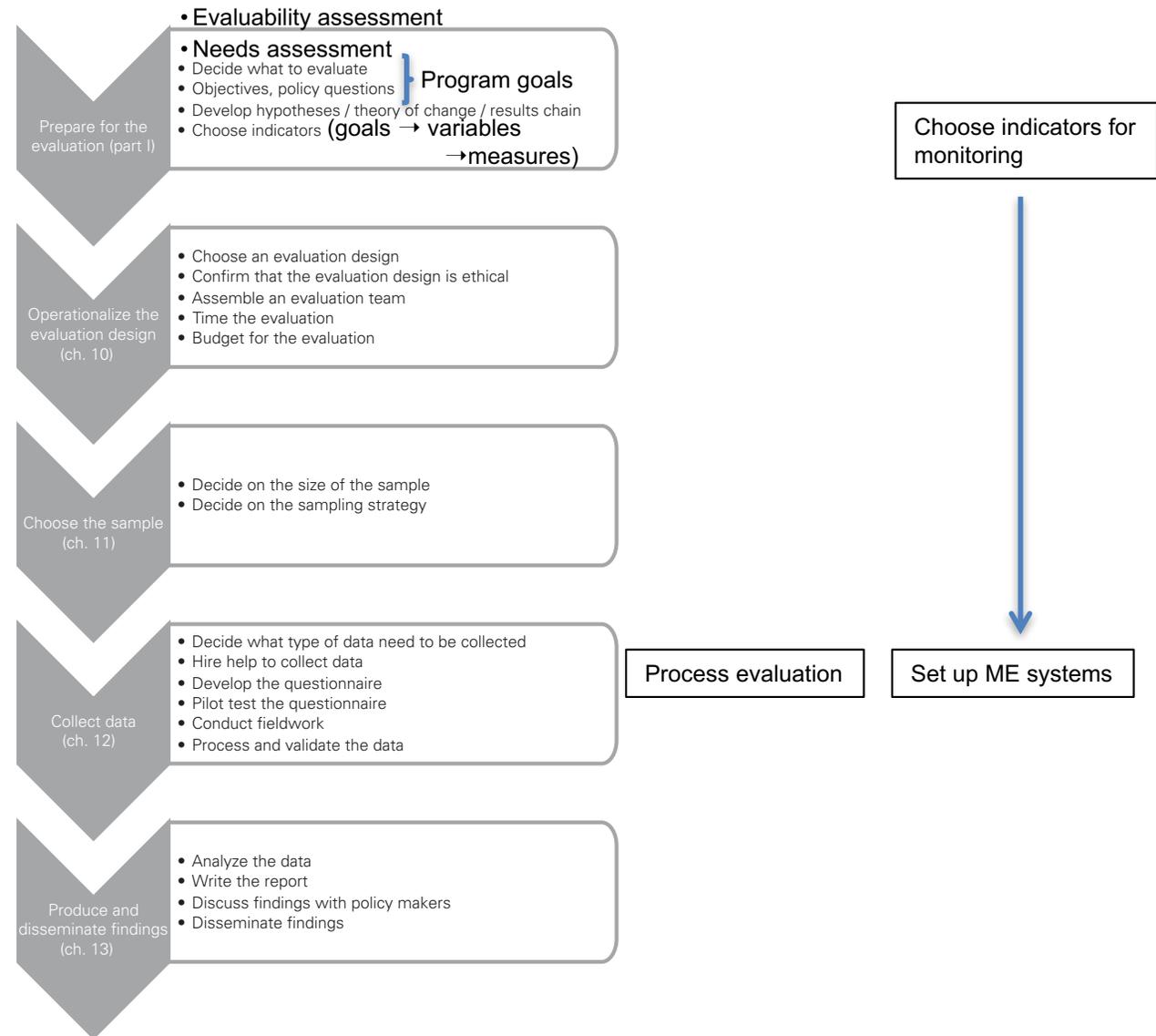


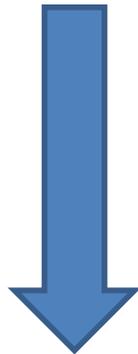
# Data Collection for Research

# Overall framework for program evaluation / research



# Impact Evaluation Project Cycle

**Design the intervention**



**Rollout the intervention**



**Ongoing Monitoring and  
Process Evaluation**

**Design the impact  
evaluation**



**Collect Baseline Data**



**Collect Follow-up Data**



**Analysis**

Scope of this  
lecture



# It's all about timing

TIMING	MONEY →	Excess demand for program (limited resources)		No excess demand for program (fully resourced)	
	TARGETING RULES →	Continuous targeting or ranking & cutoff (1)	No continuous targeting or ranking & cutoff (2)	Continuous targeting or ranking & cutoff (3)	No continuous targeting or ranking & cutoff (4)
	Phased implementation over time (A)	<b>CELL A1</b> (3.1) Randomized assignment (4) RDD	<b>CELL A2</b> (3.1) Randomized assignment (3.2) Randomized promotion (5) DD with (6) Matching	<b>CELL A3</b> (3.1) Randomized assignment to phases (4) RDD	<b>CELL A4</b> (3.1) Randomized assignment to phases (3.2) Randomized promotion to early take-up (5) DD with (6) Matching
Immediate implementation (B)	<b>CELL B1</b> (3.1) Randomized assignment (4) RDD	<b>CELL B2</b> (3.1) Randomized assignment (3.2) Randomized promotion (5) DD with (6) Matching	<b>CELL B3</b> (4) RDD	<b>CELL B4</b> If less than full takeup: (3.2) Randomized promotion (5) DD with (6) Matching	

# Assumptions for this discussion

1. We're planning a prospective research project.
2. We need baseline and follow-up survey data on treatment and control group(s) to measure program impact\*
3. We are going to collect our own data.

