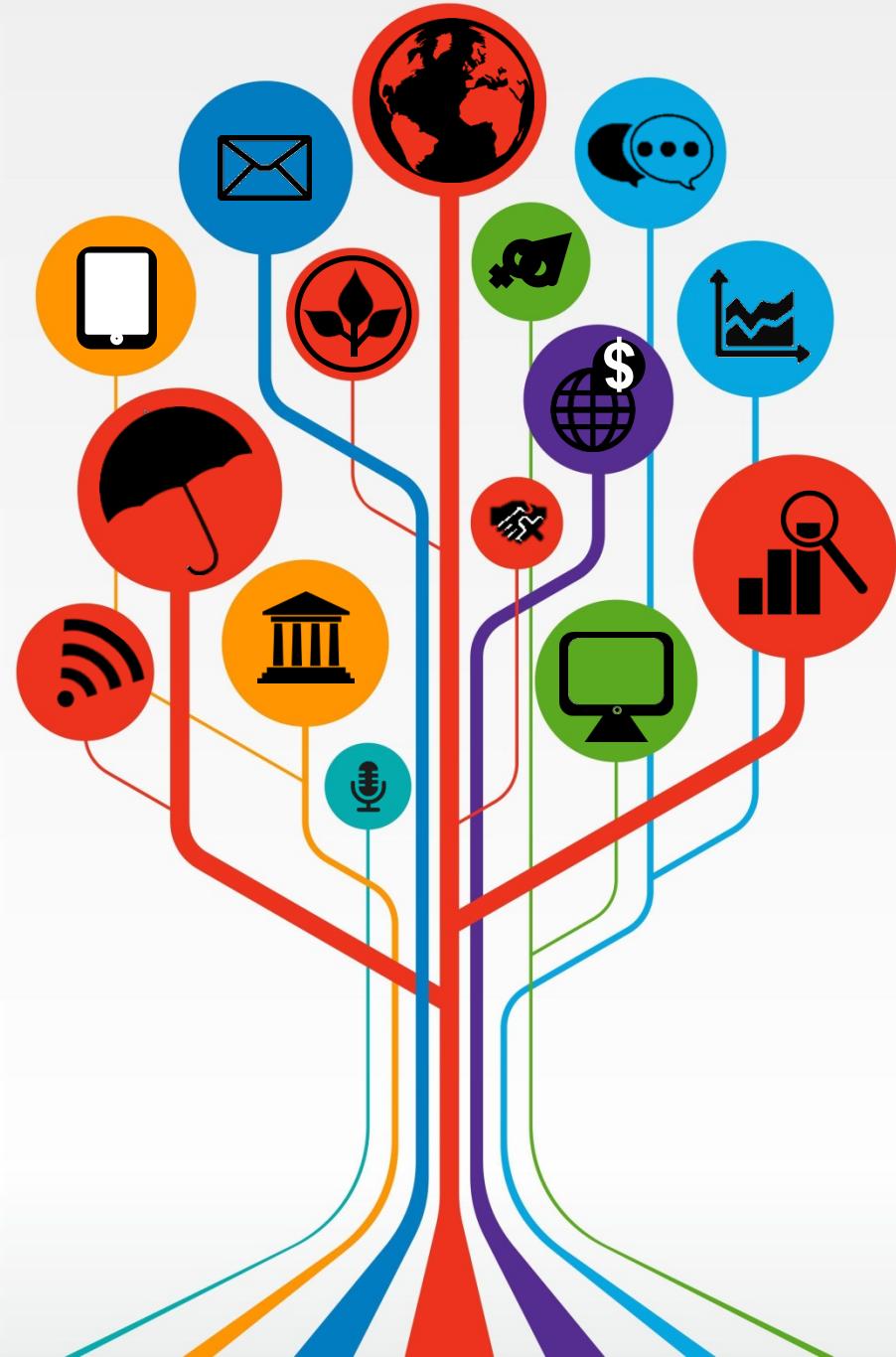


FIELD COORDINATOR WORKSHOP

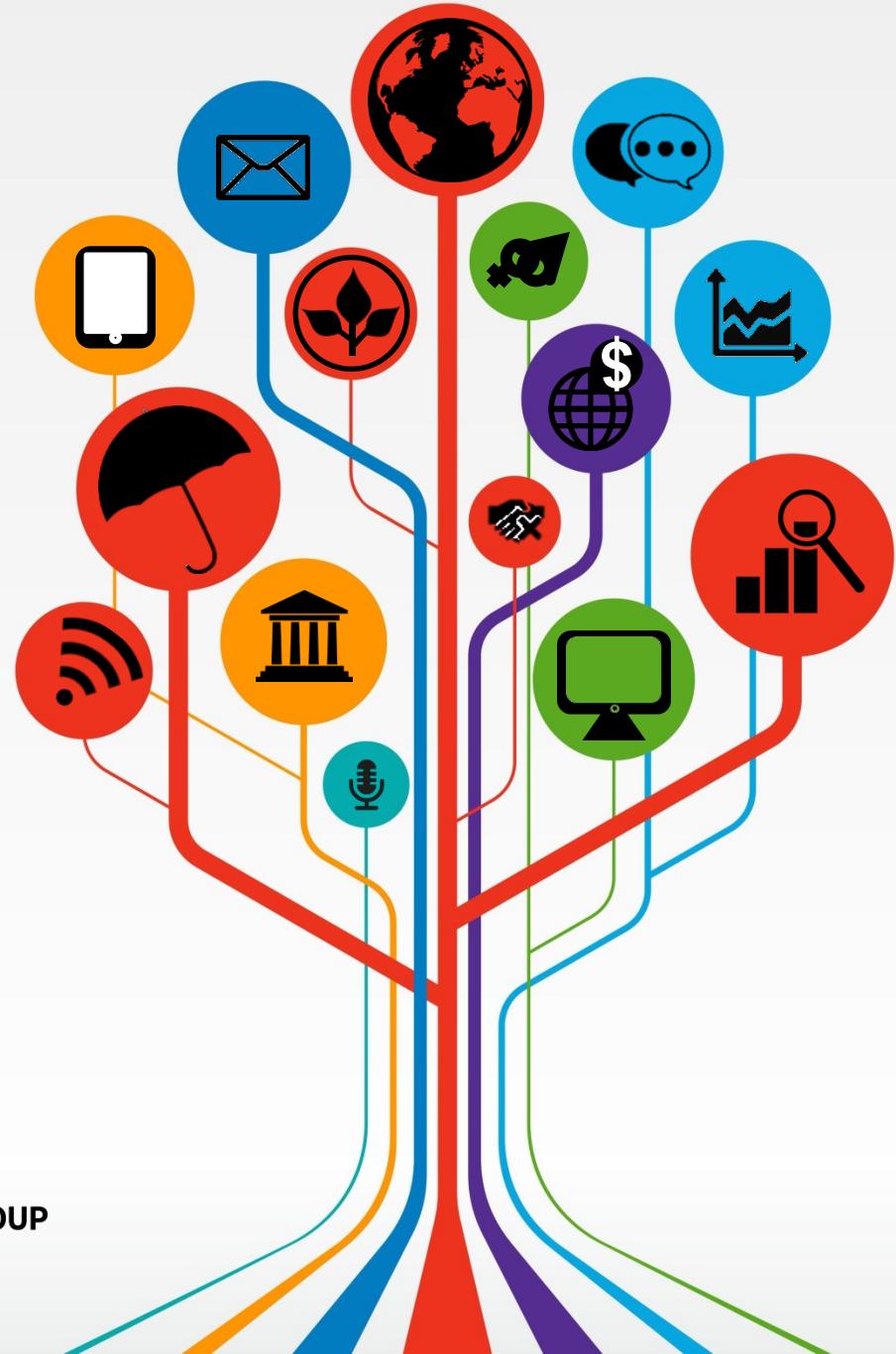
Manage Successful
Impact Evaluations

18 - 22 JUNE 2018
WASHINGTON, DC



Data Quality Assurance Plan (1/2)

Laura Costica
Aurelie Rigaud
Steven Glover
19th June 2018





Data Quality

Think of everything that might go wrong ...



