part of the label for next questions or as a conditional construct (constraint or relevant).

In this case, we want to calculate the number of children in the household based on the difference between the total number of people in the household and the number of adults in the household. We also want to present the results to the enumerator in a note. Our Excel 'survey' worksheet will look like this:

Excel 'survey' worksheet

type	name	label	calculation
integer	hh_people_nb	A.1.8 How many people live in this household?	
integer	hh_adult_nb	A.1.9 How many adults live in this household?	
calculate	hh_children_nb	(Label is not required because calculates do not display any content.)	$\$\{hh_people_nb\} - \$\{hh_adult_nb\}$
note	hh_summary_nb_display	Confirm that the household is composed of $\$\{hh_adult_nb\}$ adults and $\$\{hh_children_nb\}$ children	

The screenshots show the ODK Collect app interface. The first screen displays the question "A.1.8 How many people live in this household?" with the answer "11". The second screen displays the question "A.1.9 How many adults live in this household?" with the answer "3". The third screen displays a note stating "The household is composed of 3 adults and 8 children". All screens are titled "ODK Collect > household_survey" and show "Section A - Household Information".

The survey goes directly from question A.1.9 to the note. Behind the curtains the calculation was performed and saved in the variable 'hh_children_nb'. Notice how the note is presenting dynamic content and in particular notice that the enumerator never asked or entered any data about the number of children, the calculate field determined there were 8 children in this household.

	A	B	C
1	sectionA/hh_people_nb	sectionA/hh_adult_nb	sectionA/hh_children_nb
2	11	3	8
3	3	1	2
4			
5			
6			

The survey result data does not make any distinction between the data entered manually and a calculated variable.

More details about the language that javarosa uses, and the operators that are available are on [this page](#).

Warning: + - * correspond to add, subtract, and multiply. Division however, is special, and you need to use the word "div" to do division (for example, the ratio of households in the household could be calculated by the formula: $\$\{hh_adult_nb\}$ div $\$\{hh_people_nb\}$). This is a decision made by the xpath language, and unfortunately, we cannot do much with it in XLSform at the moment.

5 - Multiple conditions

Binding conditions and formulas can be as complex as you want. You can combine and, or and not logical operators with arithmetic operators. Our question B.1.1 allows to select multiple options for the sources of

water, but what if the respondent answers “I don’t know”? In this case we should probably add an additional choice with the “I don’t know option” but this adds one issue. The surveyor might check “House tap” and “Bottled water” and also check the “I don’t know” check box. This would not make much sense; you could add a condition to avoid this situation. The following example presents this situation. For simplicity purposes we have removed some of the options that we previously had defined. We added the ‘dk’ option at the end of the choices list.

Excel 'choices' worksheet

list name	name	label
water_sources	house_tap	Household tap
water_sources	bottled	Purchased bottle water
water_sources	other	Other
water_sources	dk	I don't know

Excel 'survey' worksheet

type	name	label	constraint
select_multiple water_sources	hh_water_source	B.1.1 Where do you get the water you use at home?	not (selected(.., 'dk') and (selected(.., 'house_tap') or selected(.., 'bottled') or selected(.., 'other')))

Notice how the constraint condition becomes quite long. This constraint is telling ODK Collect to allow the surveyor to move forward only if the combination of selected items is not “I don’t know” and any other option.

6 - Cascading selects

Now we want to present the interviewee with a series of selection questions, where the answer depends on a previous selection. The classic example is asking someone to enter a district in a country. But you don’t want the user to deal with 777 options at once (example from Nigeria), so you present enumerators with a question asking them about the zone they are in (6 in the country), the state they are in depending on the zone (less than 10 per zone), and then, finally, the LGA, depending on the state they are in.

This is possible with xlsforms without cascading select, but it requires you to write out one question per state in this case (36 different questions), with a fairly complex criteria in the relevant column. And even then, the end information you care about is what zone, state, and LGA (local government authority) the enumerator has filled out information about. Calculating this information requires writing a gigantic if statement, with the same number of branches as there are states in the country! This is quite an unimaginable process. We now present a syntax addition to overcome this issue and collect cascading-select information.