\* Process/implementation data should also be gathered throughout the intervention

# Before collecting your own data

- Can we use existing data?
  - Regular surveys (census, cross-sections, panels)
  - Regular monitoring (annual achievement tests)
  - Administrative records (health records, school enrollment)
- In many settings, administrative data is insufficient, poor quality and low coverage
  - Ex. Turkey vs. Argentina

# Before collecting your own data

- Who should collect?
  - Bureau of Statistics: has capacity & good place to invest in further capacity
  - University: maybe cheaper, often less infrastructure and thus more monitoring
  - External firm: depends on capacity and experience
- When do we need to start?
  - Procurement, training and data collection all take time

# Objectives of data quality control

To collect data that:

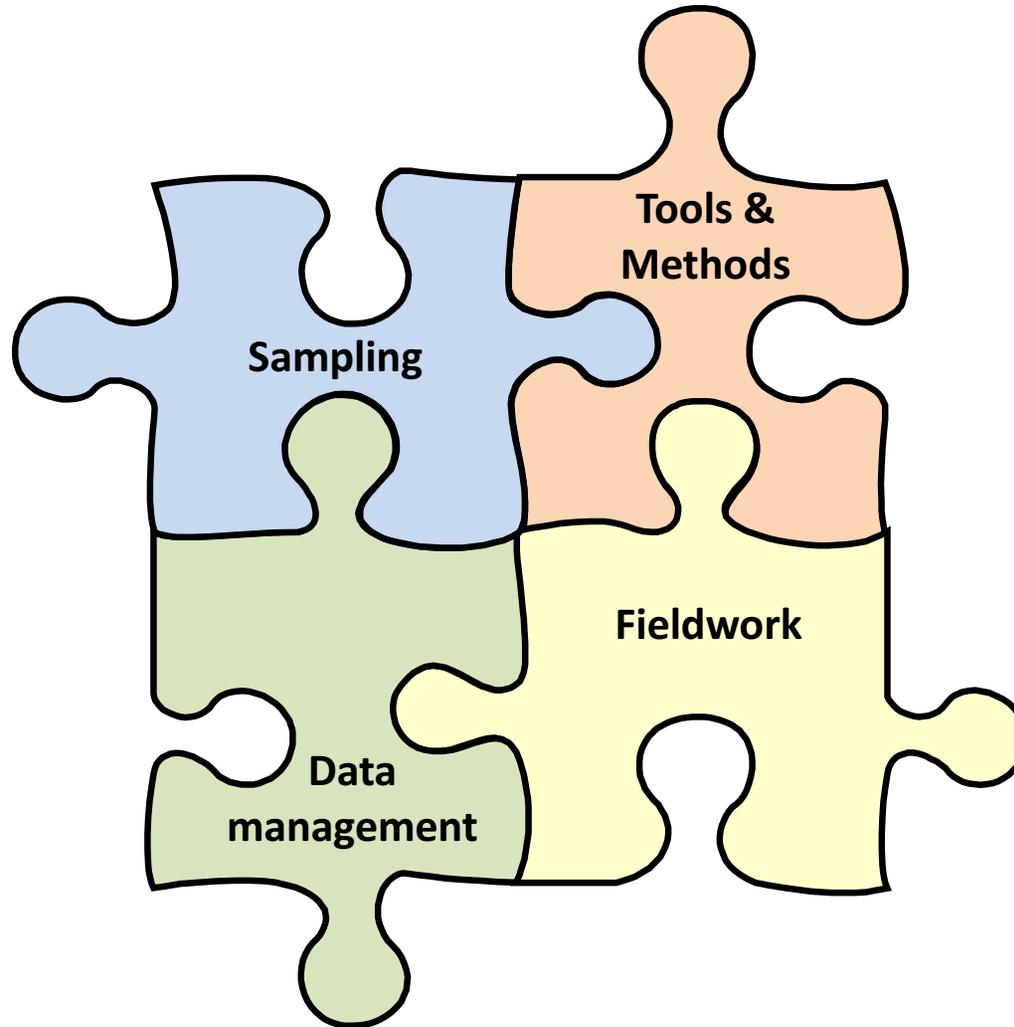
1. accurately reflects the reality of the population
2. is representative of the entire target population
3. allows policy makers and analysts to make real-time, informed decisions

 Need to control sampling and non-sampling error!

# What are sampling and non-sampling errors?

- **Sampling error**: the result of observing a sample of  $n$  households (the sample size) rather than all  $N$  households in the target population
- **Non-sampling error**: the result of errors in survey development and execution. Some examples are:
  - **measurement error** -when the answers written on the questionnaires are different from the actual values
  - **selection bias** -results from imperfections in the sample frame or deficiencies in the sample selection process
  - **non-response** -when we don't get an answer to some or all of our questions from certain households

# The Total Survey Design (TSD) approach



# Data collection plan

Includes:

1. Sampling
2. Tools and methods
  - questionnaires and instruments
3. Fieldwork
4. Data Management
5. Budget
6. Timeline of activities

# 1. Sampling

- Develop sample design – sample size and geographic distribution
- Get an updated sample frame and corresponding cartography
  - Local statistical office, Ministry of Health
- Do we have to generate a listing?
  - before the survey or in parallel?
- Compute sampling weights and sampling errors as necessary

## 2. Tools and methods: questionnaires

- What do you want to know?
  - Tailor your survey to capture outcomes of interest
- Use reliable and valid instruments
  - Be careful: what's reliable and valid in one cultural and linguistic context may not be so in another
  - Adapt it to the country specific reality and local language(s)
- Save time, money and pain: test your questionnaire!

# Best practices for questionnaires

- Define your topics and concepts to avoid confusion
- Question order matters
- Keep it short and make it user-friendly
- Phrase questions clearly
- Use established techniques to minimize respondent mistakes (e.g. calendars for event histories)

# Best practices for questions 2

- A good question is understood consistently by all respondents
- A good question is administered consistently to all respondents
- A good question elicits the kind of answers the researcher wants:
- **Bad question:** When did you move to Dar es Salaam?
  - A: In 1964
  - A: When I was 20 years old
  - A: After I finished college
- **Better question:** In what *year* did you move to Dar es Salaam?

# Best practices for questions 3

- A good question is one where the respondents have the necessary knowledge to answer
  - asking a good question of the wrong person is a source of error in your data
- A good question is one where the respondent is willing to provide the “true” answer
  - difficult for sensitive questions

# Best practices for questions 4

- Ask about firsthand experience
- Ask one question at a time
  - **Bad question:** Are you physically able to do things like walk or carry a full water bucket without difficulty?
  - **Better question:** Are you physically able to carry a full water bucket without difficulty?

# Who are your respondents?

- Different respondents require different techniques, e.g. youth are a tough crowd:
  - Often mobile
  - Not always well informed
  - Require special consent procedures
  - Cagery about socially undesirable behavior
  - Can have low literacy levels

# Best practices for sensitive questions

- Use open questions for frequencies of undesirable behavior
- Design long questions but short instrument
- Use familiar words (know local terms)
- Ask “have you ever done x” before asking “are you currently doing x” for socially undesirable behavior
- Embed threatening questions in a list of more or less threatening topics

# List randomization

- Two groups are presented with a list and ask how many apply. One group randomly assigned with an additional sensitive question.