Data Quality

Think of everything that might go wrong ...

Answers entirely
made_up

Attrition

Tablet is broken

Incorrect recording
of answers

Respondent doesn't
understand the
question

Wrong person
interviewed

Programming
bug

... and make a plan for pre-empting it



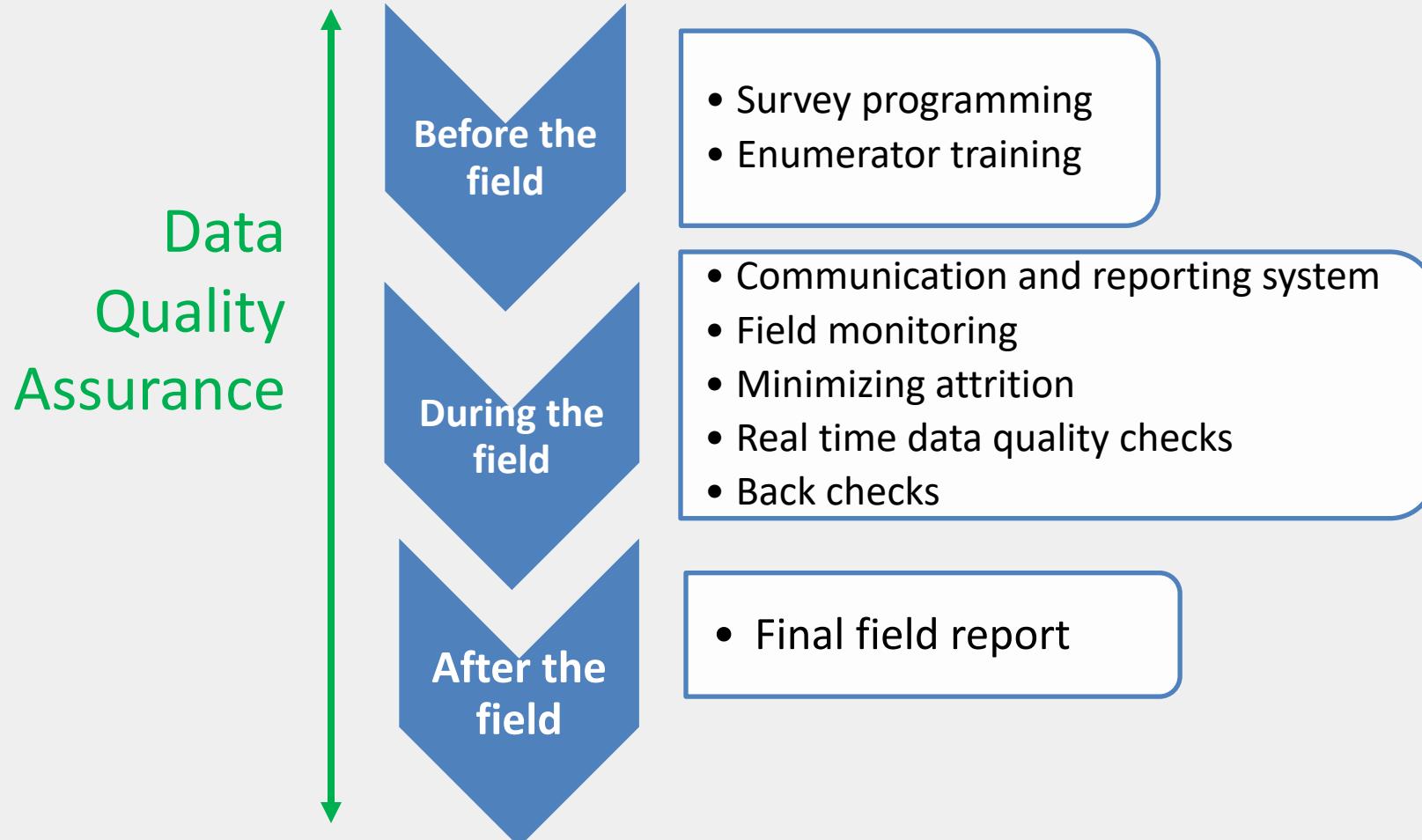
Data Quality

“The quality of the data we collect plays a key role in driving the quality of our decision-making”

