

UNIVERSITY OF ILLINOIS  
AT URBANA-CHAMPAIGN

**Survey Design with Open  
Data Kit (ODK)**



[illinois.edu](http://illinois.edu)

# Outline for the Session

1. The *whats* and *whys* of ODK
2. ODK Environment
3. Building a survey for ODK



# The *whats* and *whys* of ODK



# What is ODK?

- “Open Data Kit is an open source suite of tools that helps organizations author, field, and manage mobile data collection solutions.”
  - Open-source suite of tools
  - Mobile data collection
    - *Android OS-Phones, Tablets*



# Why use paperless data collection, generally?

- Prevents loss of paper surveys
- Data quality controls
  - Data entry errors
  - Use of constraints and relevance
  - Ability to capture other media
    - *Voice, pictures, GPS*
- Quick access to data in a central location



## So then, why use ODK?

- Low cost of use
  - Open-source means free, with large online support system:  
[tsp://groups.google.com/forum/#forum/opendatakit](https://groups.google.com/forum/#forum/opendatakit)
  - Android is relatively low cost
- Effective for use in developing world
  - Limited technology infrastructure
  - Able to do large number of surveys for limited amount of money



# So what is ODK, again?

- “...open-source suite of tools that helps organizations author, field, and manage mobile data collection solutions.”



# ODK Environment





# ODK Environment

- **Build:** create data collection form or survey
  - Simple online system
  - XLSForm
- **Collect:** compile data on a mobile device and send it to a server
- **Aggregate:** put together collected data on a server and extract it in useful formats



BUILD



## ...to be continued...

network\_survey - Excel

File Home Insert Page Layout Formulas Data Review View FOXIT PDF Tell me what you want to do

Clipboard Font Alignment Number Styles Cells Editing

H14 You cannot select both None and a household

	A	B	C	D	E	F	G	H	I	J	K	L	M	
1	type	name	label	hint	default	appearance	constraint	constraint message	relevance	disabled	required	required message	read only	calculat
3	end	endtime												
4	deviceid	deviceid												
5	subscriberid	subscriberid												
6	smserial	simid												
7	comments	comment												
8	select_one_enumid	enumerator	Enumerator								yes			
9	select_one_village	village	Village			search('choices_exp','matches','list_name_key','vid')					yes			
10	select_one_partid	partid	Participant ID	Please select the participant that you are interviewing							yes			
11	select_one_hhid	hhid	Household			search('choices_exp','matches','list_name_key','hhid','village_filter',\$village)		You cannot select both None and a household			yes			
12	select_one_part_pic	nw0	Which one of these photos is you?			compact-3	not(selected(, '0') an household				yes			
13	note	nwON	The participant did not correctly identify themselves - please leave a comment on why you think that is.						not(selected(\$nw0,\$partid))					
14	select_multiple part_pic	nw1	Which one of these HOUSEHOLDS are your relatives?	Households are those that share the		compact-3	not(selected(, '0') an household	You cannot select both None and a household			yes			
15	select_multiple part_pic	nw2	Which one of these HOUSEHOLDS do you talk to on a regular basis?	About anything		compact-3	not(selected(, '0') an household	You cannot select both None and a household			yes			
16	select_one_yesno	microfin	Are you in a microfinance group?								yes			
17	select_multiple part_pic	nw3	Which of these HOUSEHOLDS are in your microfinance group?			compact-3	not(selected(, '0') an household	You cannot select both None and a household	\$microfin=1		yes			
18	select_one_yesno	farmgroup	Are you in a women's group / farming group?								yes			
19	select_multiple part_pic	nw4	Which of these HOUSEHOLDS are in your women's group / farming group?			compact-3	not(selected(, '0') an household	You cannot select both None and a household	\$farmgroup=1		yes			
20	select_multiple part_pic	nw5	Which of these HOUSEHOLDS' fields do you WALK/BIKE BY regularly?			compact-3	not(selected(, '0') an household	You cannot select both None and a household			yes			
								Please select only 3 households						



UNIVERSITY OF ILLINOIS AT URBANA-CHAMPAIGN

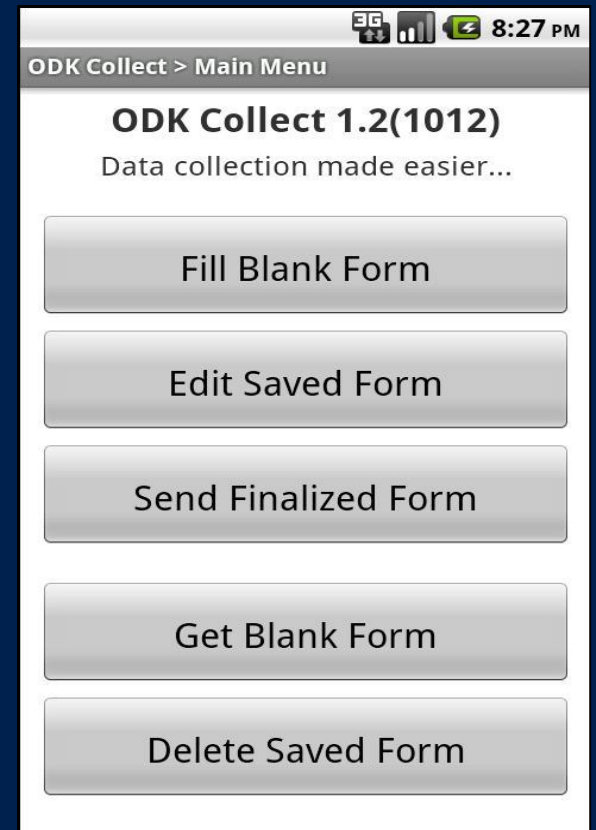
COLLECT



[illinois.edu](http://illinois.edu)