Clients were randomly selected to be presented with one of four possible groups of three to six statements<sup>3</sup>. All clients received the following three statements: “I used part of my Arariwa loan to buy *merchandise* for my economic activity”, “I used part of my Arariwa loan to buy *equipment* for my economic activity” and “I *shared* my loan with another person”. Clients in group A (n = 408) only received these statements. Clients in group B (n = 414) additionally received the following statement: “I used at least a quarter of my Arariwa loan on *household items*, such as food, a TV, a radio, etc.” Group C (n = 388) received the four previous statements, and the statement, “I used at least a quarter of my Arariwa loan to pay for my family's *medical expenses*.” Group D (n = 401) received the previous five statements and the statement, “I used at least a quarter of my Arariwa loan to pay for my family's *educational expenses*.” By subtracting the mean number of true statements for group A from the mean number of true statements for group B, we get the proportion of clients that used a quarter of their loan for household items. We similarly subtract B from C and C from D to get the proportions of clients using their loans on education or medical expenses.

# List randomization

**Table 2**

Comparison of direct report and list randomization estimates from ICT financial literacy project in Peru.

	Loan Use:	Household Items (1)	Health (2)	Education (3)
Direct report				
Proportion reporting this use		0.077	0.022	0.072
	SE	(0.007)	(0.004)	(0.006)
	N	1650	1650	1650
List randomization				
Mean of "Yes" responses for short list		1.213	1.527	1.758
	SE	(0.031)	(0.038)	(0.049)
	N	408	414	388
Mean of "Yes" responses for long list		1.527	1.758	2.090
	SE	(0.038)	(0.049)	(0.055)
	N	414	388	401
Difference (proportion reporting this use)		0.313	0.231	0.332
SE of difference		(0.049)	(0.062)	(0.074)
p-value from ttest		0.000	0.000	0.000
	N	822	802	789
Comparison of direct report and list randomization				
List randomization minus direct report		0.236***	0.209***	0.261***
Z-test statistic for difference in proportions		4.752	3.386	3.512

# How will you capture the data?

- Self Administered Questionnaire (SAQ)
- Telephone (CATI,RDD)
- Paper and pencil Interview (PAPI)
  - Data entry
  - Scannable forms
- Computer Assisted Personal Interview (CAPI)
- Audio Computer Assisted Self Interview (ACASI)

# Questionnaire development

- The survey researchers' axiom:
  - Everyone thinks they can design a questionnaire
- Include an experienced survey research organization in the development of your questionnaire
- Begin with examples of successful questionnaires fielded in similar settings

# Questionnaire testing

- Evaluate your questions
  - Conduct focus groups
  - Conduct cognitive testing
- Evaluate your questionnaire and procedures
  - Conduct full field pretest
- Revise questionnaire and procedures

# 3. Fieldwork: staffing and training

- Interviewer staffing and training is directly related to data quality
- Recruit supervisors, interviewers and data entry operators very carefully
- All materials must be finalized before training begins:
  - Training materials
  - Field manuals
  - Questionnaires
  - Data-entry program

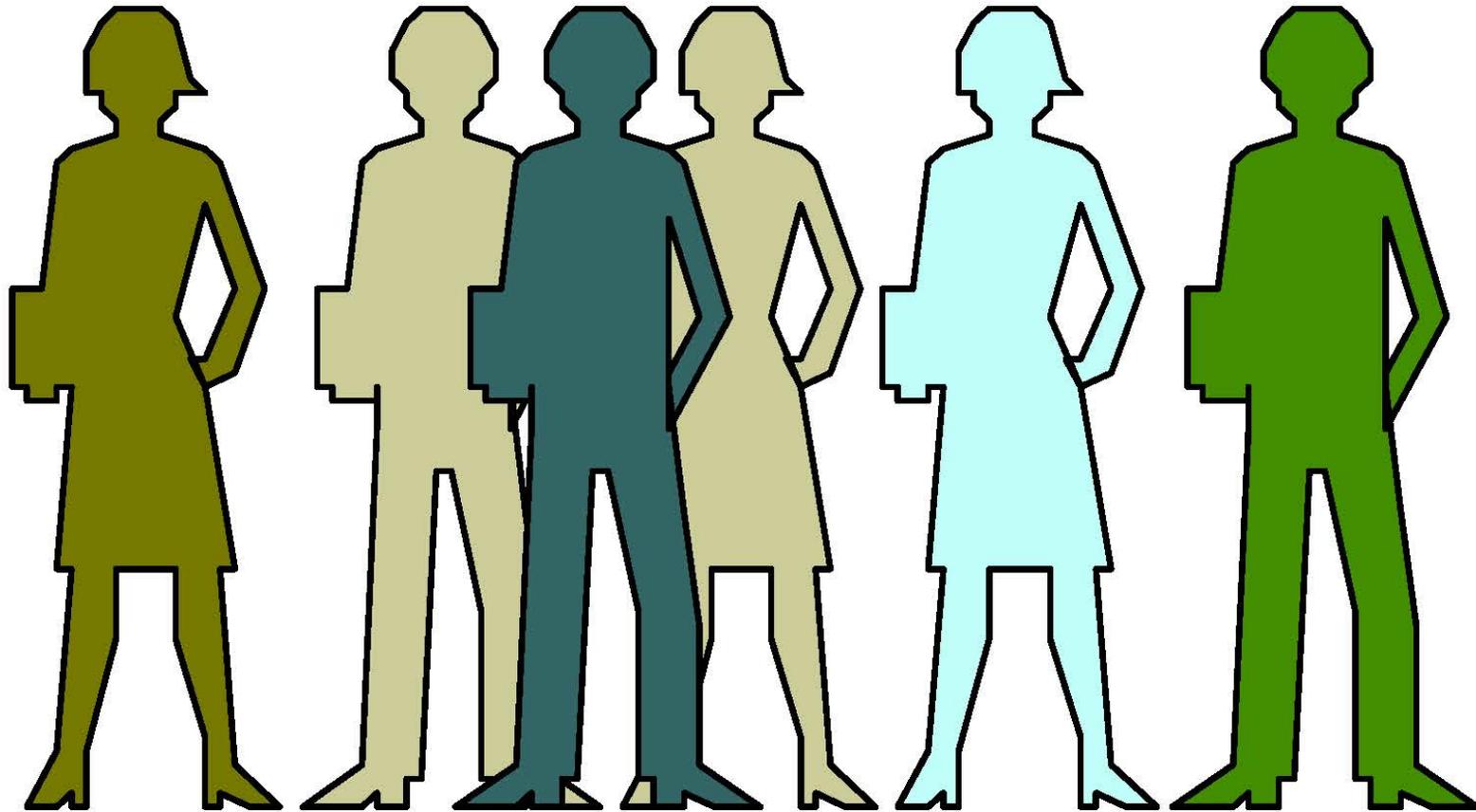
# Train interviewers thoroughly

- Training should have four main elements
  1. plenary sessions: introduction and lectures
  2. practical work in small groups:
    - simulated interviews, role playing, interpreting inconsistencies, etc.
  3. practice fieldwork
  4. candidate assessment
- High-quality training takes time
  - Plan to spend roughly 2-4 weeks depending on questionnaire complexity
  - This is often underestimated