Christopher Robert, SurveyCTO 2017

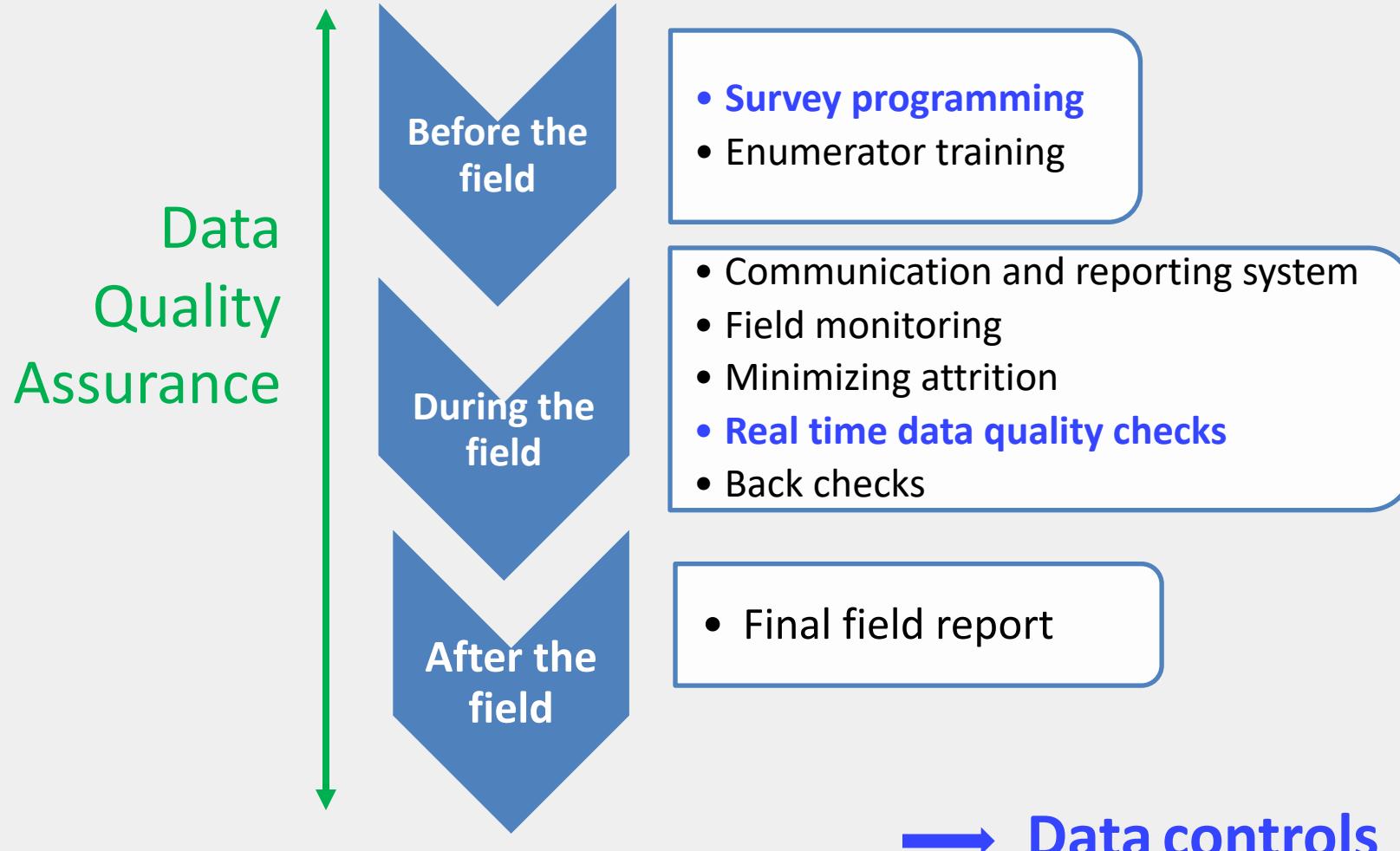


Content



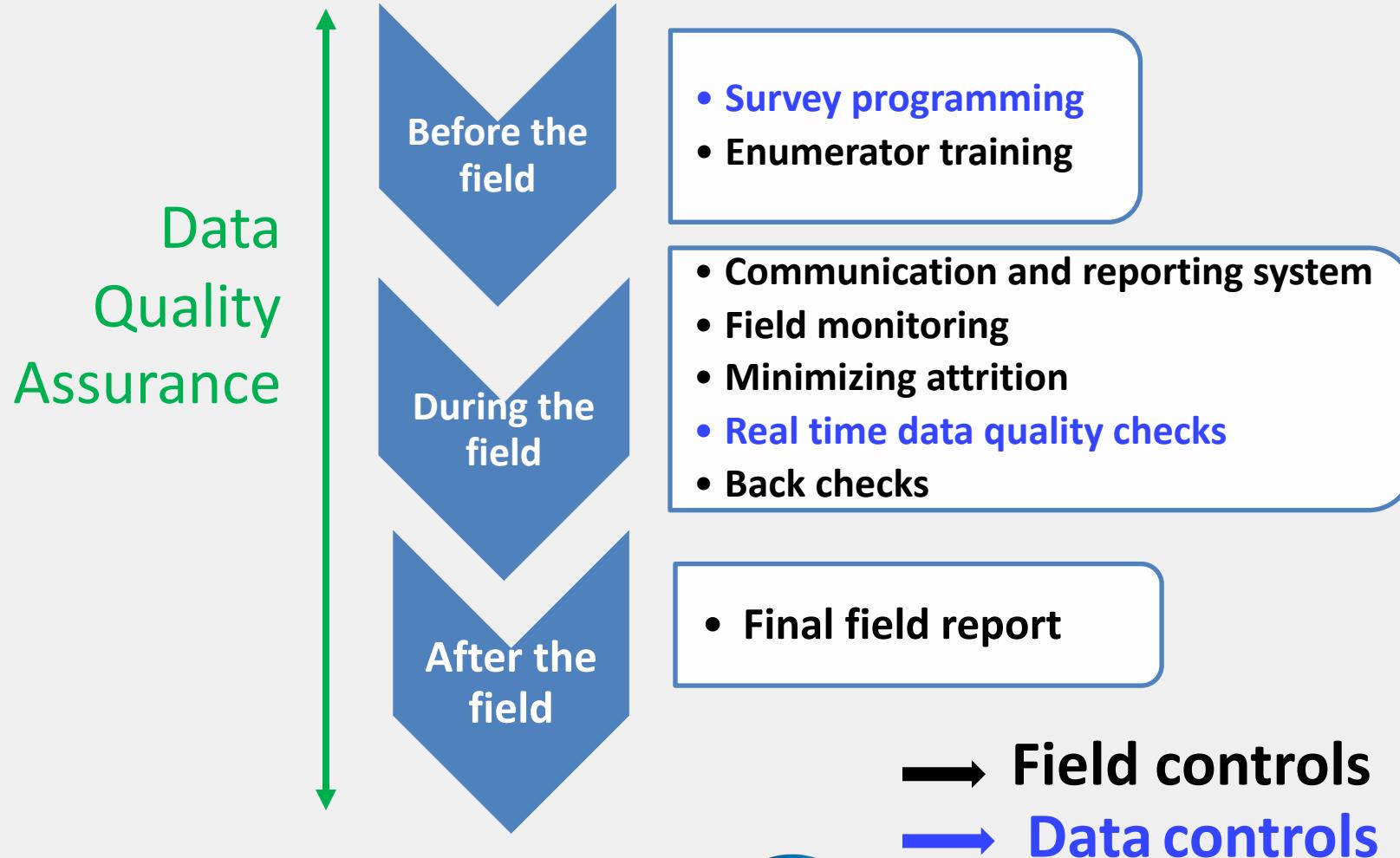


Content





Content





Data Quality

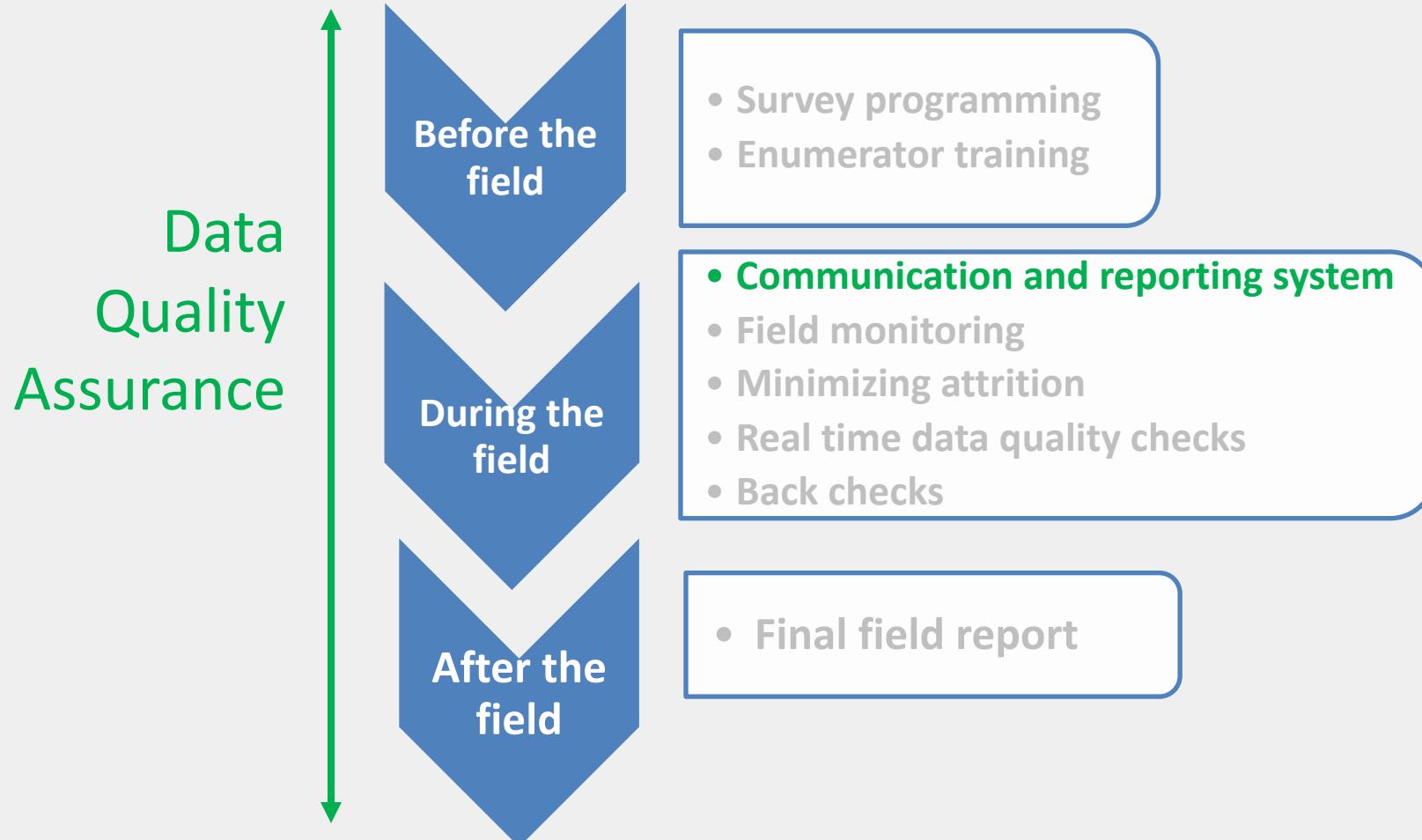
Part 1- Field Controls

Field Controls - Content



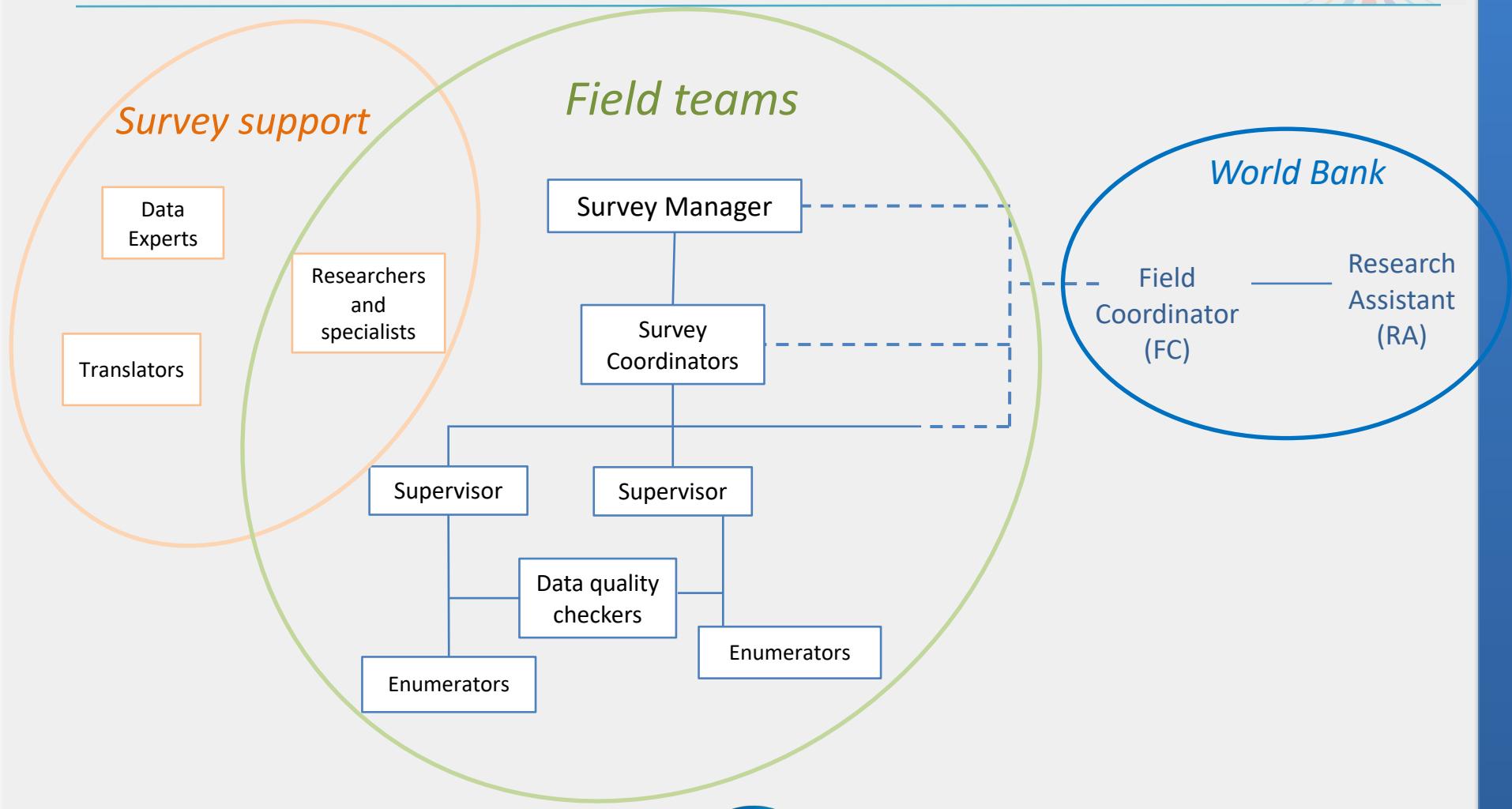
- Communication and Reporting System
- Field Monitoring
- Back Checks
- Minimizing attrition
- Final Field Report

Communication and reporting system



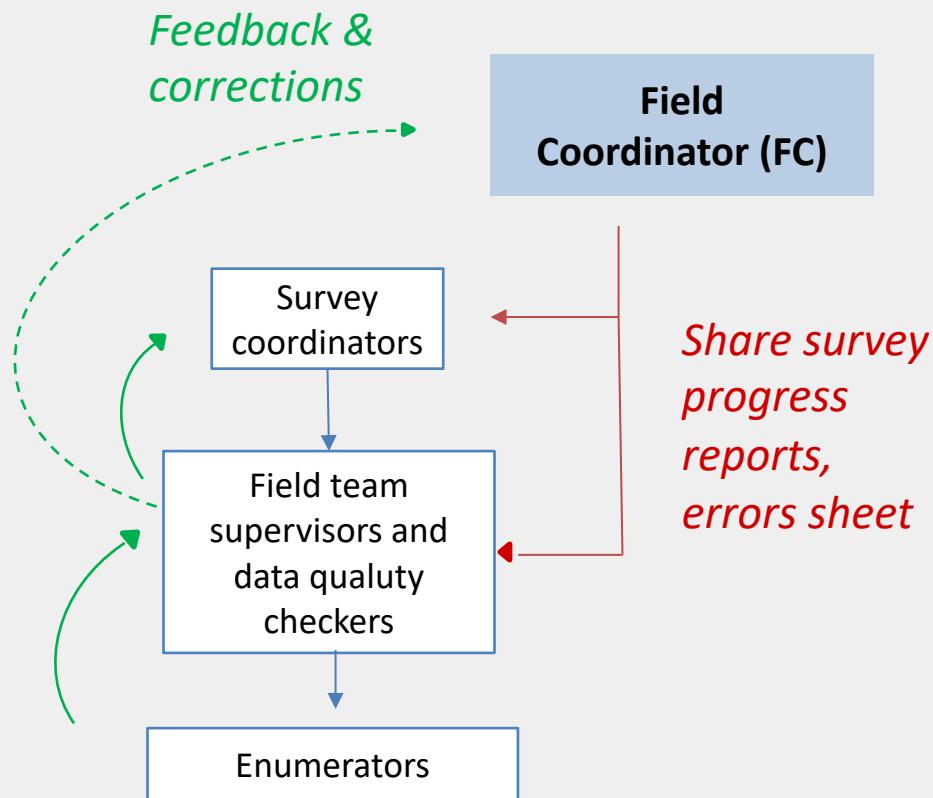


Staff Organigram



Good communication is key!

→ Within the team



- Share reports, observations, errors sheets, etc. with the survey coordinator, and/or directly to the field team supervisors (make sure the survey firm agrees with this communication channel)
- Enumerators usually should not report to you, field team supervisors can

Good communication is key!

→ Within the team



- Set up a good system for reporting feedback!
 - Create **Whatsapp group** with you, survey managers/coordinators and supervisors
 - Whatsapp conversation can be exported & stored on DB
 - Each team meets at the end of the days for feedbacks, **sharing experiences** on challenges and success
 - **Get in touch with** your survey manager / coordinators everyday for feedback if you couldn't meet them during the day
 - Keep a **diary** or **log** of conversation with teams

Good communication is key!

→ Within the team



Surveys in remote areas can present connectivity issues:

- Cannot send survey forms every day
- Receiving and sending HFCs and backcheck feedback
- Charging tablets
- Phone network

What can you do?

- Train the supervisors well! – many problems don't need FC interaction to resolve
- Set maximum period before having to check in again / sync tablets (not > 48 hours)
- Be creative

Good communication is key!

