

# ***Impact Evaluation Technical Course 2013: How to Design, Manage, and Conduct Impact Evaluations***

**CLEAR-AFDC  
April 9-19, 2013  
Shanghai, China**

**Session 19  
Data Collection and Data Management**

## I. Introduction

Many, if not most, impact evaluations collect new data in order to compare individuals who have been “treated” and individuals who are “controls”. **Data collection and data management are very important tasks that should not be assigned to a group that has no particular interest in the evaluation.** This session provides recommendations on data collection and management.

There are two general goals for data collection, which may at times conflict:

1. **Timeliness:** The data must be collected and analyzed relatively quickly, since policymakers are waiting for the results.
2. **Quality:** The data must be accurate and useful for the planned evaluation.



The rest of the session gives specific recommendations, in chronological order.

## II. The Steps Involved in Data Collection and Data Management

The overall process of collecting and managing data can be divided into 5 steps:

- *Establish Procedures for Collecting the Data*
- *Collect the Data (Including Monitoring of Data Quality)*
- *Check Data Quality*
- *Create Data Files for Analysis and Dissemination*
- *Establish a System to Store, Revise and Disseminate the Data*

The rest of this session will discuss each of these steps in more detail. While most of the session focuses on the collection of new data, much of the advice also applies to evaluations that are based on existing data. Indeed, many impact evaluations have combined new data with existing data.

### III. Establish Procedures for Collecting and Managing the Data

The first task in data collection and management is to put together a plan that describes how the data will be collected and managed. This task can be divided into several distinct components, which will now be presented, roughly in chronological order.

#### Choose a Data Manager

In general, the first decision to be made to begin establishing a set of procedures for data management and collection is to **choose a data manager** for the impact evaluation. As will be explained in the *Survey Management* Session, the overall impact evaluation needs to establish a team to work on the evaluation, and one of the key positions is that of data manager.



In some situations, it may be useful to have two data co-managers, one being a high level official from the organization that is managing the research project, and the other an “outsider” with experience in collecting and managing data for the purposes of conducting an impact evaluation.

## Draw the Sample

Once the population has been chosen for the impact evaluation, a sample of households or other entities (e.g. schools or communities) must be drawn.



For randomized experiments, drawing the sample for data collection can be combined with random assignment of the treatment. For example, from a population of 1000 schools, 100 can be randomly drawn to be assigned to the program while another 100 can be randomly drawn to be the controls.

For non-experimental methods, a sample of households (or schools or communities) must also be drawn. If either the “treated” or the “control” households are relatively rare, it may be necessary to oversample that group to ensure that there are adequate numbers of people who are treated and untreated. When participation in treatment is a choice and the sampling is based on that choice, the sampling is called a “choice-based sample,” which is a special type of stratified sampling. For a discussion of stratified and choice-based sampling, see, e.g. Amemiya’s *Advanced Econometrics*.

## Link New Data Collection with any Existing Data that Will Be Used

In some impact evaluations it is very useful to link newly collected data with existing data that are already collected by the government or some other organization. For example, some countries (e.g. Brazil) collect a school census every year that includes useful information on all schools in the country.



The first step in linking existing data to new data is to **obtain permission to use the existing data**. In some cases, it is difficult to obtain such permission, so you should not assume that you can get it.

Perhaps the **main issue**, after permission to use the data has been granted, is **matching the existing data to the new data**. In many, but not all, cases, there is some kind of code that can be used to match. For example, there may be unique school codes in the existing data, and the same codes should be used in the newly collected data. In some countries each individual may have a unique national identity number (e.g. social security number in the U.S.) that can be used to link new to existing data, and so – **if possible** – it will be useful to collect the national identity number in the household questionnaire.

## Organize a System of Data Entry

In general, the last step before data collection is to train interviewers, focusing on the specifics of the new questionnaire. Almost always, during this final training some errors are found in the questionnaire, so it will have to be corrected and then the final version of the questionnaire needs to be printed in large quantities.