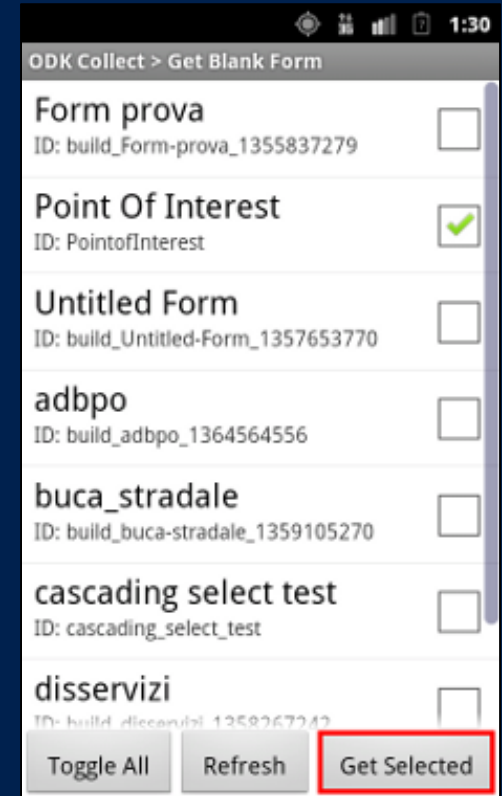
# COLLECT

- Download “ODK Collect” from Play Store
- This is what it looks like:

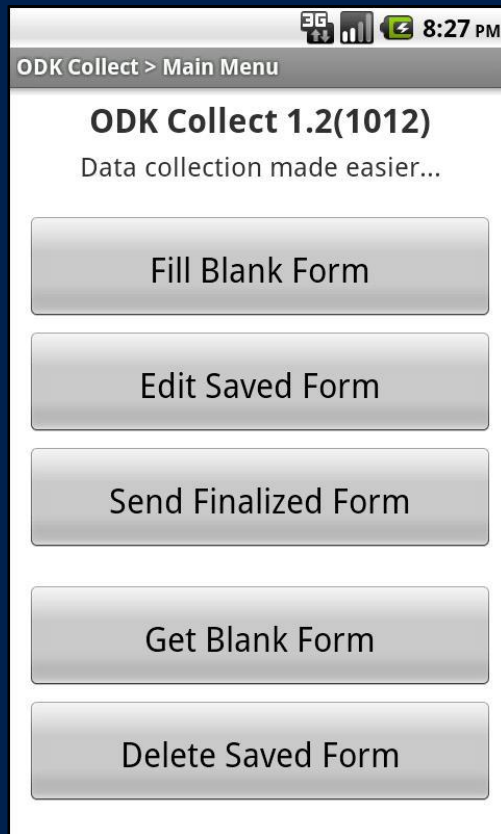


# COLLECT

- 1. GET BLANK FORM
- You will see a list of all the forms that have been uploaded on your device's server.



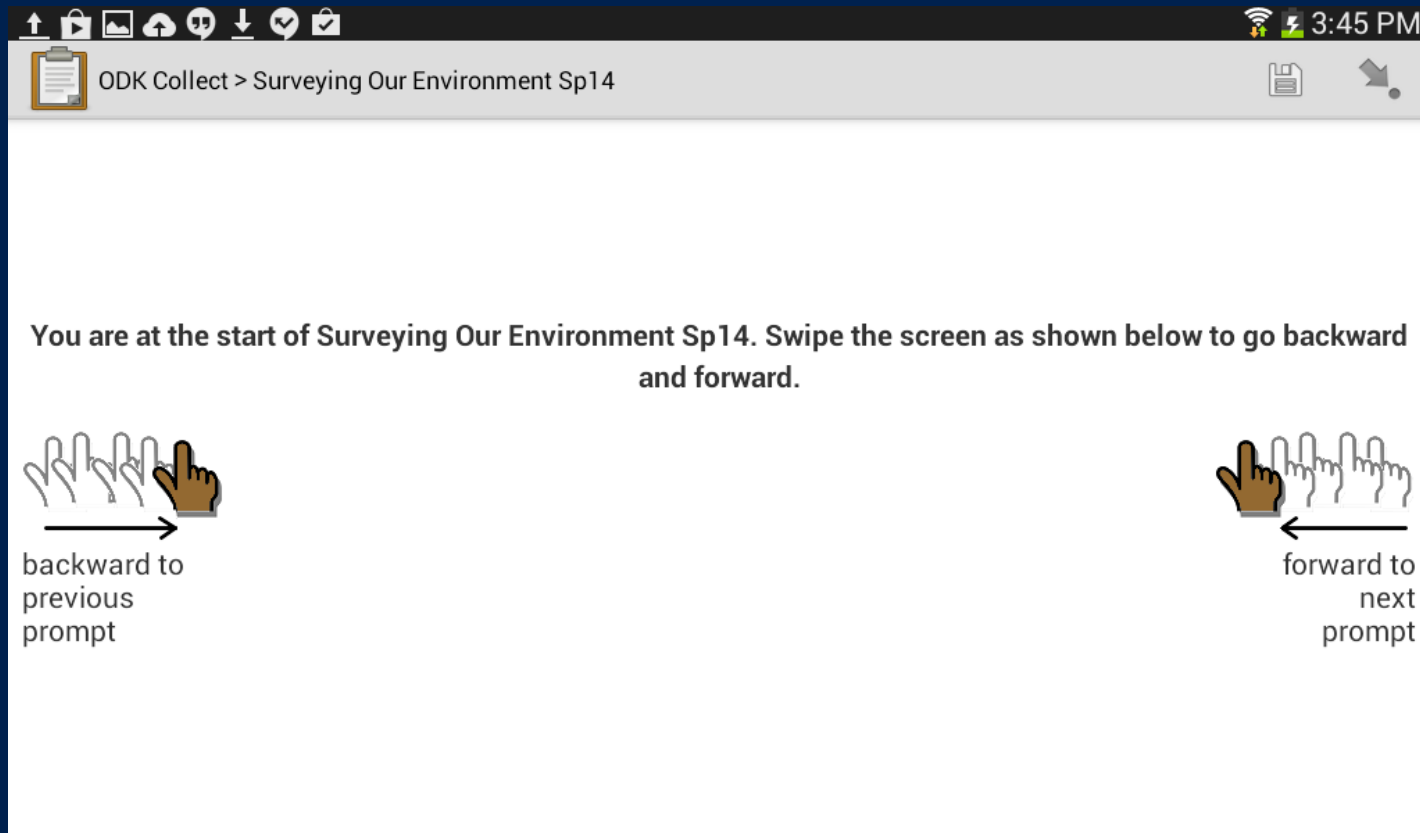
# COLLECT



- 2. FILL BLANK FORM
- Choose the latest version of the form, based on time and date
- Begin the survey!