→ With the respondent



- Ensure the field staff knows how to interact with the respondent :

In the
Consent
Form



- About the project and the evaluation
- The importance to participate (cooperation techniques)
- The confidentiality of the data
- With neutrality and respect
- Fluency in respondent's language



Training, practice, and field observations...

Good communication is key!

→ With the respondent



Translation issues:

- Many local languages, often can't be written.
- Incorrect translation leading to bad understanding and impact data quality

Good communication is key!

→ With the respondent

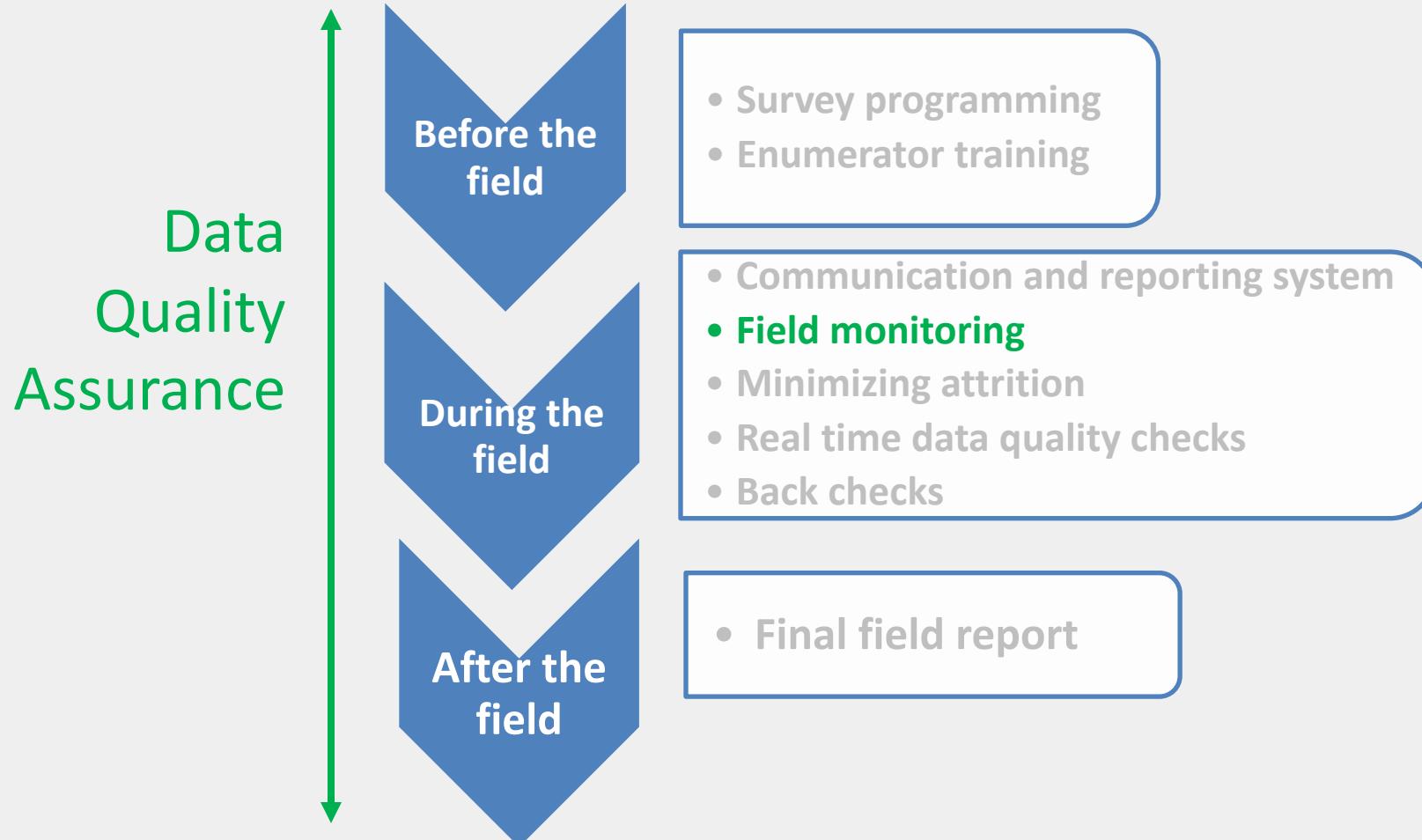


How to deal with the translation difficulties?

- During training:
 - Group sessions for translation between enumerators speaking the same local language
 - Invite the project staff in the translation session!! (They know the project's technical terms)
 - Translation on key words - make cheat-sheets
- Tablets:
 - Translated questionnaire can be programmed in the form using SurveyCTO
 - NOT recommended to program survey in a mostly unwritten language – use ‘hints’ for keywords instead
- During interview: enumerator can be helped by people of the community
- Audio records?



Field monitoring





Field Monitoring: *Field Coordinator*

- Important to follow the teams at the beginning of the data collection:
 - Give useful feedback after (not during) interviews
 - Check if questions are understood correctly by the respondent
 - Ensure reporting systems are set up and understood by supervisors / data quality checkers
 - Maintain good relationship with surveyors and strengthen staff motivation



Field Monitoring: *Supervisors*

- Supervisors are experienced enumerators who have diverse tasks, the most important of which are:
 - Organizing logistics (workplan and movement, materials, accommodation...)
 - Presenting the enumerator teams to local authorities
 - Monitoring activities
 - Observing interviews and supporting their team members in perfecting their interviewing abilities
- Give supervisors a form to write down their observations about each enumerator and ask them to submit these forms



Observing interviews

Example: Enumerator observation form

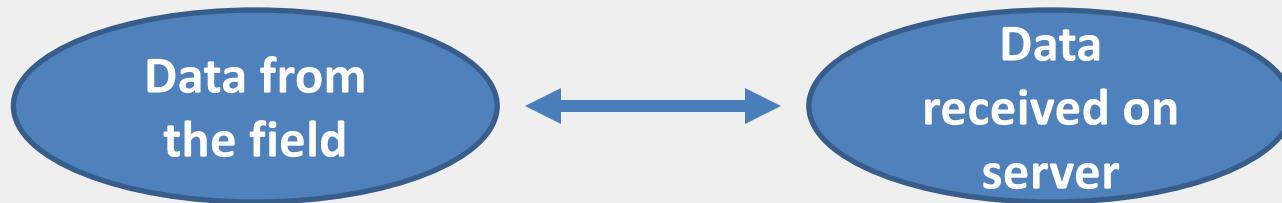
Description	Score (1-5)
Presentation	
Does the enumerator have all the required equipment?	
The enumerator correctly introduces himself / herself to the respondent and correctly asks the respondent for his / her informed consent.	
Consistency	
The enumerator reads out the questions the way they appear in the questionnaire.	
The enumerator prompts and probes the respondent at the right times and to the right extent.	
The enumerator correctly reads the introductory sentences at the beginning of each section before asking the questions.	
Interaction	
The enumerator is familiar and confident with the questionnaire and how to ask the questions.	
The enumerator is proficient in the language to communicate the questions effectively.	
The enumerator shows patience and respect towards the respondent.	
The enumerator remains neutral / objective throughout the interview.	
The enumerator is able to keep the attention of the respondent well and to keep the interview going at a good pace.	
The enumerator creates an environment of comfort and privacy for the interview.	
COMMENTS	
TOTAL	

- Each surveyor should be observed at least one time in the initial days of the survey
- Fill in an enumerator observation form



Field Monitoring: Survey Progress

- Ensure surveys conducted match the original respondents list → good reporting system



- Reconcile registers of completed survey with the field teams **every day**. Have an up to date list always accessible to the supervisors.
- Missing cases can create major issues with the survey firm
- Often have cases where surveys don't leave the tablets, forms not finalized, forget to add replacements, etc.