A general recommendation for **data entry** is that it **should be done as soon as possible after the interviews have been completed**. Indeed, if interviewers use laptop computers or smaller electronic devices to conduct the interviews, data entry is “automatically” done at the time of interview. Rapid data entry allows the data entry software to find errors **in time for interviewers to return to respondents to correct those errors**.

However, in many cases data are still collected using paper questionnaires, so data entry takes place after the paper questionnaires have been filled out. In such cases, a system should be set up that ensures that data entry will be done within at most a few days of the date of the interview.

With either type of questionnaire (paper or electronic), a data entry program is needed for storing the data electronically and, more importantly, to check the internal consistency of the data. Some specific points on this are:

- There are several software programs available. For a recent assessment of computer assisted personal interviewing (CAPI) software, see <http://siteresources.worldbank.org/INTSURAGRI/Resources/7420178-1294259038276/CAPI.Software.Assessment.Main.Report.pdf>
- The software should check for internal consistency and “beep” (or send some other signal) when inconsistencies are found. If this occurs during the interview, the interviewer can re-ask questions to resolve the inconsistency. If data are entered after the interview, ideally there is time to return to the household to resolve the inconsistency.
- Four kinds of data checks are: 1. range checks; 2. checks against reference tables (e.g. anthropometric data); 3. check that skip codes are correctly followed; and 4. mutually inconsistent data.
- The draft program should be checked during the field test and the training.



## Interviewer and Supervisor Training

Once the questionnaire has been finalized and all other preparations have been done, interviewers and supervisors should be trained in the use of the questionnaires and other survey instruments. **This is done immediately before the data collection is supposed to start**, although (as mentioned above) it is very likely that some errors will be found in the questionnaires and so some minor adjustments need to be made and new questionnaires need to be prepared (and the data entry program must be finalized) before the data collection can start.



Here is some general advice on interviewer and supervisor training:

- Interviewers must understand the informational goal of each question.
- All field team members must clearly understand the data collection goals.
- All field team members must clearly understand the intervention (if an experiment, they should know how the randomization was done, and what the control group can expect).

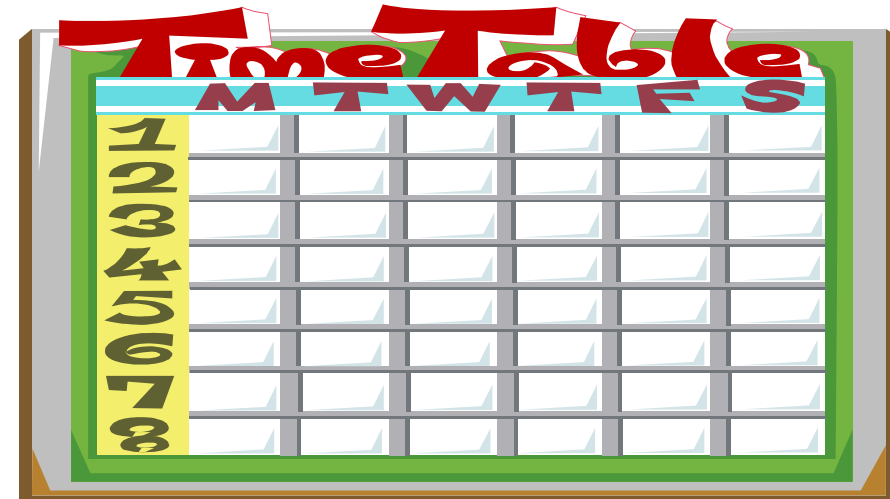
- If respondents are to receive financial incentives, it is VERY important that they be made clear to all involved. One option is to announce them publicly.
- Interviewers should be trained how to interact in a polite and respectful way with all respondents.
- To ensure cooperation by households, the data collection team should:
  - a) Train all team members to be polite and respectful
  - b) Use mass media or local media to publicize, if appropriate
  - c) Obtain cooperation from local authorities
  - d) Consider material incentives for households' cooperation (though this should be done carefully)
  - e) Use appropriate interviewers for particular circumstances, e.g. woman interviewers to ask women about family planning and fertility questions.