# Field management for quality data

- Maximize response rates:
  - Schedule visits sensibly
  - Follow up with non-respondents
  - Consider incentives
  - Gather locating data for follow up
- Employ Computer Assisted Field Entry (CAFÉ)

# Composition of a “field team”



**Supervisor**

**Interviewers**

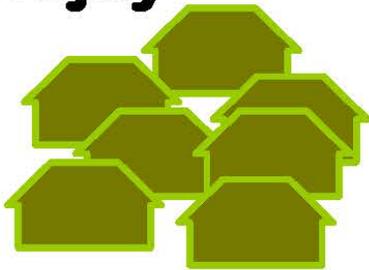
**Anthropo  
-metrist**

**Data entry  
operator**

# Field management

*Mobile teams with integrated data entry*

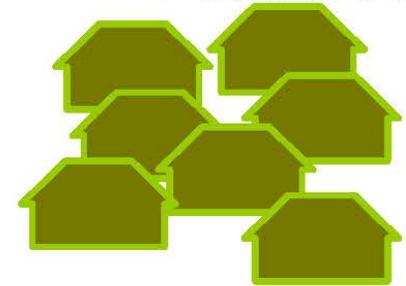
**Jujuy**



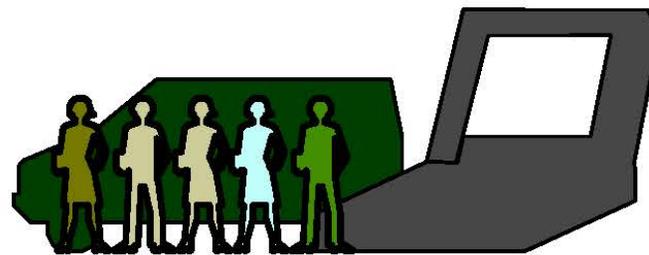
**Salta**



**Entre Rios**



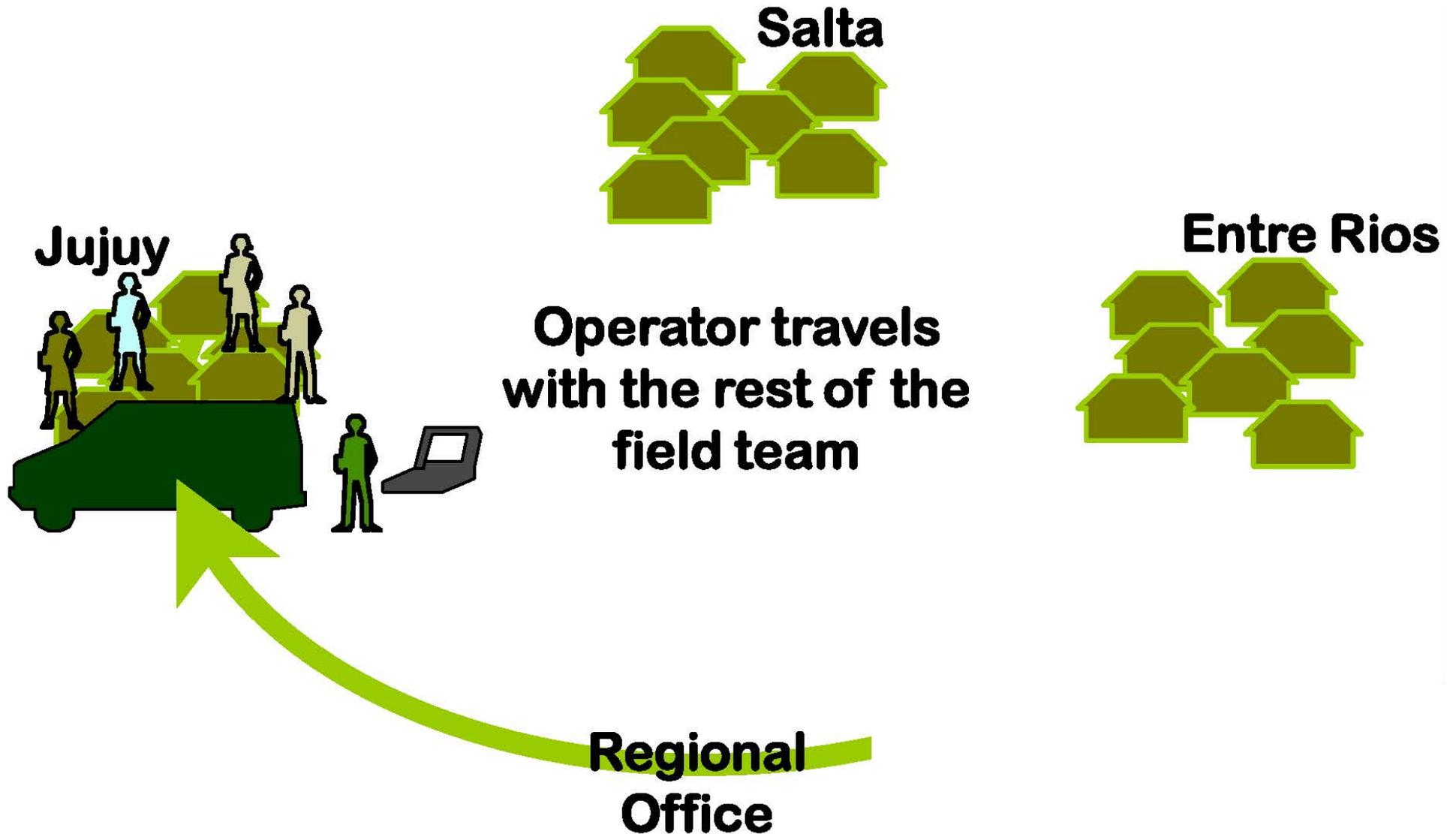
**Team works with  
portable computers  
and printers**