Since xlsforms are computer processed, we introduced a feature called cascading-select, where the user is able to simply list the cascading relationships. In a new worksheet called ‘cascades’ you create a new column ‘name’ and then an additional column for each level of the cascading hierarchy (if you are working with continents and countries you would have three columns ‘name’ ‘continent’ and ‘country’). In our case

we have three levels so we would need a total of four columns ‘name’ ‘zone’ ‘state’ and ‘lga’. The first non-header row is still a special row and has to start with the word ‘label’ under column ‘name’ then under each additional column you can enter some free-text with the text you want to display to the enumerator for each hierarchy level. The next rows will contain the relationships between the different elements in your hierarchy. State 1 for example has two LGA’s (1 and 2) and belongs to Zone 1. Notice how every element contains ‘choice_label’ under the ‘name’ column. Also notice that if there are 777 LGA’s in the country you will need 777 rows regardless of the number of Zones or States, the lower level in the hierarchy stipulates the size of the sheet.

Excel ‘cascades’ worksheet

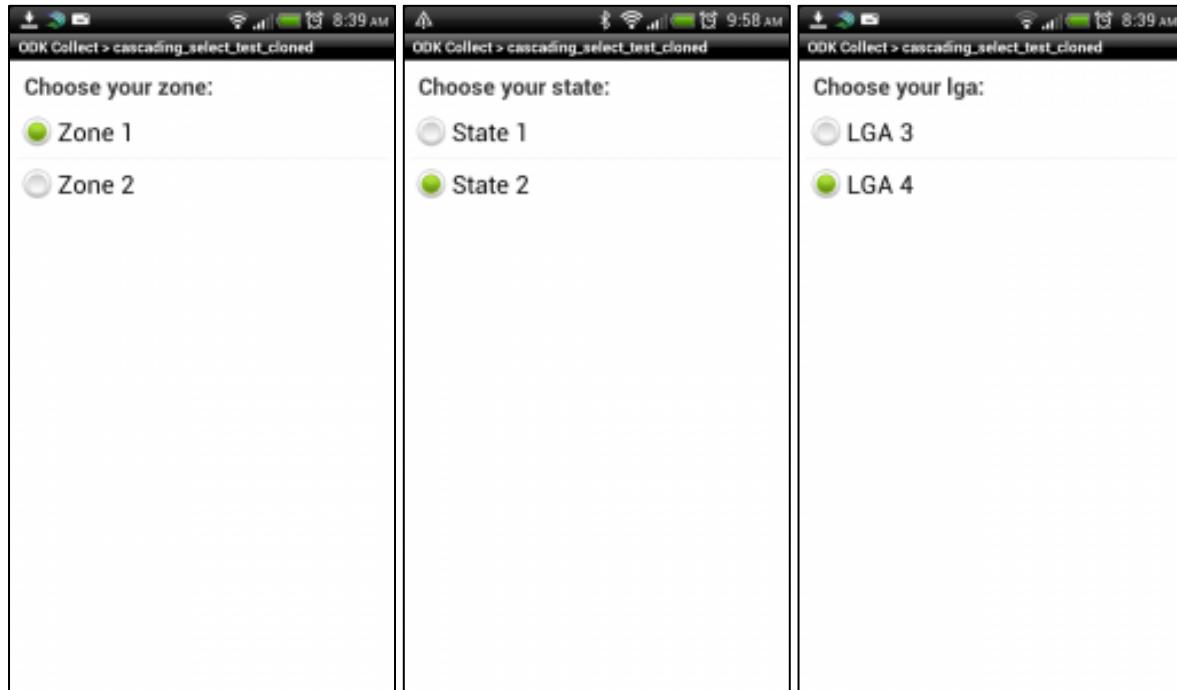
name	zone	state	lga
label	Choose your zone:	Choose your state:	Choose your lga:
choice_label	Zone 1	State 1	LGA 1
choice_label	Zone 1	State 1	LGA 2
choice_label	Zone 1	State 2	LGA 3
choice_label	Zone 1	State 2	LGA 4
choice_label	Zone 2	State 3	LGA 5
choice_label	Zone 2	State 3	LGA 6
choice_label	Zone 2	State 4	LGA 1
choice_label	Zone 2	State 4	LGA 2