## IV. Collect the Data (Including Monitoring of Data Quality)

Once the procedures have been established for collecting the data, the interviewers and supervisors have been trained, and the questionnaire and data entry program have been finalized, the survey teams can begin collecting the data. The data collection procedures will vary according to the program being evaluated, but some general points can be made.

## Field Work Schedule

The pilot test should provide an accurate estimate of how long it takes to collect data from a typical household (or school or community). For the sample size chosen, this can be used to determine how many interviewers and other team members are needed, and how much time they will spend in the field.



It is also important to account for transportation, both in terms of cost and in terms of time. In many surveys, field staff spend more time moving from one area to another (and arranging accommodations, introducing themselves to local officials) than they do actually interviewing respondents.

Some other points to keep in mind regarding the field work schedule are:

- Data should be collected at approximately the same time for the treatment group and comparison group. This is especially true in rural areas where work and other activities (e.g. school attendance) vary according to agricultural seasons.

- Ideally, there should be a supervisor for every 3-4 interviewers. Supervisors should do some or all the following:
  - a) **Observe** some interviews to check that interviewers are following protocol.
  - b) **Check completed questionnaires** (either paper or electronic files) to see if they are complete, and completed correctly (e.g. follow skip codes).
  - c) Randomly **re-interview households** to ensure that basic information is accurate (e.g. all household members are in the questionnaire).
- The top management for the evaluation should also conduct unannounced supervisory visits, to check not only the interviewers but also the supervisors.

## Methods to Check Data Quality in the Field



The following is general advice for ensuring data quality during the field work:

- The data entry program should be used as soon as possible after the interview has been completed. The interviewer should return to the household as quickly as possible to resolve errors and inconsistencies detected by the data entry program
- It is quite possible that errors will be found in the questionnaire and/or the data entry program. New instructions for modifying the questionnaire should be provided as soon as possible for all teams, and a corrected data entry program should also be sent to all teams as soon as possible.
- All problems with data quality that cannot be resolved in the field should be reported to the top management immediately.

## V. Further Checks of Data Quality after the Field Work

After all data have been collected, researchers can begin to work on it to prepare it for analysis. Yet before conducting the analysis further checks should be conducted which are difficult or impossible to conduct in the field.

1. Check for duplicate households (or schools, clinics or communities) as indicated by ID codes; it is surprising how often duplicates are found!
2. Key variables – such as income, total expenditure, quantities of food consumed, prices (including prices calculated by dividing expenditure by quantity), test scores, and health indicators – should be checked for unusually large or small values. Paper questionnaires can be checked for data entry errors, and interviewers can also be contacted for clarification.
3. If previous data were collected from the same households (or schools or communities), they should be matched and consistency checks should be conducted to ensure that the matches are correct.

4. Summary statistics (income, school enrollment, disease prevalence, etc.) can be compared with other sources (e.g. administrative data).
5. Important constructed variables (e.g. total income or expenditure) should be created and compared the same variables to previous surveys.
6. All changes to the data should be documented, and original (uncorrected) data files should be kept if later researchers want to try different methods to handle data problems.



## VI. Create Data Files for Analysis and Dissemination

Even before the data quality checks conducted after the field work are done, it is necessary to organize the data from each household (or school, clinic, community) into files that are organized by the type of data.

For example, many household questionnaires will be divided into sections that cover different topics, such as the household roster (list of household members with basic demographic information), education, health, employment and income. It is not very useful to have a separate data file for each household. Instead, files should be created that have household roster for all households, education for all households, etc. These can be merged using community and household ID codes to conduct analysis.

Once the data files have been rearranged in this way, and the quality of the data have been thoroughly (but hopefully quickly) checked, they are ready for analysis.