**Regional  
Office**

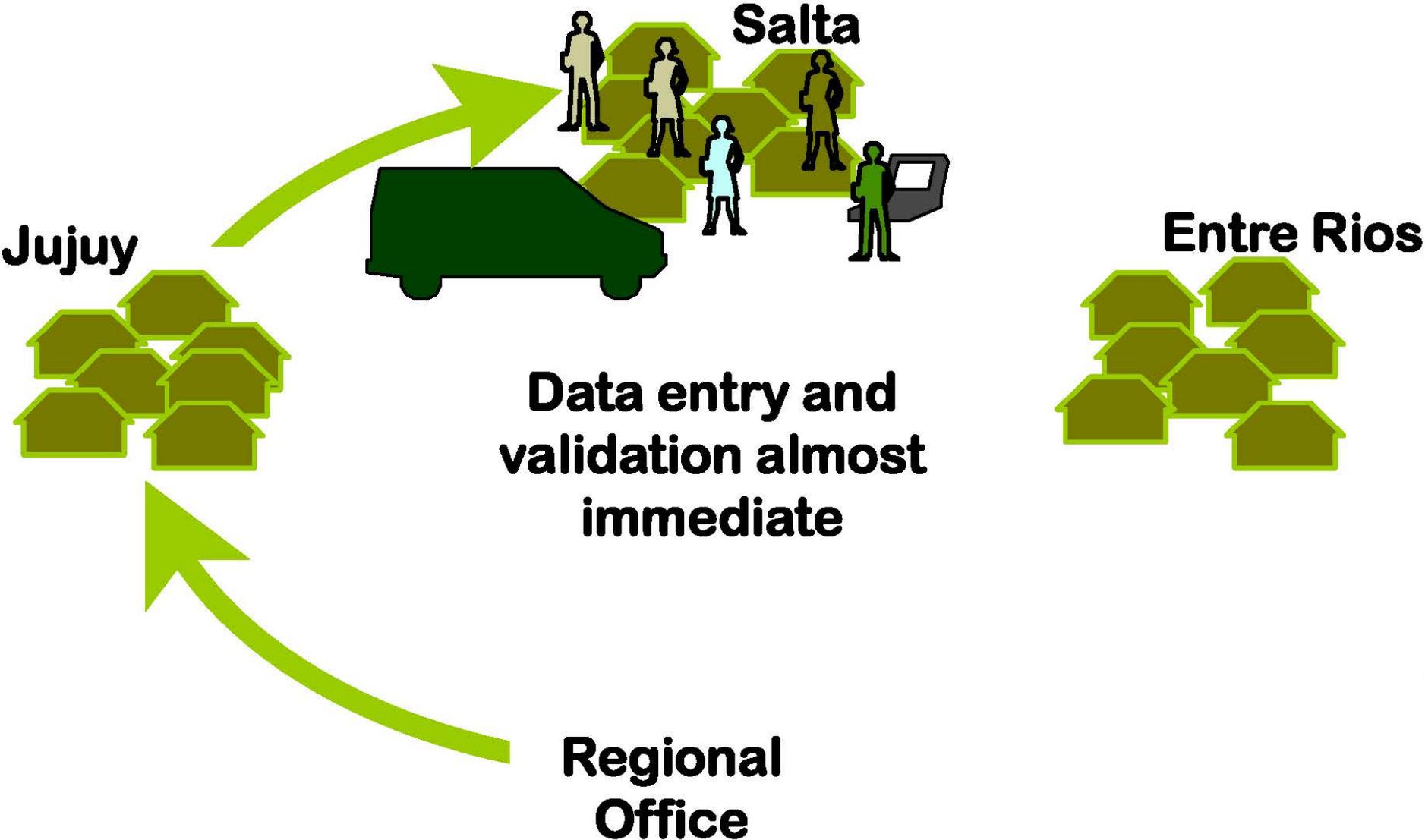
# Field management

*Mobile teams with integrated data entry*



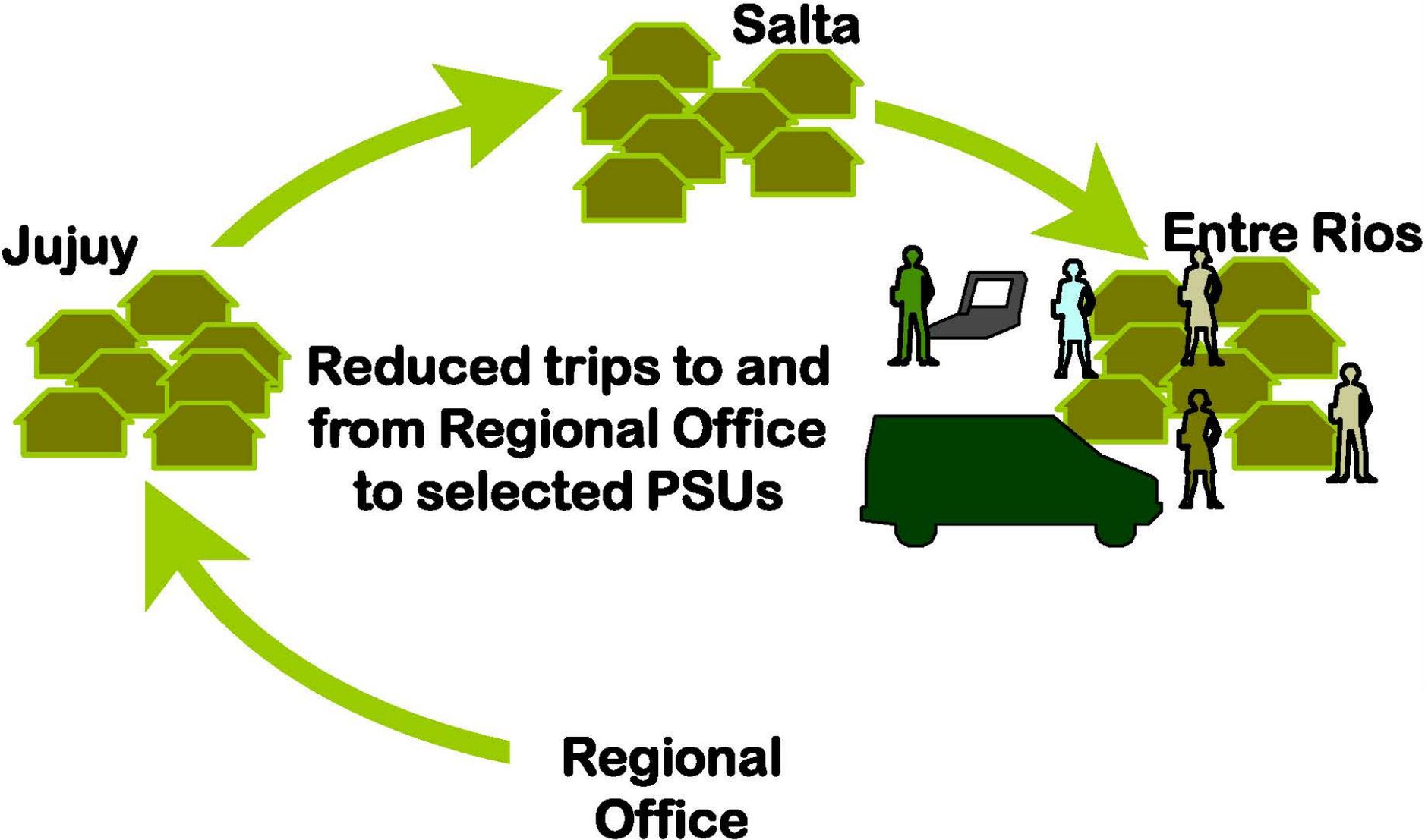
# Field management

*Mobile teams with integrated data entry*



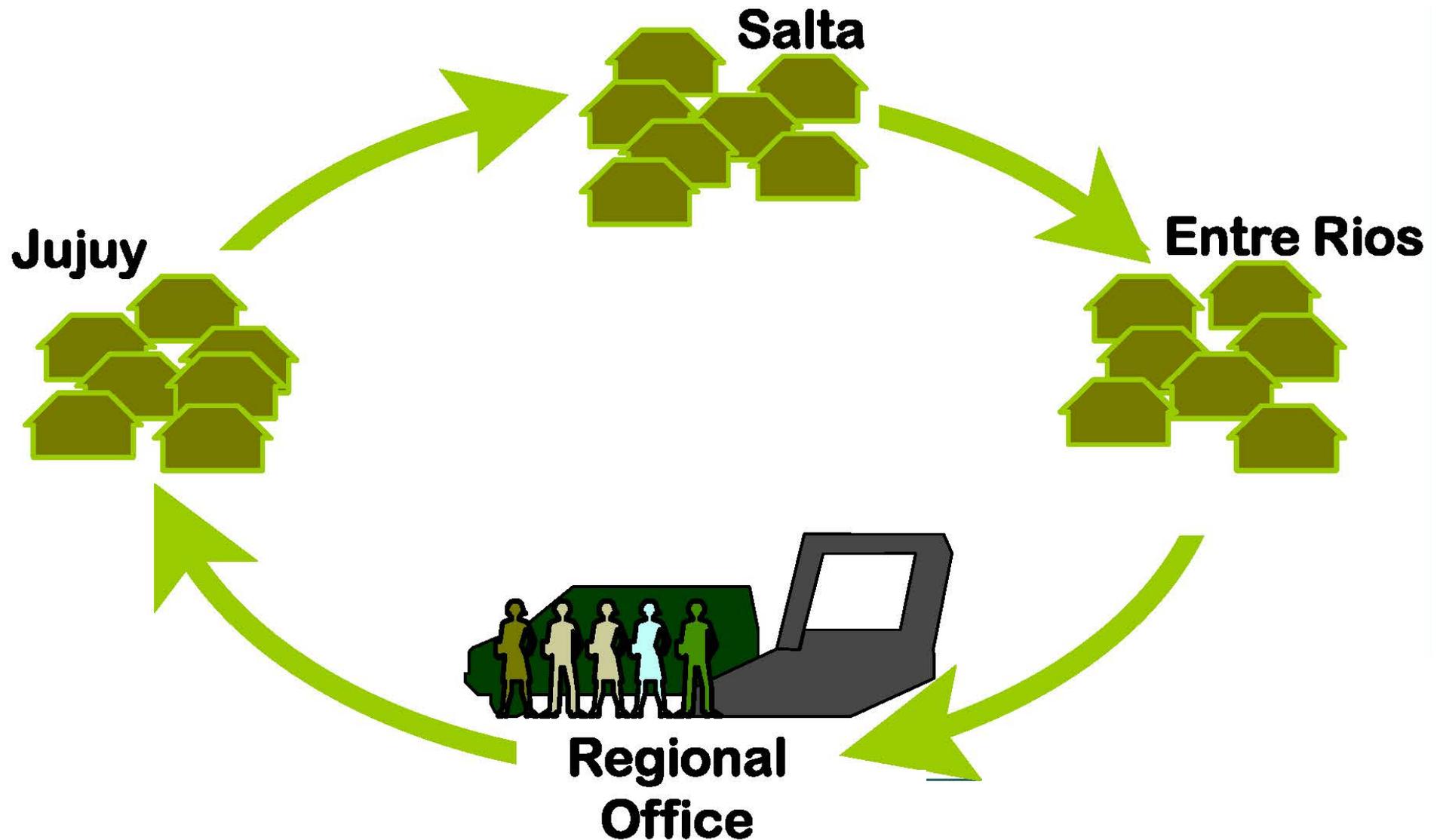
# Field management

*Mobile teams with integrated data entry*



# Field management

*Mobile teams with integrated data entry*



# Computer-assisted field entry (CAFÉ)

- provides immediate feedback on the performance of the field staff