In the survey sheet we will add a new row of type ‘cascading_select’ followed by the name of the column of the lowest level (rightmost column in the Excel ‘cascades’ worksheet) of the hierarchy (in this case ‘lga’). The ‘name’ of this row element is user defined, in this case ‘mylga’. For the content of the ‘label’ column in the Excel ‘survey’ sheet it is important to notice that the ‘cascading_select’ works differently than any of the question types we have been working with in the sense that a ‘cascading_select’ does not correspond to just one question but to as many questions as there are hierarchy levels defined in the Excel ‘cascades’ worksheet. In this case we have three levels so we will have three questions and thus three labels to display on the Android smartphone screen. This might sound confusing but it’s just how the cascading select works. The ‘label’ column on the Excel ‘survey’ sheet is overridden by the ‘label’ row on the Excel ‘cascades’ sheet. If you are confused just try it out and test it. You can then replace this example with your own cascading select.

Excel ‘survey’ worksheet

type	name
cascading_select lga	mylga

This will present the user with three questions (which one shows up totally depends on previous input).



If you want to know the zone, the state, or the LGA of your submission, you can just look for variables called mylga, mylga_state, and mylga_zone in the resulting data files. The “mylga” bit was obtained from the name assigned to the cascading select in the survey sheet, and the “state” and the “zone” respectively from the ‘name’ row in the Excel ‘cascades’ sheet.

mylga_zone	mylga_state_i n_zone_1	mylga_state_i n_zone_2	mylga_state	mylga_lga_in_ state_1	mylga_lga_in_ state_2	mylga_lga_in_ state_3	mylga_lga_in_ state_4	mylga
zone_1	state_2	n/a	state_2	n/a	lga_4	n/a	n/a	lga_4

The above is a simple test case, and may not seem impressive. But when you're dealing with 777 LGAs, like Nigeria, the feature can be very effective.

[1] Try it our yourself! Go to [the forms page on formhub](#), search for “cascade”, and clone any of the forms that show up there. The one I have worked above is available [here](#) or [here](#).

7 - Media

You can also make questions that show the user an ‘image’, a ‘video’ or an ‘audio’ file as a way of asking a question. In order to do this, you will need to put the media in /odk/formname-media folder on your phone, and reference them from xlsform in a media column.

8 - Logo

There is an easy way to show a logo at the top/beginning of your form. Simply upload an image file in the png format called form_logo.png. You can do this in the Media section of the form settings page.

Appendix I - All Question Types

text	Text input.
integer	Integer (ie, whole number) input.
decimal	Decimal input.
select_one [options]	Multiple choice question; only one answer can be selected (the [] notation indicates that you will need to replace [options] by a user defined choice list variable name).
select_multiple [options]	Multiple choice question; multiple answers can be selected (the [] notation indicates that you will need to replace [options] by a user defined choice list variable name).
note	Display a note on the screen, takes no input.
geopoint	Collect GPS coordinates.
image	Take a photograph.
barcode	Scan a barcode, requires the barcode scanner app is installed.
date	Date input.
datetime	Accepts a date and a time input.
audio	Take an audio recording.
video	Take a video recording.
calculate	Perform a calculation; see “calculates” below.

Appendix II - More Resources

Your first stop for more resources should be [formhub university](#), an account on formhub where we have put some example forms that delve deeper into some of the elements we have reviewed so far. You can download the forms to your Android smartphone and inspect them to learn about the most advanced features.

The Formhub community is very active and growing as users around the world start using the system and asking for improvements. We have created a google group where you can ask questions, review other users conversations about the most advanced features and solutions to common problems and of course meet people from all over the world who are working on mobile rapid data-gathering projects. Become a member of formhub-users@googlegroups.com and let us know about your experience with the platform. Another important resource is our [blog](#), Formhub is an ongoing project and every month we publish a roundup of the improvements made to the system. If you want to learn what's new this is the go to place.

The following resources will act as important reference once you have gone through the contents of this document:

- The [XLSform standard document](#) can guide you through the specific input types, column headers, and so on that are legitimate syntax in XLSforms.
- If you want to dig in deeper to understand xforms and go beyond xlsforms, here are some resources to understand them:
- Form guidelines: <http://opendatakit.org/help/form-design/>
- The javarosa formula language (used in constraint, relevant, and calculation expressions): <http://opendatakit.org/help/form-design/binding/> (See section under “operators”)
- Sample forms: http://formhub.org/formhub_u <http://formhub.org/forms>