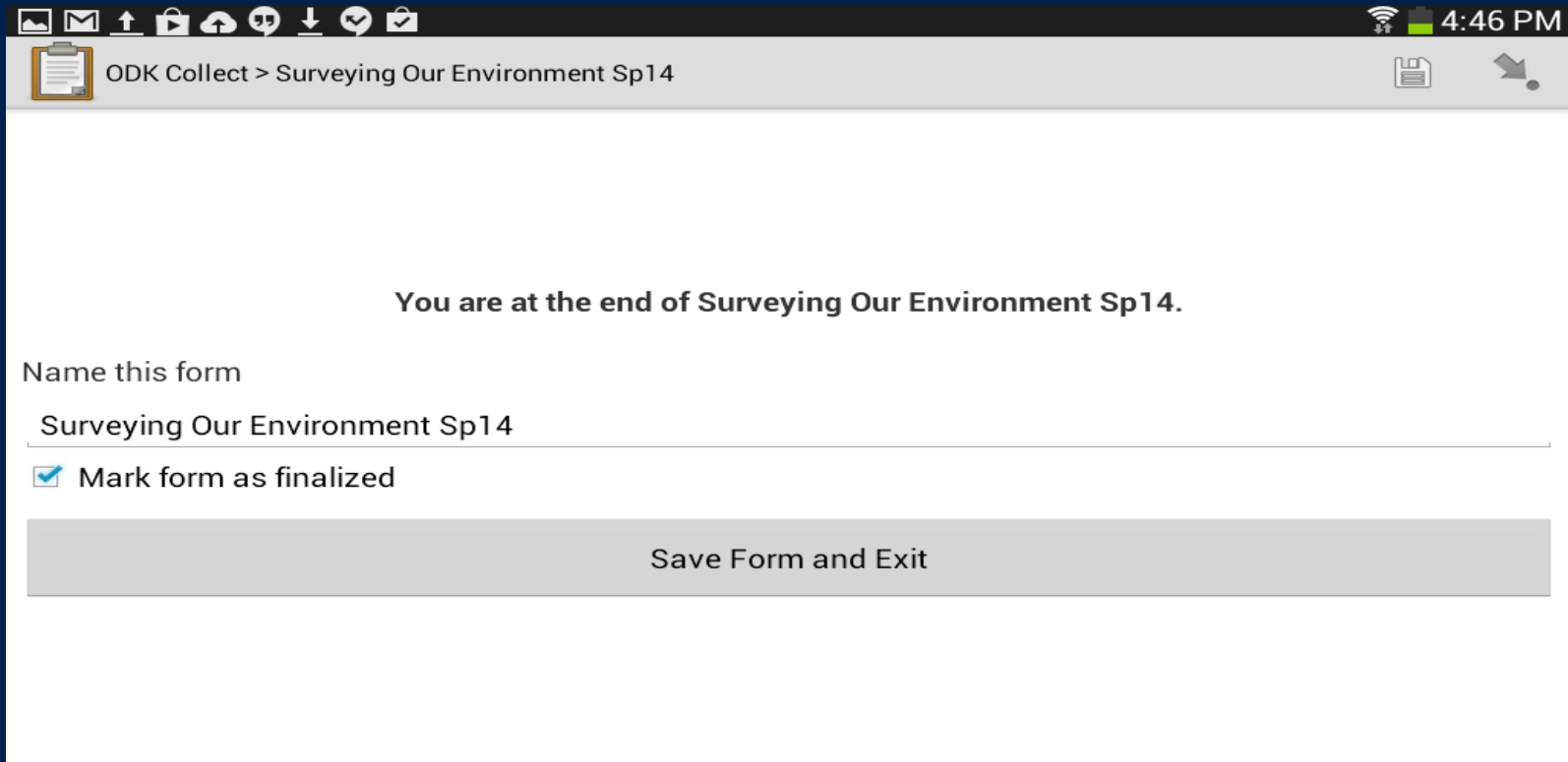
# COLLECT





# COLLECT

- Finalize data and send to server

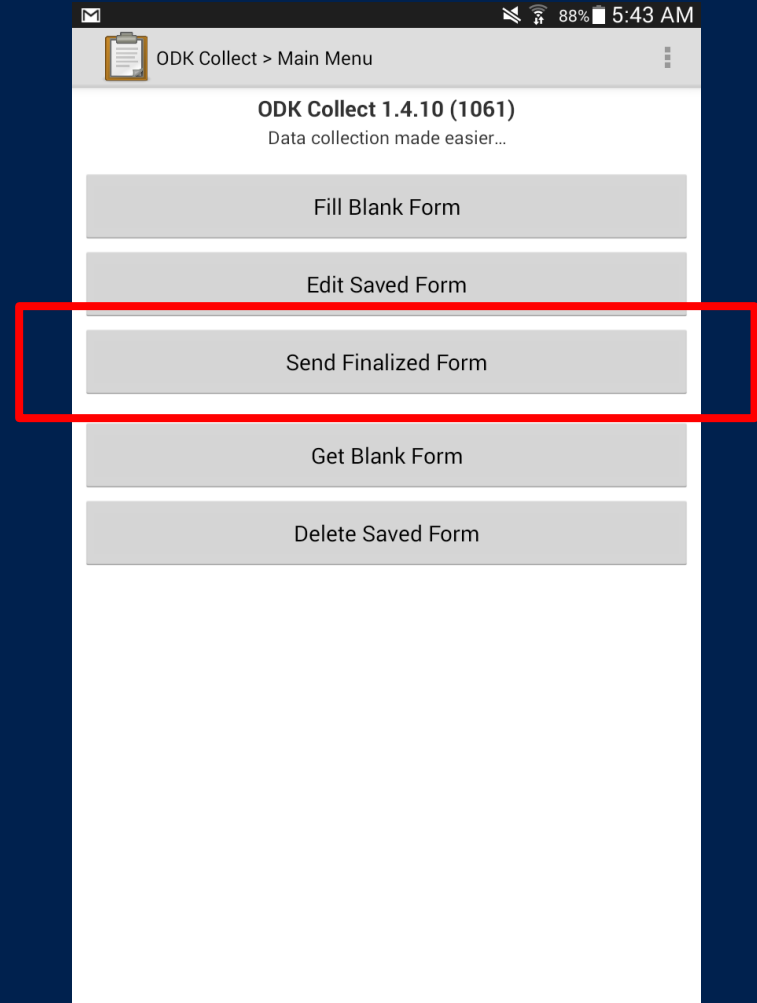


The screenshot shows the ODK Collect mobile application interface. At the top, there is a status bar with various icons and the time 4:46 PM. Below the status bar is a header bar with the text "ODK Collect > Surveying Our Environment Sp14" and a clipboard icon. The main content area displays the message "You are at the end of Surveying Our Environment Sp14." Below this message, there is a section titled "Name this form" with a text input field containing "Surveying Our Environment Sp14". Underneath the input field, there is a checkbox labeled "Mark form as finalized" which is checked. At the bottom of the form, there is a large grey button labeled "Save Form and Exit".