Field Monitoring: Survey Progress

Example: Log book in Google Sheets, not shared with supervisors

Sample					Server					Back-check				
assoc	hhid	name	prov	sample	status	comments	nbr_visits	visit1	visit2	visit3	status	notes	selected_backch	date_backch
BADZA ROTANDA	10101	JOREGE DENIASSE	Manica	1	Complete		1	5/29/2018			Interviewed	Terminada		
BADZA ROTANDA	10102	TICHA JOAO	Manica	1	Complete		1	5/29/2018			Interviewed	O inquérito correu com sucesso!		
BADZA ROTANDA	10103	GUIVIMO FARAI	Manica	1										
BADZA ROTANDA	10105	CHAITA SAISSE	Manica	1										
BADZA ROTANDA	10106	DAVIDE JUDA	Manica	1	Complete		1	6/1/2018			Interviewed	Nada a comentar		
BADZA ROTANDA	10108	MARIA MOISES	Manica	1										
BADZA ROTANDA	10109	MANUEL GOSTAVO	Manica	1										
BADZA ROTANDA	10110	VAILETE QUEFASSE	Manica	1	Complete		1	5/31/2018			Interviewed	Entrevista termin		1
BADZA ROTANDA	10111	HELENA SIEREMBE	Manica	1	Complete		1	5/29/2018			Interviewed	Nada a comentar		
BADZA ROTANDA	10113	DORCA ALFACE	Manica	1	Complete		1	5/30/2018			Interviewed	Nada a comentá		1
BADZA ROTANDA	10114	IDISSE HELTON	Manica	1	Complete		1	5/30/2018			Interviewed	Nada a comentar		
BADZA ROTANDA	10115	BEATRIZ QUEFASSE	Manica	1										
BADZA ROTANDA	10117	MARTA RUISENNE	Manica	1	Complete		1	5/29/2018			Interviewed	Nada a comentar		1
BADZA ROTANDA	10118	FARAI MATSVAI	Manica	1	Complete		1	5/30/2018			Interviewed	A entrevista correu bem		
BADZA ROTANDA	10120	ROSITA CHEMADA	Manica	1										
BADZA ROTANDA	10121	MARIA CHPAUMIRE	Manica	1										
BADZA ROTANDA	10122	WANDA LAVU	Manica	1	Complete		1	5/29/2018			Interviewed	A correu bem		1
BADZA ROTANDA	10124	AUGUSTO FARAI	Manica	1										
BADZA ROTANDA	10125	CHENJERAI FARAI	Manica	1										
BADZA ROTANDA	10127	SUZENE FIBIONE	Manica	1	Complete		1	5/30/2018			Interviewed	A entrevista corre		1
BADZA ROTANDA	10128	ELISA DAVIDE	Manica	0										
BADZA ROTANDA	10129	MATEUS WANDA	Manica	1	Complete		1	5/29/2018			Interviewed	Nada a comentar		
BADZA ROTANDA	10130	MEQUE NEUVICO	Manica	1	Complete									
BADZA ROTANDA	10131	FILIPE CHIMORE	Manica	1	Complete		1	5/30/2018			Interviewed	Nada a comentar		
BADZA ROTANDA	10132	LUCAS ANTONIO	Manica	0										
BADZA ROTANDA	10134	SUZENE ZOMBA	Manica	1										



Field Monitoring: Survey Progress

Completed surveys at
the end of the day

Finalizados e enviados

45103
45112
45116
45118
45124
45127
45136
45143
45144
45145
45147
45180
45181
45185
45187

13:44

Recebidos, obrigado!

14:46 ✓

Luiza Andrade

IDs para back check:

45147
45144
45192
45143
45103

20:48

Manuel Cossa

Recebido. Obrigado.

23:07

Replacements, IDs to
backcheck, other issues

Desconhecidos

46139
46140
46141
45183
45184
45201

Mesmo AF

44101>4413 (São cônjuges. Será entrevistado o
44101)

13:31

44101>4413 (São cônjuges. Será entrevistado o
44101)

--> Confirme o ID, 4413 falta digito!

15:01 ✓

46102 -> 44165 (Paz)
46139 -> 43144 (Mziva 2)
46140 -> 68202 (Thewe)
46141 -> 47282 (Aremo)
45183 -> 42164 (Mziva 1)
45184 -> 42156 (Mziva 1)
45201 -> 68111 (Thewe)

15:04 ✓

Manuel Cossa

You

44101>4413 (São cônjuges. Será entrevistado o
44101) --> Confirme o ID, 4413 falta digito!

Vou confirmar mas acho que é 44113. Mas cou
confirmar.

15:13



Field Monitoring: *Data Quality Checkers*

- Report any concerns about **data quality** (errors and inconsistencies found in the data) to data quality checkers
- (Ideally) have a separate team of backcheckers and auditors to:
 - Peer edit and questionnaire review before submission
 - Process errors and correction sheets from HFCs
 - Interview observation and monitoring
 - Spot checks and backchecks



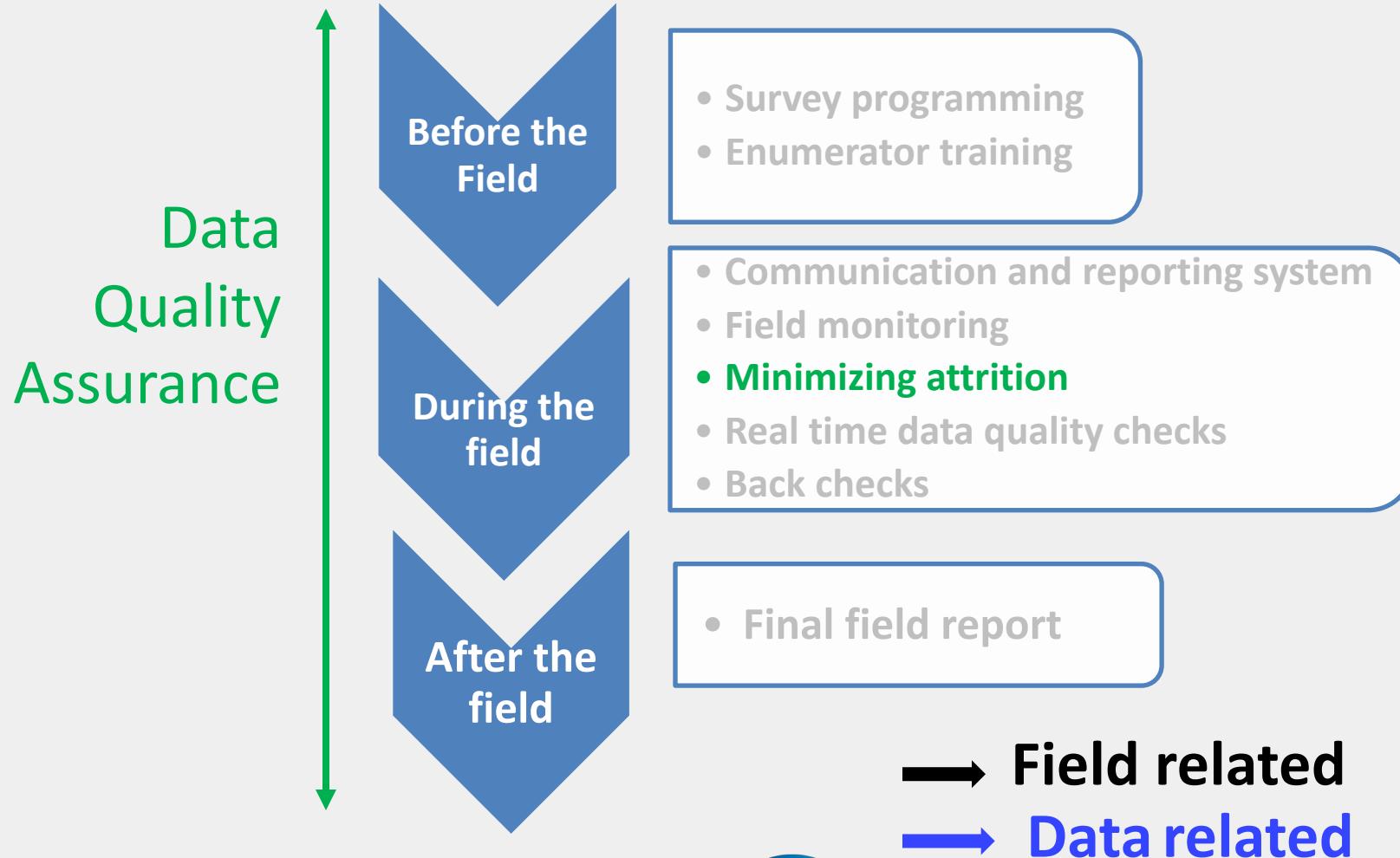
Field Monitoring: *Reviewing Forms*

How to review a form?