Some other points regarding preparing data for analysis are:

- Some widely used variables should be created from the data, such as total income, total expenditure, Z-scores of anthropometric measurements, and total scores on academic tests.
- The different “thematic” data sets must have all ID codes so that they can be merged with other data sets. For example, health and education data are usually individual specific, so such data from a household questionnaire should have not only the household ID number but also the individual ID number.
- The “levels” of data sets can also differ. Some data may be at the individual level, while other data at the household level or community level. In a survey of 2,000 households that has 10,000 individuals, there is no need for household level data to have 10,000 observations; 2,000 observations contain all the needed information, and this can always be merged with individual level data using household ID codes.
- Data from other sources, such as administrative data, should be combined with the newly collected data to enrich the possible analysis. The ID codes from this other data may need to be modified to match those in the newly collected data.

- It may be convenient to make the data available in several different formats, such as SAS files, STATA files, etc. However, this may not be very important given that software is available that quickly transforms data from one type to another type.
- A document should be written that explains how the data sets were created and any modifications that were done. It should contain any code numbers that are not in the survey questionnaires, and it should explain how any “summary” variables (e.g. total income, total expenditures) were created. In the long run, creating this document will reduce the time spent answering questions from data users.

## VII. Establish a System to Store, Revise and Disseminate the Data

The organization undertaking the impact evaluation should take the responsibility for storing the data, revising it (as researchers point out errors), and disseminating it to interested researchers. In some cases that organization is not expected to continue indefinitely, in which case the data could be transferred to another organization (university or research institute, government evaluation unit).

Here are some comments and advice on how to do this:

1. Many funding organizations now require researchers to make all data collected publicly available.
2. All information that could identify survey respondents (names, addresses, national ID codes) should be removed. In some cases, the names of schools, health clinics and local communities may also have to be removed. These can be replaced by codes to identify the individuals. The evaluation team should take care not to overlook this step.
3. To defray costs and reduce frivolous requests for data, modest charges could be set for providing the data to outside researchers.

4. To publicize the availability of the data, a research dissemination conference could be organized, with a large amount of publicity direct toward national and international researchers, as well as the general public.
5. Training course could be set up to teach local researchers how to use the data. Local universities could incorporate the data into their research method classes.
6. If the questionnaires are in local languages, they should be translated into major international languages (English, French, Spanish), which will make them available to more users. (Perhaps add Arabic, Russian and Chinese.)

# ***Impact Evaluation Technical Course 2013: How to Design, Manage, and Conduct Impact Evaluations***

**CLEAR-AFDC  
April 9-19, 2013  
Shanghai, China**

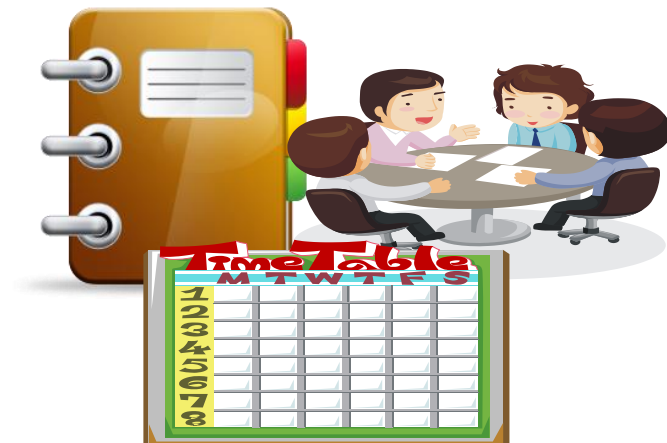
## **Session 20 Survey Management**

## I. Introduction

Implementing a survey requires a variety of skills and a great deal of logistical coordination, including each of the following:

- Budgeting and developing an overall Plan of Activities
- Human resources management
- Logistical coordination (equipment, transportation)
- Community relations

This session covers each of these topics.



## II. Budgeting and Developing an Overall Plan of Activities

The first steps in planning an impact evaluation are to set up a budget and develop an overall plan of activities.