# COLLECT

- Send finalized form
- Download data from the server
- Export data into Stata / R/ desired software econometric package



UNIVERSITY OF ILLINOIS AT URBANA-CHAMPAIGN

# AGGREGATE



[illinois.edu](http://illinois.edu)

# AGGREGATE

- Provides blank forms to ODK Collect
- Accept finalized forms from ODK Collect and manage those data
- Visualize, export, and publish data
- <https://opendatakit.org/use/aggregate/>



## Alternatives...

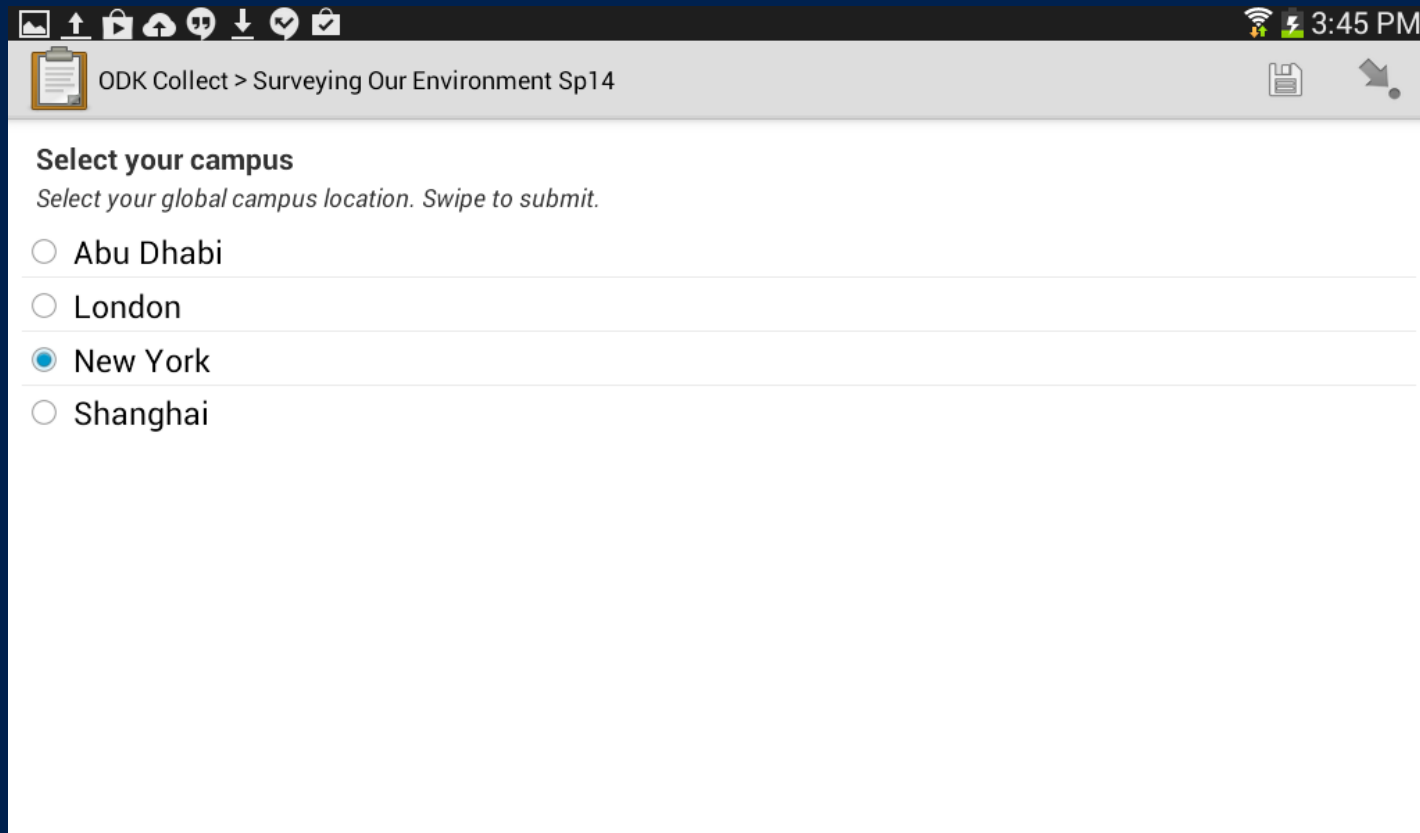
- ODK has aggregate
  - But, it is not always intuitive to use
- ONA, Survey CTO, Tomcat, other alternatives
  - Subscription
  - Front-end, as well as server space



