



# **YOUNG LIVES SCHOOL SURVEY**

**INDIA  
2010/2011**

## **DATA USER GUIDE**

**ZOE JAMES**

**July 2013**

**Please refer to accompanying justification documents for further details on the development and use of the Young Lives school survey questionnaires.**

## Background

This note relates to the first round of the Young Lives India School Survey, conducted in 2010-11 with a subsample of the Younger Cohort children. The survey aimed to add information on the schools and classes<sup>1</sup> attended by Young Lives children, to the existing Young Lives household panel data.

## Sample Design

Conducting the school survey in India posed a particular challenge to sample design, because those Younger Cohort children who attended school in Andhra Pradesh in 2009 were distributed across some 807 different schools, of which 538 were schools attended by a single Young Lives child. A stratified sampling approach was therefore taken, with the key requirements that:

- Adequate variation was retained across school types and urban/rural locations
- The number of schools remained logistically manageable (~300 max.)

The sampling frame consisted of all the Younger Cohort children who were enrolled in Andhra Pradesh-based schools in household survey Round 3 (2009)<sup>2</sup>. This comprised some 1880 children.

Ensuring that adequate variation was maintained across different school-types was a key consideration. The six school types that we wanted to cover and which formed the strata of the sample were:

- Urban private recognized schools
- Rural private recognized schools
- Urban private unrecognized schools<sup>3</sup>
- Rural private unrecognized schools
- Urban public schools
- Rural public schools

In each stratum, a pre-determined number of children were drawn randomly. All additional Young Lives Younger Cohort children studying in the same schools were then included, since the marginal effort of adding children was low and this maximised the within-school variation of the sample.

35 initial draws were taken from each of the unrecognized rural, unrecognized urban and urban public categories; 75 initial draws were taken from the private recognized category in both urban and rural areas; 150 draws were taken for rural public schools. The number of initial draws varied for different school types for two reasons: a) the number of sample children enrolled in each school type differed vastly, such that an equal number of draws would have been unfeasible; and b) the average

---

<sup>1</sup> Note that unlike some of the other Young Lives school surveys this is not a grade-specific survey. Instead, it tracks a subsample of Young Lives younger cohort children to the grades they are studying in in the 2010 academic year.

<sup>2</sup> Children attending school outside Andhra Pradesh were excluded because tracking was logistically infeasible and because all questionnaires, tests and procedures were designed for the Andhra Pradesh education system.

<sup>3</sup> It is worth noting here that the household Round 3 data itself did not have details about whether the school was recognized or unrecognized. To construct the above categories and then sample within them, the lack of a school census code (DISE code) was taken as a proxy for unrecognized schools

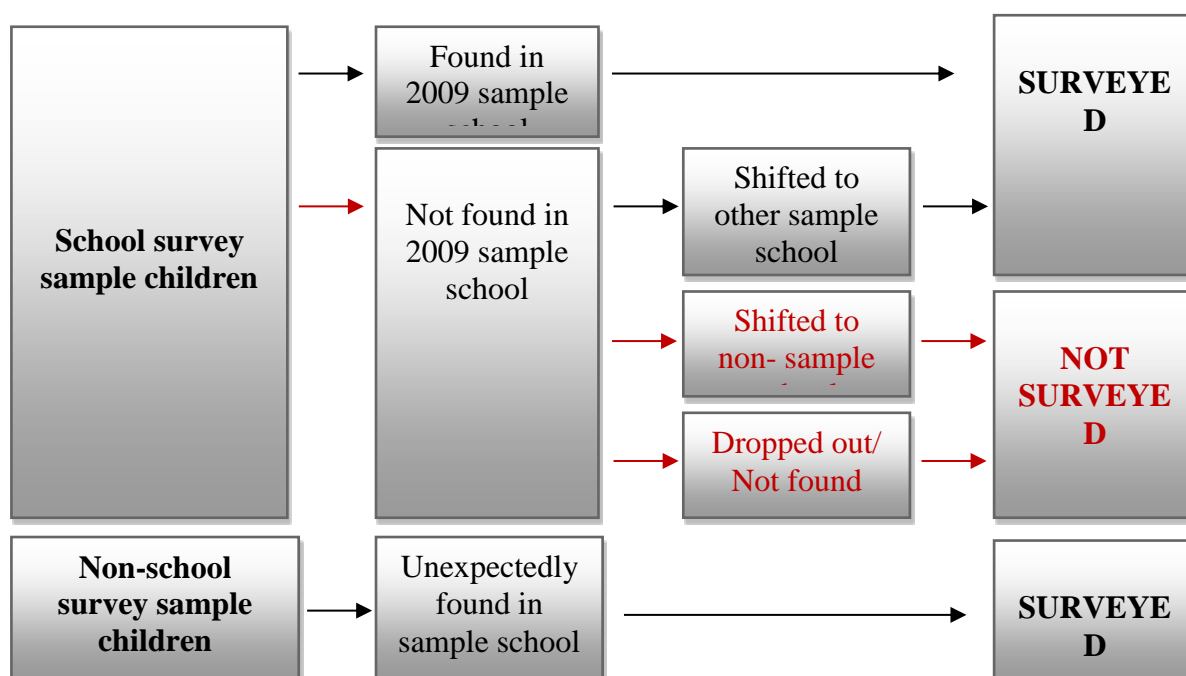
number of children per school in each of these school type categories differed vastly, such that different numbers of initial draws were necessary both to preserve some within-school variation and to keep the number of schools within the bounds of logistical feasibility.