### Budgeting

A budget should include the following items:

#### *1. Personnel*

- Management: this varies by survey, but often includes a survey manager, data manager and fieldwork manager.
- Team supervisors
- Interviewers
- Data entry personnel
- Drivers
- Specialized data collection personnel (e.g. if doing educational testing, health evaluations, anthropometrics)
- Accident insurance for field staff
- Training costs (space rental, food)



## 2. *Transportation*

- Vehicles
- Gas
- Insurance



## 3. *Equipment and supplies*

- Copies of the surveys
- Hand-held devices to enter data (if used)
- Laptops for use in the field
- Computers at central office
- Software for data entry process
- Cell phones, credit for field staff and central office
- Cash, food or other goods as compensation for respondents
- Cost of community meetings if to be held in conjunction with field work (space rental, food)



## 4. *Unforeseen expenses* (allow an extra 10% for unforeseen supplies)

**Discussion:** How much do all of these cost in your case? If you don't know, how do you find out?

## Overall Plan of Activities

Once the survey has been designed and the sample has been selected, the next step is to develop an overall plan. This step must come before you can do budgeting or hiring. Once the survey manager is hired, a more detailed plan may be developed.

This overall plan will include:

- Personnel: Who is needed and with what qualifications?
- Training: On what do they need to be trained? How long does this take?
- Transportation: What mode of transportation will be used that is safe, efficient and cost-effective?
- Timeline: How long will it take to administer the survey? What is the best time of year to implement the season? Will a follow-up survey be done later?
- Budget: Is the plan within the available budget?

Consider the following hypothetical plan for survey activities (This is only one example of the timing of activities required to implement a survey; actual survey times vary widely depending on survey requirements.):

## A Hypothetical Time-Line for Survey Activities

Month	Task	Questions to Consider
1-2	Prepare draft survey questionnaires	<i>Are other types of instruments, such as tests, needed? How much time is needed for an (average) interview?</i>
2	Draw up a budget and approximate timeline.	<i>What is more important: fewer interviewers and high supervision, or do it quickly with many interviewers?</i>
2	Hire survey manager	
3	Revise budget and plan with manager	
3	Recruit, train some interviewers	
4	Pilot the survey in one or more communities. If in multiple languages, pilot in each.	<i>Did the survey go well? Did respondents understand the questions? What can be improved?</i>
4	Revise survey and training methods	
4	Hire and train remaining interviewers	<i>Pay respondents for their time? Money or in kind?</i>
5	Contact community leaders before team's arrival	
5-6	Visit and conduct interviews in communities	
5-6	If baseline, collect information for future contact	
5-6	Follow-up for households missed by current survey	
5-7	Data entry (as soon as possible after interviews)	
7	Data cleaning	
7	Evaluate data collection experience.	<i>What can be improved for follow-up data collection (if another wave of data collection will be done)?</i>

***Discussion:*** Take a moment to look over this timeline. Discuss with a neighbor how it compares to the time you anticipate spending on your next data collection.

## Example of Another Timeline, with More Visual Depiction of Time

This study phased new software into 3 NGOs gradually, involving piloting and training in each.

		2010	2011									2012			
Activity			April	May	June	July	Aug.	Sept.	Oct.	Nov.	Dec.	Jan.	Feb.	Mar.	April
Prep.	Agreements with Min. of Health, NGOs														
	Finalization of intervention design														
	Software development														
Implementation	Software installation, 1st NGO														
	Training, 1st NGO														
	Pilot implementation, 1st NGO														
	Software installation, 2nd NGO														
	Training, 2nd NGO														
	Pilot implementation, 2nd NGO														
	Software installation, 3rd NGO														
	Training, 3rd NGO														
	Pilot implementation, 3rd NGO														
	Baseline survey - health workers, NGO 1														
	Baseline survey - health workers, NGO 2														
	Baseline survey - health workers, NGO 3														
	Household surveys, 1st NGO area														
	Household surveys, 2nd NGO area														
	Household surveys, 3rd NGO area														
	Analysis of monitoring data														
	Final analysis														
	Write paper														