- allows early detection of any inappropriate behavior
- enables inconsistencies to be corrected at their source by re-visiting the HH, rather than ex-post “data cleaning” based on assumptions

# 4. Data management

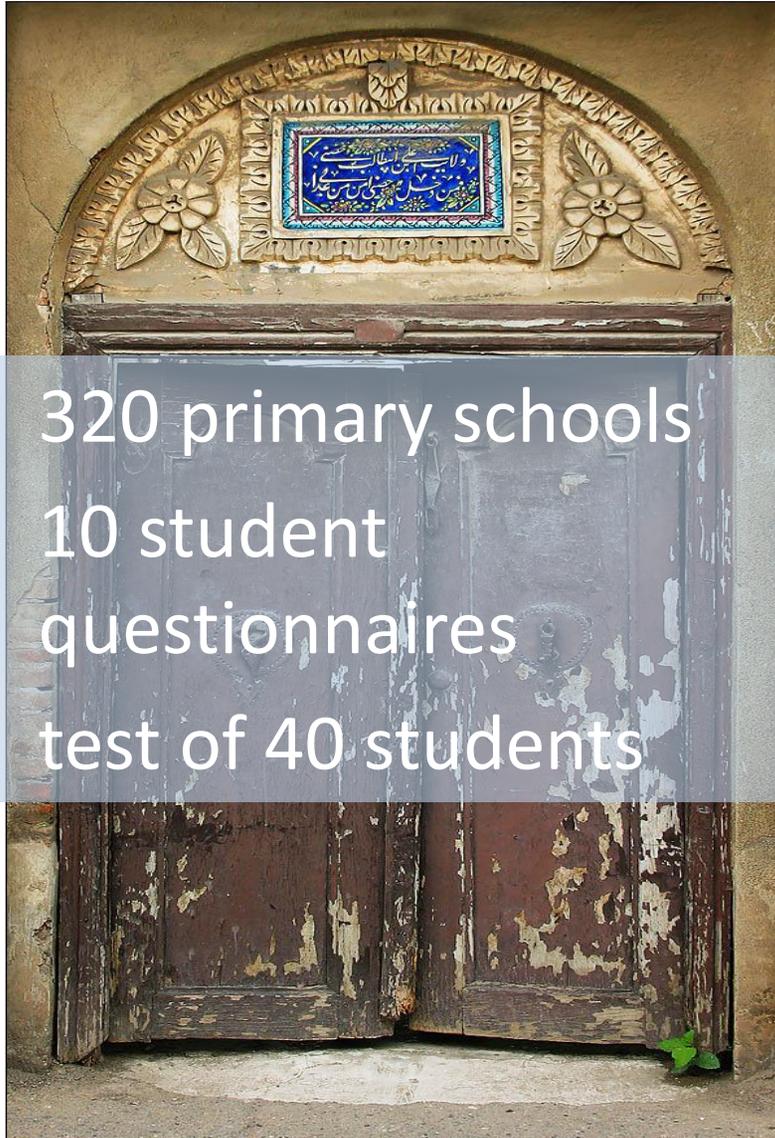
- Develop a data entry program that includes quality control checks
- Include data management procedures that emphasize confidentiality
  - Remove unique identifiers when transmitting and storing the data
- Monitor your sample carefully using frequent supervision reports