- Can be done by Supervisor, Data Quality Checker, or even other enumerator.
- ‘*Go to prompt*’ lists all questions of SurveyCTO form on a tablet
- Not all questions! Key ones that determine repeat group count – HH size, number of plots, number of crops, etc.
- DO NOT swipe through whole survey page-by-page, it can accidentally change responses!
- Often key questions are buried at the bottom of numerous layers of repeat groups (E.g. crop sales for a plot in a season). Confusing structure and not easy for supervisors to locate – also not recommended to dig for these.



Minimizing attrition



Data quality: minimizing attrition



- Tracking respondents in a follow-up survey
 - Data analysis is often done on the sample with both baseline and follow-up data
 - Attrition over 5% will be frowned upon by peer reviewers
 - Different attrition rates between treatment and control group will bias the follow-up sample
- Reasons for attrition - the respondent:
 1. moved away from the community
 2. cannot be found
 3. refuses to participate
 4. was interviewed with the wrong ID



Data quality: minimizing attrition

Problem	Solution
Moved away from the community	<ul style="list-style-type: none">Have a mop-up plan, including the budget for it
Cannot be found	<ul style="list-style-type: none">Record as much identifying information as possible during baseline: address, phone number, nickname, contact details of others to contactUse GPS coordinates to find their baseline locationAsk nearby study participants and neighbours if they know the person

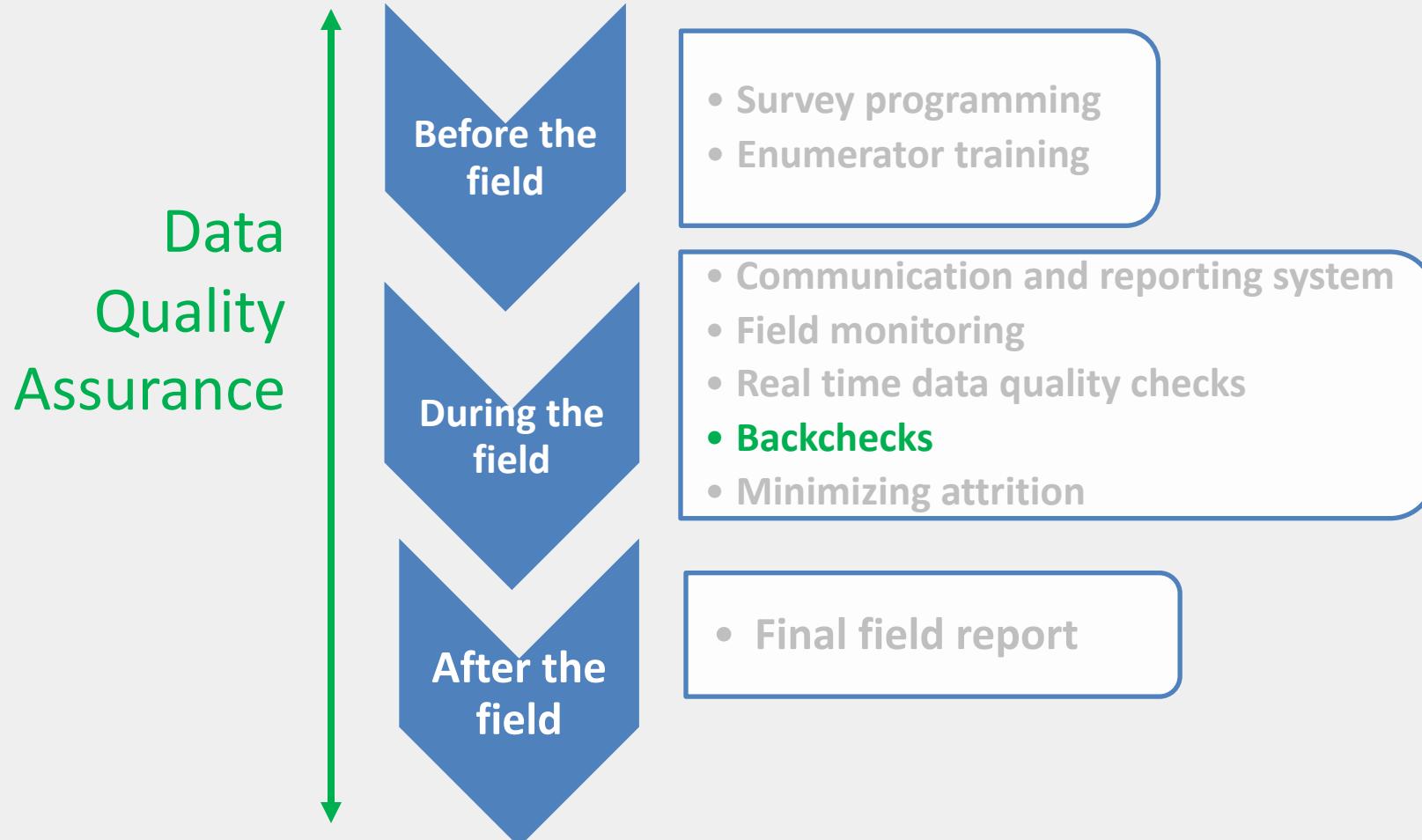


Data quality: minimizing attrition

Problem	Solution
Refuses to participate	<ul style="list-style-type: none">• Consent form to put respondent at ease• Cooperation techniques• Adapt to the respondent's schedule• Respondent gift
Was interviewed with the wrong ID	<p>Ensuring consistency of IDs with the baseline Mismatched IDs are a real headache!</p> <p>Preload all IDs → Enter ID → Pop-up message: You are about to interview Nancy Mwansa. Is this correct? <i>If this isn't your respondent's name, go back and correct the ID.</i></p> <p><input type="checkbox"/> Yes <input type="checkbox"/> No</p>



Backchecks





Backchecks

Definition: An experienced enumerator re-visits shortly after the interview and asks selected questions again (10-15 min)

- The answers are compared with the original survey
- Every team and every surveyor must be backchecked as soon as possible, and regularly. Frontload the backcheck sample.
- **10-20% of sample, with 20% being administered in the first 2 weeks of field work.**
 - Random sub-sample – select it in advance if data is available
 - Include **missing respondents** to verify that your team is not biasing your sample by not tracking hard to find respondents.
 - Observations flagged in other quality tests
 - Surveys of enumerators suspected of cheating



Backcheck Survey

Identifying Respondent and Interview Information

- Identifying information - make sure it's the right person
- Were you interviewed? On which day, what time of day?
- Was the enumerator friendly, annoying, rude, or (s)he was just ok?
- Who was present during the back-check interview
- If the back-check cannot be done, reasons why



Categories of Backcheck Variables

Type 1

Straightforward questions where we expect very little variation.

e.g. age, education, rel. to household head, floor type

Type 2

Questions where we expect capable enumerators to get the true answer.

e.g. sensitive Qs., Qs that involve calculations

Type 3

Are we asking the question in the right way? Are respondents changing their answers because they don't know?

e.g. It depends on your instrument. Type 3 Qs are about the survey, not enumerator performance

Introduction for the Backcheck Survey



- The **conduct of the backchecker** will influence the respondent's opinion of the study (and their eventual participation in the follow-up).
- No gift required.
- Write a **introduction script** for back-checkers to memorize:
 - Following up on a previous visit to check how was the interaction with the surveyor and to confirm we recorded their answers well
 - A short consent form - confidentiality & voluntary participation
 - That we are just asking a few questions and will not take more than 10-15 minutes of their time



Back Checks: Error Source

Is a discrepancy because:

1. the respondent understood the question both times but decided to change their answer
2. the respondent hadn't understood the question initially
3. the respondent didn't understand the question during the back-check
4. the enumerator recorded the response incorrectly accidentally
5. the enumerator recorded the response incorrectly maliciously (to save time by skipping sections, repeat counts, etc.)