## Example of a Budget

	Unit	Cost per unit (US\$)	No. of units	Total cost (US\$)	Unit	Cost per unit (US\$)	No. of units	Total cost (US\$)
A. Staff salaries	Weeks	\$7,500	2	\$15,000	Weeks	\$7,500	2	\$15,000
B. Consultant fees				\$10,250				\$27,940
International consultant (1)	Days	\$450	15	\$6,750	Days	\$450	0	\$0
International consultant (2)	Days	\$350	10	\$3,500	Days	\$350	10	\$3,500
Research assistant/field coordinator	Days	\$188	0	\$0	Days	\$188	130	\$24,440
C. Travel & subsistence								
Staff: international airfare	Trips	\$3,350	1	\$3,350	Trips	\$3,350	1	\$3,350
Staff: hotel and per diem	Days	\$150	5	\$750	Days	\$150	5	\$750
International consultants: international airfare	Trips	\$3,500	2	\$7,000	Trips	\$3,500	2	\$7,000
International consultants: hotel & per diem	Days	\$150	20	\$3,000	Days	\$150	20	\$3,000
Field coordinator: international airfare	Trips				Trips	\$1,350	1	\$1,350
Field coordinator: hotel & per diem	Days				Days	\$150	0	\$0
D. Data collection								
Data type 1: consent					School	\$120	100	\$12,000
Data type 2: education outcomes					Child	\$14	3000	\$42,000
Data type 3: health outcomes					Child	\$24	3000	\$7,200
<b>Total cost per stage:</b>				<b>\$39,350</b>				<b>\$184,390</b>

	Follow-up Data Stage I				Follow-up Data Stage II			
	Unit	Cost per unit (US\$)	No. of units	Total cost (US\$)	Unit	Cost per unit (US\$)	No. of units	Total cost (US\$)
A. Staff salaries	Weeks	\$7,500	2	\$15,000	Weeks	\$7,500	2	\$15,000
B. Consultant fees								
International consultant (1)	Days	\$450	15	\$6,750	Days	\$450	10	\$4,500
International consultant (2)	Days	\$350	20	\$7,000	Days	\$350	10	\$3,500
Research assistant/field coordinator	Days	\$188	100	\$18,800	Days	\$188	130	\$24,440
C. Travel & subsistence				\$20,000				\$20,000
Staff: international airfare	Trips	\$3,350	2	\$6,700	Trips	\$3,350	2	\$6,700
Staff: hotel and per diem	Days	\$150	10	\$1,500	Days	\$150	10	\$1,500
International consultants: international airfare	Trips	\$3,500	2	\$7,000	Trips	\$3,500	2	\$7,000
International consultants: hotel & per diem	Days	\$150	20	\$3,000	Days	\$150	20	\$3,000
Field coordinator: international airfare	Trips	\$1,350	1	\$1,350	Trips	\$1,350	1	\$1,350
Field coordinator: hotel & per diem	Days	\$150	3	\$450	Days	\$150	3	\$450
D. Data collection								\$114,000
Data type 1: consent								
Data type 2: education outcomes	Child	\$14	3000	\$42,000	Child	\$14	3000	\$42,000
Data type 3: health outcomes	Child	\$24	3000	\$72,000	Child	\$24	3000	\$72,000
E. Other								\$65,357
Workshop(s)						\$20,000	2	\$40,000
Dissemination/ reporting						\$5,000	3	\$15,000
Other 1 (clusterwide coordination overhead)						\$5,179	2	\$1,357
<b>Total cost per stage:</b>				<b>\$181,550</b>				<b>\$246,797</b>
<b>Total evaluation costs:</b>								<b>\$652,087</b>



### III. Human Resources (Personnel) Management

The composition of the survey staff varies by project. Two general issues include:

- Is it more important to finish data collection quickly (in this case, hire a large number of teams) or to have highly trained personnel that are supervised closely (hire a smaller number of teams)? This depends on time pressure and the complexity of the data collection.
- Note that turnover should be anticipated for field workers, as the work is very demanding. One approach is to hire more staff than you need to participate in training, and make a final selection after training. If staff quit, you can then hire some that participated in the training but did not “make the cut”.



## Qualifications and Responsibilities of Survey Personnel

One common approach is to have the following team members.