# 5. Budget

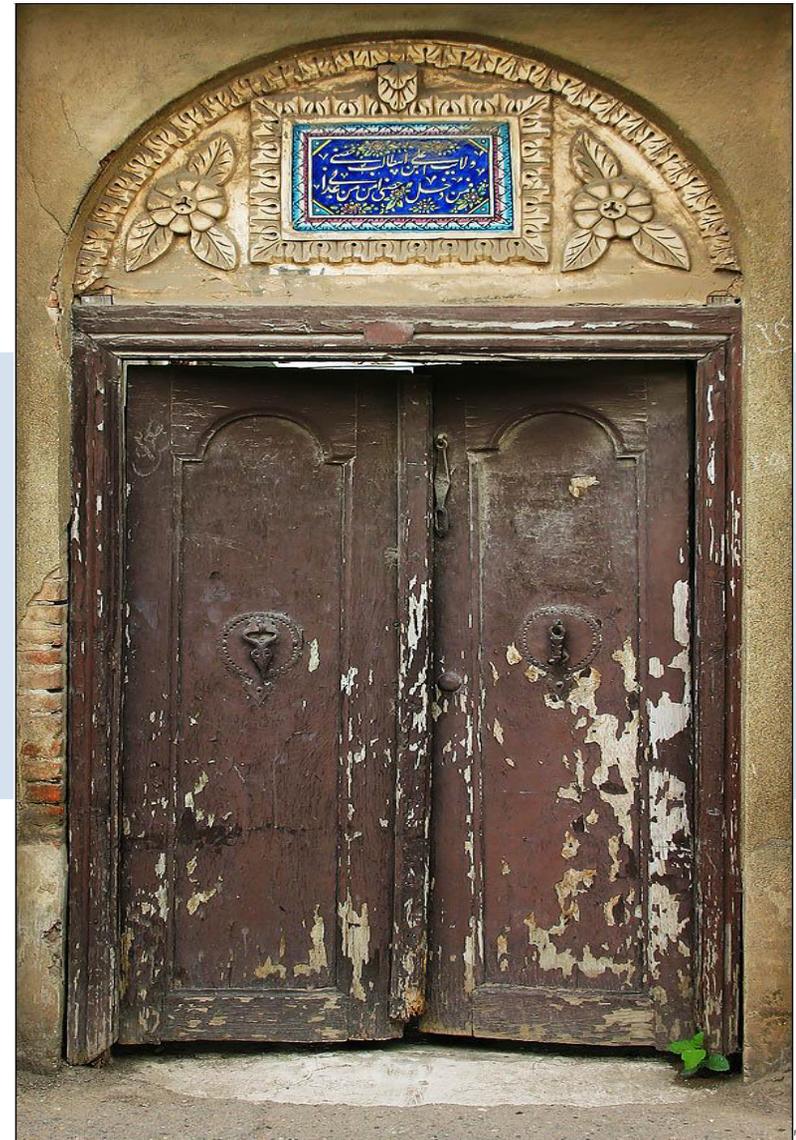
- Survey data collection can be expensive
  - But it isn't as expensive as carrying out programs that are not effective.
- The long term savings of conducting a thorough impact evaluation with quality data collection can be immeasurable in social terms
- High quality data is vital to conducting an effective impact evaluation

# How much is this going to cost?

[hint: quite a bit]

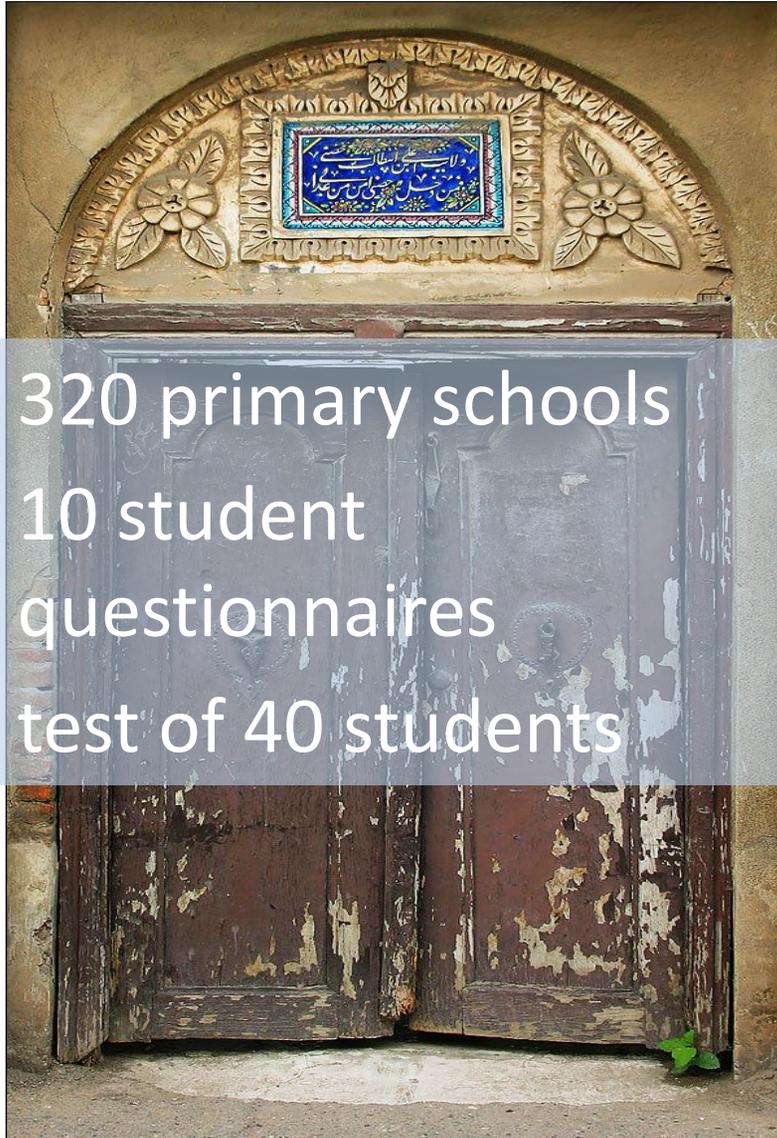


- 320 primary schools
- 10 student questionnaires
- test of 40 students



# How much is this going to cost?

[hint: quite a bit]