→ Third visit or call in case of high rates of discrepancy.

Back Checks: Dealing with Errors



Analysis framework

- Set up at the beginning to track deviations early, e.g. “enumerator X always gets his question B wrong” or “question A constantly change across the entire team”.
- Decide on error thresholds (is a discrepancy of 5% acceptable? 10%?) and what you corrective measures you will take.
- As field work matures, these decisions can be modified if necessary. The goal is to: (1) record discrepancies, (2) maintain a log of errors and (3) set realistic expectations with clear actionable steps.
- *bcstats* = J-PAL/IPA Stata package that automatically compares back-check & main survey data and provides an outline for analysis – Lab 4 Track 1

Back Checks: Dealing with Errors



Type 1

Overall error rate >10% → ⚠️ that there may be systemic problems in the questionnaire or administration.

Calculate error rates by team, surveyor and question.

Type 2

Error rates > 10% should start conversations in the management team. Possible actions: re-training, meetings to review protocols and editing the questionnaire.

Calculate error rates by team, enumerator and question.

Type 3

The error rates by question should be examined. High error rates should be discussed with PIs. Not good practice to change question wording or structure halfway through the survey!

Back Checks Example



Type 1 Q should never ever change, regardless of interviewer, location or time of day.

Examples of these questions include:

- *age at last birthday*
- *whether currently in school*
- *highest level of education*
- *relationship to household head*

Back Checks Example



Team error rates

hhidteam	error_rate	differences	total
1	0.1136	85	748
4	0.0682	57	836
3	0.0527	40	759
2	0.0391	34	869



Back Checks Example

Enumerators with error rates above 10%

enum_id	error_rate	differences	total
314	0.2121	7	33
324	0.1818	8	44
224	0.1688	13	77
122	0.1636	9	55
113	0.1591	14	88
121	0.1558	12	77
223	0.1558	12	77
115	0.1515	15	99
425	0.1414	14	99
123	0.1182	13	110

Back Checks Example



Questions with error rates above 10%

variable	error_rate	differences	total
<code>mod1rel_h_head</code>	0.1336	39	292
<code>mod1highest_school_level</code>	0.1062	31	292

`mod1rel_h_head` - *How is the head of this household related to you?*

Data quality field checks: one more action you can take

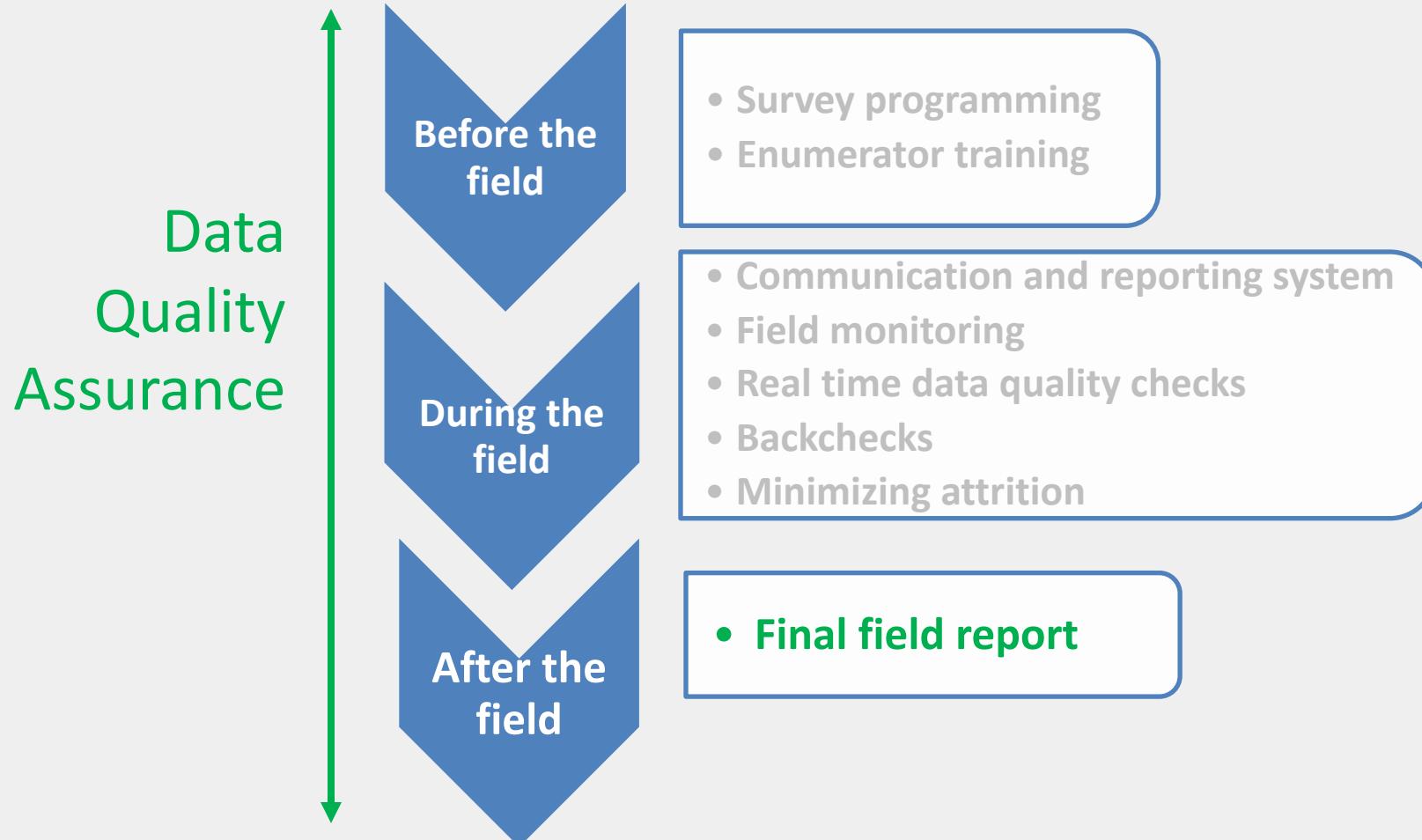


Refresher training

- **When?**
 - 2 weeks into fieldwork
 - Can be planned ahead of time and included in the survey firm TOR
- **Goal?**
 - Share experiences from the field
 - Discuss questions that are not being understood
 - Review errors uncovered by back-checks & real time data quality checks
 - Discuss strategies to avoid errors
 - Re-train on specific areas of the survey
 - Discuss challenges and strategies to locate respondents



Final field report





Final Field Report

- Survey firm usually provides a final field report at the end of the data collection
- Can be used to contribute to data quality in the last stage of the process, including qualitative data:
 - Provide information about everything that could not been captured by the survey instrument:
 - Respondents never understood a question and answered randomly
 - Limited option choices for specific questions
 - Any enumerator feedback regarding respondents feeling and understanding of questionnaire
 - Share information about communities size and structure of communities (useful to calculate sample weight!)

Thank you!

Contact us!

arigaud@worldbank.org
lcostica@worldbank.org
sglover1@worldbank.org