**Survey manager:** Generally a social scientist or statistician, or is a specialist in the project area (e.g. education or health). Should have a graduate degree or at least college degree. This is a senior position.

This person has decision-making authority, and is involved in the design of the survey. He/She regularly communicates with the research team & field team.

**Data manager:** Should have strong data management experience, statistical skills and computer skills.

The data manager designs software to manage data entry, writes data entry manuals and manages data entry operators. He or she also prepares the data sets for analysis and may produce some basic tables with results.

**Field manager:** Should have strong management experience, preferably in managing surveys. Should also have good computer skills.

The field manager is in charge of all field procedures. This includes training field staff, managing local communication, and developing sampling procedures.

**Field supervisors:** Experience managing people and details is important. Secondary education is a minimum requirement.

These people do much of the same work as the field manager, but at the team level. They are a liaison between individual interviewers and the field manager.

**Interviewers:** Should have secondary education, but the turnover risk increases if they are more highly educated.

In addition to administering the survey, the interviewers' responsibilities include establishing initial contact with households, selecting individual respondents, and assuring completion of the survey.

**Specialists** (e.g. anthropometrists, test administrators): Secondary education is preferred, but they don't need to have a medical or other degree as they can be trained. Their responsibilities vary depending on the specific project.

**Data entry operators:** Need basic computer literacy.

Data entry operators may enter data on a laptop or another mobile device in the field, or may receive completed paper questionnaires at an office and input data there. They are expected to identify potential errors and notify the field team as soon as possible. They are responsible for inputting all survey data.

## Personnel: Further considerations

### Composition of survey staff

- Gender and ethnic composition: Depending on the cultural requirements of each area and security considerations, it may be important to consider if the interviewers should be male, female, or a mix. Similarly, it may be important to have interviewers of the same religious, ethnic or linguistic group as the respondents.

### Supervision

- Depending on the scale of the project and the qualifications of specific staff members, it may be possible to combine roles. For example, one person could be both data manager and field manager (if qualified for both).
- On each survey team, an interviewer may also be able to play the role of team supervisor. If their data collection is less time-consuming, this might be a specialist or field data entry person.

## Training

Training serves two purposes: training survey team members on how to implement the survey instruments correctly and consistently, and screening for qualified applicants.

Essential activities include:

- Orientation – all personnel must understand the big picture of the evaluation. Especially if it is a randomized evaluation, every team member should be able to clearly explain the evaluation activities with a consistent message.
- Role playing – every interviewer should have a chance to practice implementing the survey, and to be critiqued.
- Evaluating – training is an important opportunity to evaluate interviewers before sending them into the field. Evaluating and grading each on a variety of criteria can be very useful.

## IV. Logistical Coordination

While it may seem straightforward, keeping track of logistics may be complex for a large-scale survey, and logistics are key to a survey's success. It is important to ensure the following, at a minimum:

- Vehicles in good working order, and adequate fuel supplies
- Accommodations in the field
- Supplies (copies of the survey questionnaires and other material)
- A “cushion” of time and personnel in case team members are ill or injured, or other unforeseen events occur
- Smooth communication channels between supervisors and other members of the team (via phone and email)

## Security

One aspect of logistics that deserves special attention is security. In some areas, doing fieldwork is a risky activity. Before going into the field, it is worthwhile to try to obtain up-to-date information on local security concerns, including:

- **Transportation.** Is it safe to use local public transportation (considering risk of traffic accidents as well as crime)? Are field personnel at elevated risk if they carry equipment such as laptops?
- **Crime.** Is it risky for field staff to travel into remote areas or at night? Is it possible to work with local partners to mitigate risk?
- **Accommodations.** Are reasonably safe accommodations available for survey teams?
- **Gender-based crime.** Are women at higher risk than men in survey areas? How can this be addressed?



## V. Community Relations

Almost all impact evaluations involve collecting data from local communities, and survey personnel need to develop a good relationship with both the leaders and the general population in those communities to ensure that the work goes smoothly.