# Building a survey for ODK



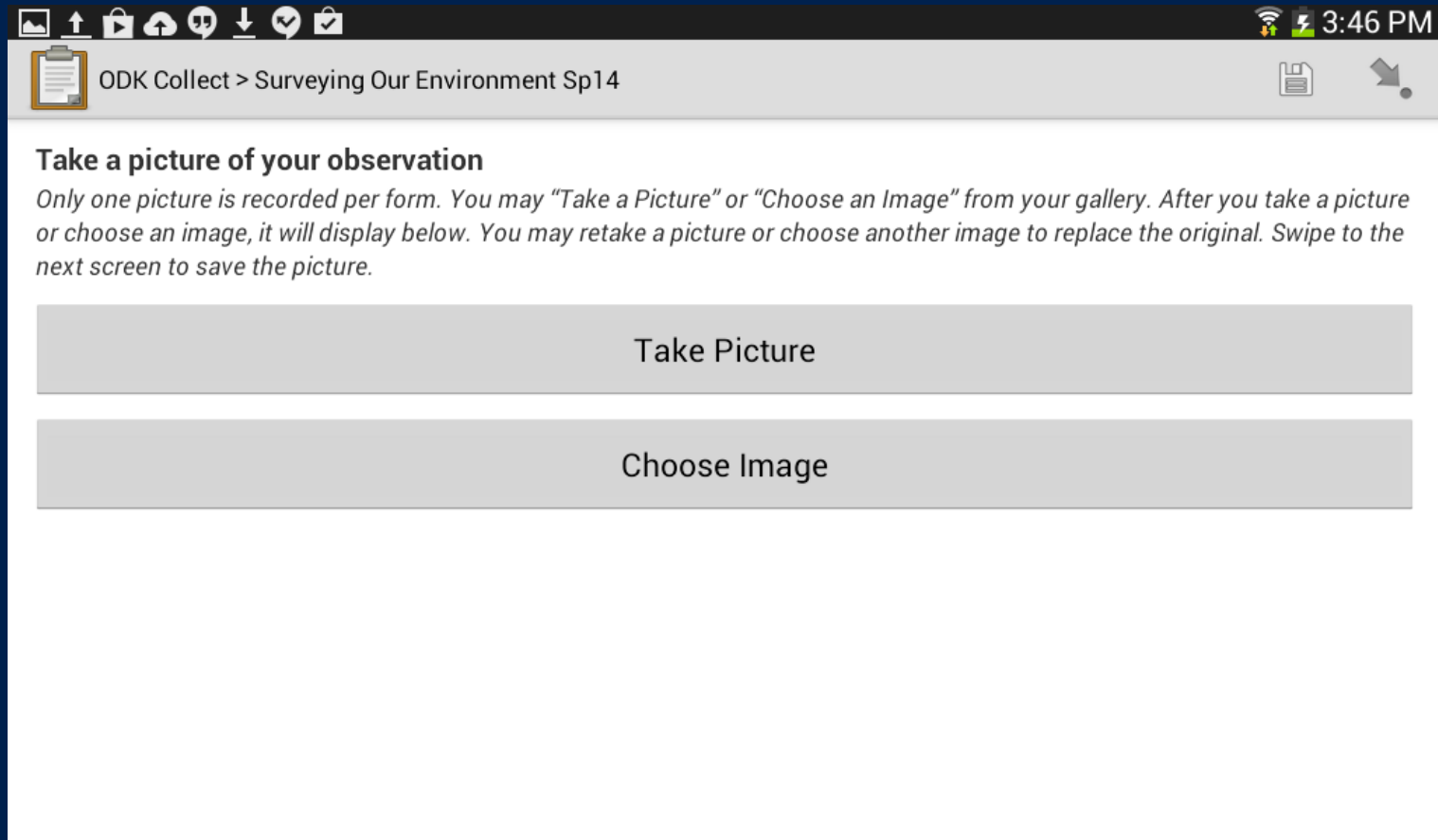
# Questions



The screenshot shows a mobile application interface for ODK Collect. At the top, there is a navigation bar with a clipboard icon and the text 'ODK Collect > Surveying Our Environment Sp14'. Below this, the survey question is displayed: 'Select your campus' followed by the instruction 'Select your global campus location. Swipe to submit.' There are four radio button options: 'Abu Dhabi', 'London', 'New York' (which is selected), and 'Shanghai'. The interface also shows standard mobile status icons at the top right, including signal strength, Wi-Fi, battery, and the time '3:45 PM'.



# Questions



ODK Collect > Surveying Our Environment Sp14

**Take a picture of your observation**

*Only one picture is recorded per form. You may "Take a Picture" or "Choose an Image" from your gallery. After you take a picture or choose an image, it will display below. You may retake a picture or choose another image to replace the original. Swipe to the next screen to save the picture.*