- 320 primary schools
- 10 student questionnaires
- test of 40 students

USD

\$200,000

# How much is this going to cost?

[hint: quite a bit]

USD

\$50,000



- 1,000 primary schools
- head teacher questionnaire
- teacher questionnaire

# Large variations in cost

## 320 schools

- USD \$200,000

## 1,000 schools

- USD \$50,000

## Big differences!

- Scale of survey
- Size of country
- Labor costs
- Difficulty in getting to schools

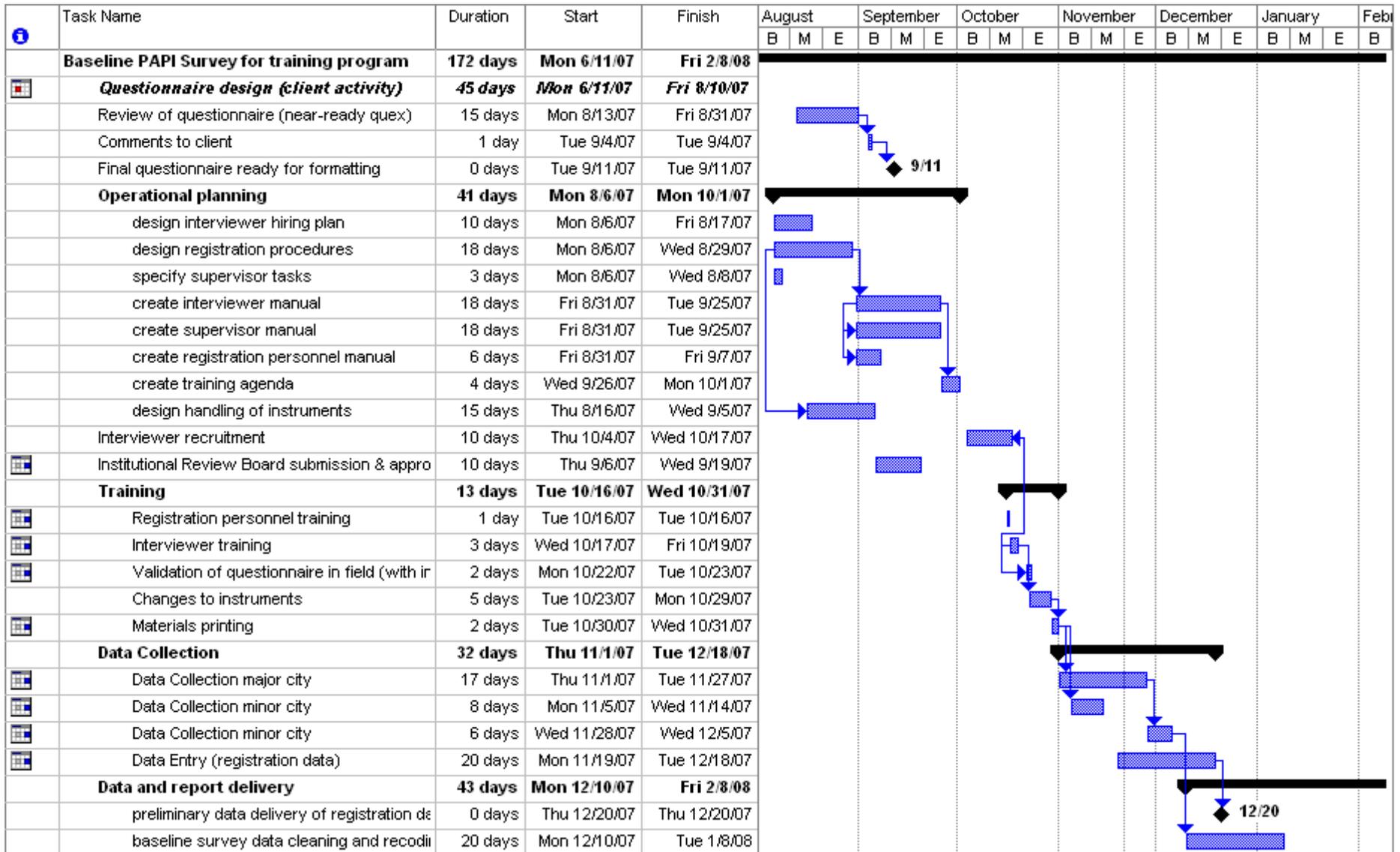
Household surveys: More expensive

## 6. Timeline: quality data collection takes time

Activity	Timing Needed
1. Procurement	2-4 months
2. Development of Sample	Depends on sample design and availability of sample frame
3. Adapt and test questionnaire to the local conditions	4-6 weeks
4. Staffing and training	2-4 weeks for recruitment and selection, 2-4 weeks for training
5. Data management	4-6 weeks
6. Data Collection	2+ months

	M 1	M 2	M 3	M 4	M 5	M 6	M 7	M 8	M 9	M 10	M 11	M 12	M 13	M 14
1. Management, Financing, Logistics and Institutional Agreements	█	█												
2. Sample			█											
3. Questionnaire design			█	█										
4. Data management				█	█									
5. Staffing and training						█	█							
6. Survey in the field. 16 weeks?							█	█	█	█	█			
7. Database preparation												█		
8. Analysis. 2 months?												█	█	

# Sample project plan: 30 minute PAPI



# Quality data collection: summary

1. Pretest data collection plan and questionnaires, leaving time for adjustments
2. Select interviewers carefully, train them thoroughly and evaluate them dutifully
3. Use “Field Team” approach and CAFÉ when possible

## Quality data collection: summary

4. Data collection is expensive: choose your mode and sample sizes wisely
5. Build a realistic timeline with sufficient time for questionnaire development, testing, training and adjustments