# Household survey on Early Childhood Development: An Impact Evaluation of an AWC-cum-crèche model in Madhya Pradesh

## Data Collection, Quality Assurance and Data Management Protocols

**July 2014**

OPM builds quality management into all aspects of any project. Our quality criteria cover everything from project planning to implementation and completion and are applicable to all staff from team leaders to administrative staff. This document lists out various ethics procedures and quality control (at various levels) OPM followed in designing and managing the baseline evaluation survey.

OPM believes that evaluations are most likely to be used when they are guided by the principals of partnership, impartiality, transparency, credibility and independence.

## Research Ethics

Our research is conducted to the highest ethical standard, in line with the principles outlined in DfID's Ethics Principles for Research and Evaluation published in July 2011[1]. This includes ensuring that expectations are not raised, confidentiality is maintained and respondents are informed about the purpose of the survey and asked to participate voluntarily.

Informed verbal consent will be obtained from the research subjects.

No personal identifiers will be used in any form of reporting or dissemination. Personal identifications will be linked with a unique identifier and were kept securely.

No information will be published that could identify the respondent. Paper copies of questionnaires will be stored for three years in a secure location; only the investigation team were able to access them. Soft copy of the dataset will be shared only with the donor for the study while maintaining the anonymity of data and following the data management protocols.

Participation in the research is voluntary and respondents are free to stop interviews at any time or skip any questions they do not want to answer. They have the right to ask questions at any point before, during or after the interview is completed.

The research staff and the participants will be informed about the purpose, methods and benefits and intended possible uses of the research.

[1] https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/67483/dfid-ethics-prcpls-rsrch-eval.pdf

All interviews will be conducted by trained staff and in conditions of privacy. All interviews at the level of the community are usually conducted at the person's dwelling, or in a private room.

Only highly trained enumerators will be allowed to take anthropometric information.

## Fieldwork

OPM brings decades of experience in both quantitative and qualitative data collection of a very high standard. OPM regularly carried out data collection activities in the low income states, and have access to high quality fieldwork staff, as well as infrastructure in the region to adequately monitor and oversee field activities. Selected data investigators will undergo an extensive training period which will cover the details of the questionnaire, as well as cover interview methods, confidentiality of information and survey best practises.

The OPM Fieldwork Director and India Country Manager will ensure adequate oversight is provided to the enumerators in the field. The field enumerators and supervisors will be recruited, trained and managed by OPM. The research supervisors will be full time OPM staff, providing greater oversight over the data collection and ensuring quality management. A small team of enumerators is expected to carry out the field work in a staggered manner, allowing for closer monitoring. Short field reports will describe the implementation of the fieldwork and any problems encountered.

We propose to use computer assisted methods for field data collection using Ultra Mobile Personal Computers (UMPC) as this is more efficient and improves the quality of the data.

One of the big contemporary debates in survey work is about Computer Assisted Personal Interviewing (CAPI) versus Pen and Paper Interviewing (PAPI). In CAPI, the interviewer reads questions to the respondent from the screen of a handheld device, preloaded with the questionnaire. The respondent's answers are immediately entered into the device, which eliminates the need for manual re-keying of the data. The computer also automates the routing through the questionnaire and enables the interviewer to run a set of consistency checks during the interview, so that anomalies can be resolved with the respondent.

There is increasing consensus in the survey community that using CAPI substantially improves data quality. This is backed up by recent formal academic research and by OPM's field experience in Pakistan, the Maldives, India and Kenya.

## Quality Control in the Field

### SAMPLE SIZE

Fieldwork will be carried out in an intensive manner across two districts (Dhar and Singrauli) of Madhya Pradesh. From each district, 100 PSUs with 1500 households and 100 ASHAs and AWWs will be covered. In total, the sample size is 200 PSUs with 3000 households and 200 ASHAs and 200 AWWs. The field work in total is expected to take 40 working days.

**TEAM SIZE AND STRUCTURE**

Given the sample size, we plan to engage a team of 25 female and 5 male investigators and 5 supervisors - a total 5 data collection teams. At the ground-level each team will consist of 5 female investigators + 1 male investigator + 1 supervisor and will attempt to track the respondent and cover 1 PSU in two days.

Two female investigators will undertake household surveys and will each cover a maximum of 4-5 household interviews per day. It is important to mention here that actual tracking and identification of the eligible household will be taking a majority time in the exercise, specifically due to such sparse distribution of the eligible respondent households.