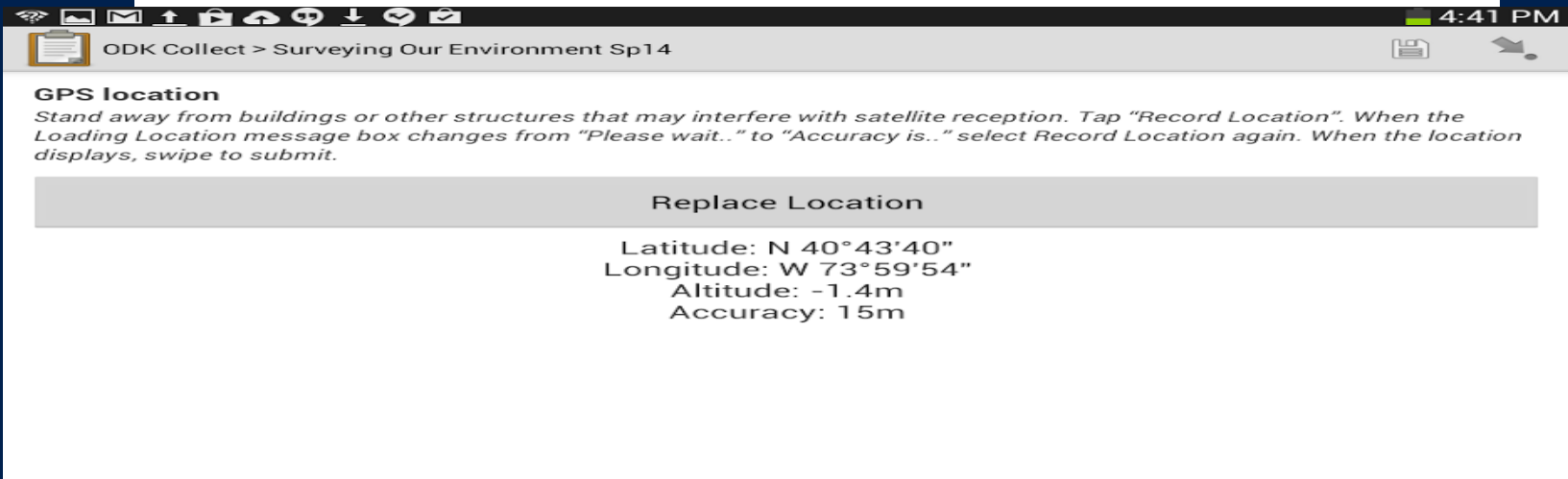
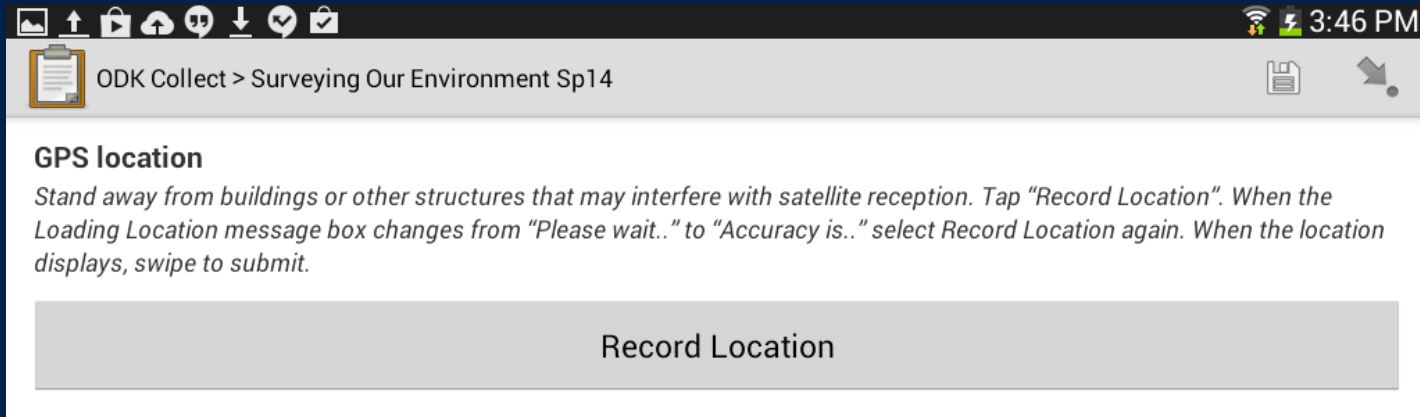
Take Picture

Choose Image

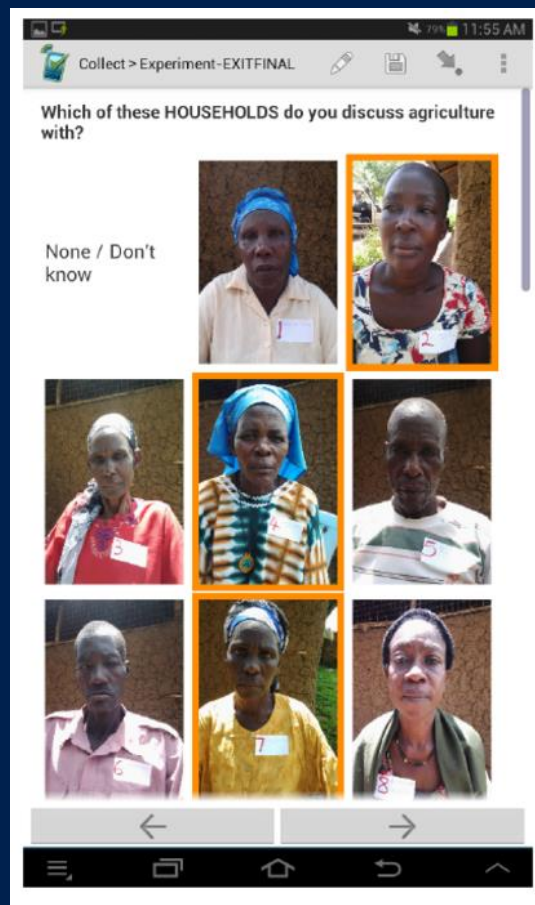




# Questions



# Questions



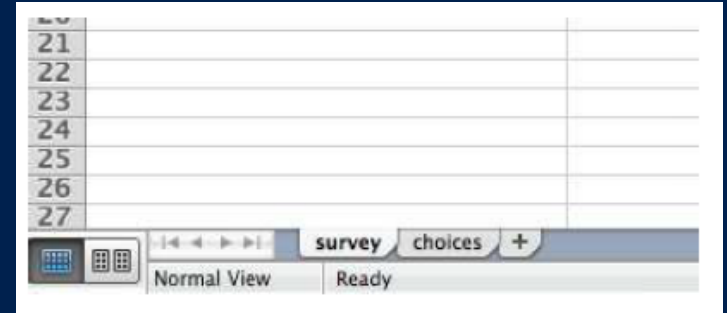
# XLSForm

- XLSForm is a standard form created to help simplify creating forms in Excel
  - Practical and familiar way to work on and share forms