Here are some suggestions for developing and maintaining good relationships with the local community:



1. Before visiting the community, a letter (or some other type of message) should be sent to local leaders telling them of the nature of the data collection, the overall purpose of the activity, and the approximate dates.
2. When arriving in the community, the survey team should first go to local leaders before making any other contact with community members.
3. Additional information on the survey and, more generally, on the impact evaluation, should be provided to community leaders.

4. Survey personnel need to treat everyone in the community with respect.
5. If the fieldworkers are recruiting individuals to participate in the study, it is important not to pressure hesitant individuals to participate.
6. What else? Perhaps token gifts to local leaders and to households?

**Discussion:** If you have participated in data collection in the past, how have you managed community relations? Did you have successful strategies to ensure clear communication?

## Informed Consent

Ethically, it is important for researchers to obtain what is known as informed consent from study participants. This means that study participants should be informed that they are part of a study, and it should be made clear to them what data about them the research team is collecting, and how the researchers will use the data. It is important that fieldworkers provide this information in a clear and culturally appropriate way to ensure that participants have an accurate understanding of their role in the study.

## VII. Lessons from (Unfortunate) Experiences

### Case Study 1: China Eyeglasses Study & the Informing Team

In an RCT examining the effect of having access to eyeglasses on educational outcomes for primary school students, there was money left over in the budget after purchasing eyeglasses for the treatment group. In a few townships, the leftover funds were used to buy eyeglasses for children in the *control group*. This obviously contaminated the control group, which should not have been given eyeglasses until after the study was done. Because of this, 40% of the sample was lost. This may have been avoided if every member of the research team understood the importance of the experimental design.

### Case Study 2: Asking about Food Stamps in Jamaica

In a survey administered in Jamaica, respondents were asked if they had received food stamps in the last month. Many food stamp recipients indicated that they had not. It was later revealed that this was because food stamps are distributed every two months. This type of error may be discovered in careful piloting of the survey instrument.

### Case Study 3: Preschool Study in Guatemala

In a recent evaluation of a preschool program in Guatemala, survey teams explained to families whose children were on a waiting list for preschool that which children would be able to attend preschool the following year would be determined randomly by a lottery. They were also informed that families that were not selected would be compensated with a bag of food for the time they spent responding to the survey. Many families misinterpreted the message and expected ongoing financial benefits. When these did not materialize, families suspected preschool teachers had stolen the goods. Frustrated by the experience, preschool teachers stopped collaborating with the study.

This demonstrates the importance of clear communication with survey teams and between survey teams and study participants.

## Case Study 4: Matching Students and Schools in Vietnam

In a survey of students, interviewers did not ask students to indicate where their school was located because it was assumed that all children would attend a school in their own commune. This proved not to be true, and it was not possible to match child data to school data in some cases.

## Case Study 5: School Census data in Brazil

In one year, respondents were asked if they receive payments from any government program. In the following year, they were asked instead whether they received payments from the *Bolsa* program specifically. The new question might have been closer to the information the research team wanted, but this variable was no longer comparable with the previous year's data.

## VIII. Useful References

1. Grosh, Margaret, and Juan Muñoz. 1996. “A Manual for Planning and Implementing the Living Standards Measurement Study Survey”. Living Standards Measurement Study Working Paper Number 126. The World Bank. Washington, DC.  
<http://documents.worldbank.org/curated/en/1996/05/438573/manual-planning-implementing-living-standards-measurement-study-survey>
2. Glewwe, Paul. 2005. “Overview of the Implementation of Household Surveys in Developing Countries”, Chapter IV in *Household Sample Surveys in Developing and Transition Countries*. Department of Economic and Social Affairs. Statistics Division. United Nations.
3. Muñoz, Juan. 2005. “A Guide for Data Management of Household Surveys”, Chapter XV in *Household Sample Surveys in Developing and Transition Countries*. Department of Economic and Social Affairs. Statistics Division. United Nations.  
[http://unstats.un.org/unsd/hhsurveys/pdf/Household\\_surveys.pdf](http://unstats.un.org/unsd/hhsurveys/pdf/Household_surveys.pdf)