Two other female investigators will conduct the ASQ tool with each of the household been surveyed while 1 female and 1 male investigator will undertake Anthropometry, ASHA and AWW questionnaire. The Supervisor will undertake PSU questionnaire while performing the primary role of field and data collection and quality supervision.

Teams will have a buffer of around 20-30 percent investigators to account for drop-outs or non-selection.

The team deployment arrangement is explained in the table below:

| No. of teams deployed | Team Structure | | |
|---|---|---|---|
| | Study Component | No. of Persons per component | Notes |
| 5 female + 1 male investigators + 1 supervisor (Each team of 6 investigators and 1 supervisor) | Mother (Primary Caregiver) | 2 Female Interviewers | |
| | Anthropometry + AWW + ASHA | 1 Female & 1 Male Interviewer | Health enumerators with previous experience in measuring height and weight of young children will be recruited. They will be separately trained in class room and field by a senior anthropometry expert. Their care and concern towards the beneficiary and hygiene practice while performing the task will be monitored thoroughly.<br><br>For anthropometry measurements, OPM are using high quality weight and height scales by the reputed brand Tanita. One baby weight machine with height measure, one weighing machine (for children who can stand) and one |

| | | | |
|---|---|---|---|
| | | | height stadiometer will be provided per team. Anthropometric measurements will be collected using paper questionnaires. |
| | ASQ | 2 Female Interviewers | Paper questionnaires will be used to allow for notes to be taken. |
| | | 1 Supervisor (Per Team) | The supervisor will also be responsible for completing a PSU questionnaire. |

**FIELD MOVEMENT PLAN**

The teams will first cover DHAR district and then move to SINGRAULI.

**MONITORING PLAN**

To ensure the data collection is of the highest possible quality, we would institute the following procedures:

- A full time supervisor for each fieldwork team dedicated solely for oversight, mentoring and assistance will be in daily contact with the survey manager. Every investigator will be accompanied with 1 supervisor who will be the first level of QA check for the teams. The supervisors will concurrently monitor the data collected by the investigators. Our Coordinators will monitor the whole work on regular basis.
- Very close supervision by senior OPM staff for the initial week of field work, since OPM experience suggests that earlier identification of mistakes, timely correction and immediate feedback dramatically improves the quality of data.
- The fieldwork supervisor checks all data entry at the end of each survey day, aided by the in-built consistency checks written into the data capture software.
- Immediate transfer of data electronically to HQ for the data processing team to check the data entry for inconsistent, impossible or unlikely data points.
- Time allocated for re-visiting interviewees in the case of queries over the data
- Spot checks by the field monitors at all stages of the fieldwork process
- Revisiting several interviewed households by senior OPM staff to ensure that the data collected by fieldworkers has been done correctly.
- On a daily basis, each team will have a meeting where the day's experiences will be discussed and corrections made.
- Daily field reports (e.g. number of interviews conducted etc.) will be sent by field supervisors to fieldwork manager.

Full time field supervisors and field monitors will try to ensure that around 10% of interviews are spot-checked and a further 5% are back checked. Throughout the period of data entry, enumerators and supervisors are expected to be available for any query on individual questionnaires where necessary.

**ANTHROPOMETRIC DATA COLLECTION AND QUALITY CONTROL**

For this assignment, height and weight of all children aged 5 years and below in the eligible sampled household will be measured as a part of the household survey by trained investigators. We understand that it is the "length" as opposed to the height of the child that shall be measured

for children below 2 years of age. A length board or mat shall be used for this purpose on which the child shall be placed lying down.

For this survey, there will be separate enumerators and anthropometric investigators in the field teams to gain greater efficiency and allow for greater specialisation, better trained enumerators and less transaction time in the field. In general, it takes 10-15 minutes for the anthropometric investigations to complete the measurements. The anthropometric measurements require two investigators in order to ensure that measurements are precise (e.g. one to measure, one to maintain the child in the correct position).

Anthropometric data collection is much more challenging than normal data collection, and the evaluation team proposed has extensive experience of the investment, quality controls and care required to achieve robust and accurate data collection. With this kind of evaluation, very small changes to weight and height are highly developmentally significant, requiring extremely precise estimates. Some of the quality control mechanisms used by the evaluation team include:

- Using the best possible equipment and calibrating it carefully. In general, we try to use Tanita weighing scales, heightboards and stadiometers as we find they are more reliable and accurate than cheaper instruments (e.g. those made by Salter which have larger gradations).
- Selection of specialist investigators, not ordinary enumerators, who have experience of collecting anthropometric data
- Rigorous prolonged training, including field visits, for at least 10 days, for the anthropometric investigators. This includes standardisation and training more enumerators than required so the weakest can be let go after the training.
- Intensive field monitoring (this will be done by the core consortium members as well as the local partners for an additional degree of scrutiny) and dedicated supervision
- 15% independent back-checks of anthropometric data by supervisors or field monitors to identify errors and weak enumerators
- Daily checking of the recorded data against WHO growth monitoring charts to identify unlikely or impossible values, and the re-collection of data-points in this rage from the field
- Aggressive performance based contracts with local partners and extensive on-site monitoring by the international team

The extensive experience of the team shows that it is very damaging to skimp on the quality control of anthropometric data collection, and a lot of the primary data collected worldwide falls below best practice levels of robustness and credibility. Often, a large proportion of the sample has to be discarded if these measures are not followed. The team understands that our unit costs may be higher than competitors, but this reflects an unwillingness to compromise on quality and rigour.

**CAPI QUALITY CONTROL**

Whilst CAPI is demonstrated to have clear benefits in terms of data quality, its success depends substantially on the effort spent programming, piloting and testing the application, as well as on careful consideration to the underlying data management and transfer systems. To ensure quality, we would implement the following:

- Detailed and extensive piloting
- Rigorous training for the field supervisors – in advance of the training for the enumerators – to ensure that they are able to support, advise and monitor the enumerators
- Daily checks of all data collected by field supervisors (aided by the automatic CSPro in-built checks), so that the field teams are able to go back and check data points the next day if necessary

**DATA ENTRY**

As we will be using in-field data entry, data collected will be transferred electronically, by the field supervisors, once a week to the data processing staff at HQ, who will undertake additional consistency checks and load the data onto the database. The data entry programme will have in-built checks for unlikely data points, and will dynamically adjust drop down menu options to reduce the scope for errors.

At HQ, more sophisticated statistical software (STATA) will be used to identify outlier data points, and all data will be visually checked for consistency as well. Any issues will be sent back to the field teams to follow up on.

**TIMELINES**

The following section details the timelines for various activity heads:

**Training:** The training of teams for 12 days. The tentative date for the start of training in Bhopal is 25th August, 2014.

**Data tracking sheet:** The data tracking sheet will be filled for every district and will be shared with the entire team at an interval of 5-6 days.

**Data-sharing:** Considering the time lag in receiving the data from the field and time required for validation, the following timelines will be adhered to in sharing the data:

- Main Data - The main data will be cleaned after completion of a district. 3 days for each district will be allocated for data cleaning and validation.
- Error Reports - The error reports will be generated after comparing the datasets.

**DATA ENTRY AND CLEANING**

As we will be using in-field data entry, data collected will be transferred electronically, by the field supervisors, daily to the data processing staff at HQ, who will undertake additional consistency checks and load the data onto the database. The electronic data collection system allows for a large proportion of the data cleaning be carried out alongside the data collection increasing efficiency.

The data entry programme will have in-built checks for unlikely data points, and will dynamically adjust drop down menu options to reduce the scope for errors. The fieldwork supervisor will check for any errors every evening after completion of fieldwork before transferring the data to HQ from where it shall be further rechecked in real time. At HQ, a more sophisticated software will be used to identify outlier data points, and all data will be visually checked for consistency as well. Any

issues will be sent back to the field teams to follow up on. The data entered shall be checked on a daily basis by the local programme manager to ensure firstly the quality of data being collected and secondly the consistency and quality of data being entered. OPM expert Prabal Singh who has extensive experience in this regard shall provide supervision and guidance to the data processing staff to ensure that data quality checks are being regularly and correctly processed. All the data collected will be anonymized which is consistent with the specifications in the TOR.

As per the ToR, datasets will be provided in STATA and ASCII formats, with all appropriate labels and descriptions along with a fieldwork report. Field progress reports and data entered in the system will also be provided to the impact evaluation team; a feasible option since the data is being collected electronically and is uploaded on the data base on a daily basis.
It must be noted that OPM also have the ability to develop a basic set of analysis tables of key variables that automatically update as data is entered into the database, to produce an abstract of the data. The variables in these tables can be agreed with the client when the survey design is finalised. On demand, the abstract will be provided to the client, showing the level of the indicators based on data entered to date. This can be a value addition to the impact evaluation as a whole since the same analysis tables can be used for the follow up surveys as well.

In line with normal evaluation procedures, OPM will:

- Safely store all original data and transcripts for three years, after which time they will be destroyed
- Create an anonymised dataset with appropriate labelling. We will discuss with the Client merits of loading the data onto specialist websites for this purpose, including the World Bank Data Repository.