	A	B	C	D	E	F	G
1	type	name	label	hint	constraint	constraint_message	required
2	text	some_text	This is a basic fill in the blank question.	Hints provide additional information about a question that is displayed less prominently.			
3	note	text_image_audio_video_test	This question shows how to use translations and media types.	In order to select a different translation press your phone's menu button and you should see a "Change Language" button. In order to add mediafiles put them in your phone's odk/forms/{formname}-media folder.			
4	integer	a_integer	Enter a integer:				
5	decimal	a_decimal	Enter a decimal:	This question has a constraint so that the decimal must be less that the integer.	. <= \${a_integer}	You can set the toast that appears when a constraint is violated.	
6	calculate	calculate					
7	note	calculate_test_output	The sum of the integer and decimal: \${calculate}				
8	select_one yes_no	select_example	This is a select_one type question.	In the type column you must include the name of a choices list that sets the available options. Choices lists are specified on the choices sheet.			
9	text	required_text	Required field	You have to enter something to continue.			yes
10	acknowledge	acknowledge_test	Acknowledge prompts can be used to indicate some action was completed. For example, a real prompt might read, "Inform the subject how the data collected on them will be used."				
11	select_one yes_no	skip_example	Skip the next question?				
12	text	skipable_question	This question will be skipped if its "relevant" column is false.	You can also have skip logic for groups.			



## Basic Format



21	
22	
23	
24	
25	
26	
27	

- **survey**: Overall structure, contains most of the content of the form
  - Includes full list of questions and information about how those questions should appear in the survey
  - 1 row = 1 question
- **choices**: Used to specify the answer choices for multiple choice questions.
  - Each row = an answer choice



## survey

	A	B	C
1	type	name	label
2	today	today	
3	select_one male_female	gender	Respondent's gender?
4	integer	age	Respondent's age?
5			
6			

- Three mandatory columns:
  - 1. **type**: specifies the type of entry you are adding
  - 2. **name**: specifies the unique variable name for that entry
    - *NO TWO ENTRIES CAN HAVE THE SAME NAME*
  - 3. **label**: contains the actual text you see in the form



## choices

	A	B	C
1	list name	name	label
2			
3	male_female	male	Male
4	male_female	female	Female
5			
6			
7			

- Three mandatory columns:
  - 1. **list name**: allows you to group a set of related answer choices
  - 2. **name**: specifies the unique variable name for that answer choice
  - 3. **label**: shows the answer choice as you want it to appear



## Basic Format

- Columns added may appear in any order
  - Optional columns can be completely omitted
- Rows can be left blank
- .xls format is ignored
  - Use dividing lines, shading, other font to make it readable to you
- **Syntax must be precise**
  - Choices ≠ choices



# UNIVERSITY OF ILLINOIS AT URBANA-CHAMPAIGN

Question Type	Answer Input
<b>integer</b>	Integer (whole number) input
<b>decimal</b>	Decimal input
<b>text</b>	Free text response
<b>select_one</b>	Multiple choice question – only one can be selected
<b>select_multiple</b>	Multiple choice question – multiple answers can be selected
<b>note</b>	Display note on the screen, takes no input
<b>geopoint</b>	Collect a single GPS coordinate
<b>geotrace</b>	Record a line of two or more GPS coordinates
<b>geoshape</b>	Records a polygon of multiple GPS coordinates, where the last point is the same as the first





Question Type	Answer Input
<b>date</b>	Date input
<b>time</b>	Time input
<b>dateTime</b>	Accepts a date and time input
<b>image</b>	Take a picture
<b>audio</b>	Take an audio recording
<b>video</b>	Take a video recording
<b>barcode</b>	Scan a barcode – this requires that a barcode app is installed
<b>calculate</b>	Perform a calculation
<b>acknowledge</b>	Acknowledge a prompt that sets value to “OK” is selected



survey

# ADDITIONAL OPTIONS



# GPS Options

- Able to collect with threshold of accuracy
  - Column: `body::accuracyThreshold`

survey			
	type	name	label
	text	store_name	What is the name of this store?
	geopoint	store_gps	Collect the GPS coordinates of this store.

survey				
	type	name	label	body::accuracyThreshold
	geopoint	store_gps	Collect the GPS coordinates of this store.	1.5



# Multiple Choice Options

- `or_other`
  - Create “**other**” option at the end of a list
- BUT:
  - Separate column is created
  - Not recommended for large data sets

survey			
	type	name	label
	select_multiple pizza_toppings	favorite_toppings	What are your favorite pizza toppings?
<b>choices</b>			
	list name	name	label
	pizza_toppings	cheese	Cheese
	pizza_toppings	pepperoni	Pepperoni
	pizza_toppings	sausage	Sausage

survey			
	type	name	label
	select_multiple pizza_toppings or_other	favorite_topping	What are your favorite pizza toppings?
<b>choices</b>			



# Metadata Options

Metadata Type	Meaning
start	Start date and time of the survey
end	End date and time of the survey
today	Day of the survey
deviceid	International Mobile Equipment Identity
subscriberid	International Mobile Subscriber Identity
simserial	SIM serial number
phonenummer	Phone number

survey			
	type	name	label
	start	start	
	end	end	
	today	today	
	deviceid	deviceid	
	subscriberid	subscriberid	
	simserial	simserial	
	phonenummer	phonenummer	



# Hints Option

<b>survey</b>				
	type	name	label	hint
	text	name	What is the name of this store?	Look on the signboard if the store has a signboard.
	geopoint	geopoint	Collect the GPS coordinates of this store.	

- Will add a hint to your question
  - Column: [hint](#)



# Constraints Option

survey					
	type	name	label	constraint	constraint_message
	integer	respondent_age	Respondent's age	. >=18	Respondent must be 18 or older to complete the survey.

- Improves data quality by including constraints
  - Column: **constraint**
- Can also add a message
  - Column: **constraint\_message**



# Relevant Option

- Creates additional feature based on response to a previous question
  - Column: **relevant**

	type	name	label	relevant
	select_one yes_no	likes_pizza	Do you like pizza?	
	select_multiple pizza_toppings or_other	favorite_topping	Favorite toppings	\${likes_pizza} = 'yes'
	text	favorite_cheese	What is your favorite type of cheese?	selected(\${favorite_topping}, 'cheese')
<b>choices</b>				
	list name	name	label	
	pizza_toppings	cheese	Cheese	
	pizza_toppings	pepperoni	Pepperoni	
	pizza_toppings	sausage	Sausage	





# Formulas Options

- Able to use formulas in **constraint**, **relevant**, and **calculations** columns
- Allows for additional functionality and another way of including data quality measures
  - Functions: +, \*, -, etc.



# Required Options

<b>survey</b>					
	type	name	label	constraint	required
	integer	age	How old are you?	, <= 150	yes

- User will not be able to move to the next question / submit the form, without entering an answer
  - Column: **required**: yes
- Also able to add required message
  - Column: **required\_message**



# Grouping Questions / Nesting in Group Option

survey			
	type	name	label
	begin group	hospital	Hospital
	text	name	What is the name of this hospital?
	begin group	hiv_medication	HIV Medication
	select_one yes_no	have_hiv_medication	Does this hospital have HIV medication?
	end group		
	end group		

- Put together related questions
  - Row: begin group
  - Row: end group
- Can nest groups within one another
  - Need a “begin group” for every “end group”



# Grouping Questions and Skipping Option

type	name	label	relevant
integer	age	How old are you?	
begin group	child	Child	$\${age} \leq 5$
integer	muac	Record this child's mid-upper arm circumference,	
select_one yes_no	mrtd	Is the child's rapid diagnostic test positive?	
end group			

- Skip a group of questions by combining group with relevant syntax
  - Put **relevant** column at the beginning of group



# Grouping Questions and Repeating Option

survey			
	type	name	label
	begin repeat	child_repeat	
	text	name	Child's name
	decimal	birthweight	Child's birthweight
	select_one male_female	sex	Child's sex
	end repeat		
<b>choices</b>	list name	name	label
	male_female	male	Male
	male_female	female	Female

- Repeat sets of questions for which you want multiple answers
- Can also limit number of repeats
  - Column: **repeat\_count**
    - *Set count (3) or dynamics ( $\${num\_hh\_members}\$$ )*



# Multiple Language Support Option

	type	name	label::English	label::Español	constraint
	integer	age	How old are you?	¿Cuántos años tienes?	. <= 150

- Able to add multiple languages
  - Column: `label::language1`, `label::language2`



# Media Option

survey					
	type	name	label	media::image	media::video
	note	media_example	Media example	example.jpg	example.mp4

- Include images / play media or audio files
  - Column: [media::image](#)
  - Column: [media::video](#)



# Upload Pre-Existing Data Option

survey				
	type	name	label	calculation
	calculate	fruit		<code>pulldata('fruits', 'name', 'name_key', 'mango')</code>
	note	note_fruit	The fruit \${fruit} pulled from csv.	

- Reference pre-existing data
  - Uploading the same way that one includes a media file
  - Column: **calculation**
    - *Command: `pulldata(...)`*





# Other Abilities

<b>survey</b>				
	type	name	label	default
	today	today		
	date	survey_date	Survey date?	2010-06-15

- Default
  - Question will be pre-populated with an answer
  - Helpful if there is a commonly selected answer
- Read only
  - Cannot be edited



# Other Abilities: Editing Appearance

Appearance Attribute	Question Type	Description
minimal	select_one, select_multiple	Answer choice appear in a pull-down menu
compact	select_one, select_multiple [choices with images]	Arranges image answer choices side by side
field-list	groups	Entire group of questions appears on one screen
label	select_one, select_multiple	Display answer choice labels
signature	image	Allows you to trace your signature on a form
draw	image	Allows you to sketch drawing with your finger on device screen



# setting

<b>settings</b>					
	form_title	form_id	public_key	submission_url	default_language
	Example Title	example_id	IIBljANBg...	https://example-odk- aggregate.appspot.com/submission	English

- Optional worksheet, allows further customization
  - Differentiate form identification
  - Encryption
  - Specify submission URL / name



# UNIVERSITY OF ILLINOIS AT URBANA-CHAMPAIGN

	A	B	C	D	E	F	G	H	I	
1	type	name	label	hint	constraint	constraint_message	required	appearance	default	relevant
2	text	some_text	This is a basic fill in the blank question.	Hints provide additional information about a question that is displayed less prominently.						
3	note	text_image_audio_video_test	This question shows how to use translations and media types.	In order to select a different translation press your phone's menu button and you should see a "Change Language" button. In order to add mediafiles put them in your phone's odk/forms/[formname]-media folder.						
4	integer	a_integer	Enter a integer:							123
5	decimal	a_decimal	Enter a decimal:	This question has a constraint so that the decimal must be less that the integer.	. <= \${a_integer}	You can set the toast that appears when a constraint is violated				
6	calculate	calculate								
7	note	calculate_test_output	The sum of the integer and decimal: \${calculate}							
8	select_one yes_no	select_example	This is a select_one type question.	In the type column you must include the name of a choices list that sets the available options. Choices lists are specified on the choices sheet.						
9	text	required_text	Required field	You have to enter something to continue.			yes			
10	acknowledge	acknowledge_test	Acknowledge prompts can be used to indicate some action was completed. For example, a real prompt might read, "Inform the subject how the data collected on them will be used."							
11	select_one yes_no	skip_example	Skip the next question?							
12	text	skipable_question	This question will be skipped if its "relevant" column is false.	You can also have skip logic for groups.						selected
13	begin repeat	repeat_test	A repeat							
14	text	repeating_question	This is a repeating question.	You can also repeat groups of questions.						
15	end repeat	repeat_test								
16	begin group	group_test	A group					field-list		
17	note	field_list_note	Using a group with a field-list appearance you can make multiple prompts appear on the same screen.							
18	select_multiple yes_no	select_multiple_1	Select multiple 1		not(selected(., 'yes') and selected(., 'no'))	You can't select yes and no.	no	minimal		
19	select_multiple a_b	select_multiple_2	select multiple 2							
20	end group	group_test								