The drawn sample comprised 1111 children across 299 schools. It was decided that if sample children had moved schools between 2009 and 2010 they would not be tracked to their new schools, unless the new school was already within the sample. This was largely a logistical decision related to the high marginal cost (both in terms of fieldworker time and resources) of adding an unknown number of schools at varying distances from the core Young Lives sites, which in many cases would be expected to contain only one Young Lives child. Any additional Young Lives children unexpectedly found in sample schools (e.g. because they had moved there since 2009) were to be included, because of the low the marginal cost of adding children once fieldwork was already taking place in a school.

### Achieved sample

In the final survey, approximately 20% of sample children were found to be no longer enrolled in their 2009 school in the 2010 academic year, such that the final achieved sample comprised 953 Young Lives children across 249 schools. The below diagram outlines the multiple reasons why children in the originally drawn school sample were not in the achieved school sample. Full details of this can be found in the roster/tracking file which enables a clear comparison of the household sample, the originally drawn school sample, and the achieved school sample.

**Figure 1: Overview of attrition and changes to the survey sample during fieldwork**



## Survey content

The survey included data collection at the school, class and pupil level, and involved the Principal, the maths teacher, and the Young Lives child<sup>4</sup>. The instruments included in the survey are detailed below in Table 1, together with a short note on their administration. All survey questionnaires are available for download at [www.younglives.org.uk/our-themes/education](http://www.younglives.org.uk/our-themes/education).

**Table 2: Survey Instruments**

<b>Instrument</b>	<b>Adminstration details</b>	<b>Languages in which translated</b>
Principal questionnaire	- Administered individually to principals	Telugu and English
Teacher questionnaires and assessment	- Administered individually to maths teachers - Teacher assessment component (of mathematical pedagogical content knowledge) self-administered under fieldworker supervision -	Telugu, English, Urdu (for self-administered component)
Child questionnaire and assessments	- Administered in groups of 6 - Questionnaire was fieldworker led and directed. - Assessments in Maths, Telugu and English were self-administered	Telugu, English, Urdu
School observation	- Fieldworker completed during time in school	Telugu and English
Classroom observations	- Fieldworker completed observation of YL child maths lesson - Both YL child/children, class peers (to make it up to 6 children per class) and the teacher were observed at intervals across each lesson and their activities recorded - Additional class-specific questions asked of teacher	Telugu and English

## Survey development

### ***Pilot work and translation of instruments***

The first piloting took place in August-September 2010 across two districts of Andhra Pradesh and was conducted by members of the Young Lives India and Oxford teams. Eight schools were involved, of different types (public/private; rural/urban) and an equal number of boys and girls were tested and surveyed at the different

---

<sup>4</sup> Note that unlike some of the other Young Lives school surveys, this survey did not collect background and assessment data on non-Young Lives children and as such a balanced class and school level sample is not achieved in all cases. The only observations which relate to non-Young Lives children come from the classroom observation, which included non-Young Lives children to achieve a balanced class-level observation of at least 6 children per class.

primary grades in which sample children were expected to be found. Following this pilot work, survey instruments were adjusted and translated.

### ***Training***

The fieldworker training took place during a period of three weeks in November 2010 at The Center for Economic and Social Studies (CESS). Each of the survey modules was introduced and discussed during training sessions. Each team of fieldworkers were then asked to conduct three practice surveys in schools of their choice under the supervision of their respective supervisors. This procedure allowed the fieldworkers to familiarize themselves with the instruments and enabled them to clarify any doubts before the final fieldwork.

### **Fieldwork**

Fieldwork was conducted across the 7 districts and 20 mandals which constitute the Young Lives sample, between 1<sup>st</sup> December 2010 - 15<sup>th</sup> March 2011.

Fieldwork teams consisted of one supervisor and two pairs of fieldworkers. Teams travelled to sites as a group, and split into pairs to conduct the survey, with the supervisor responsible for overseeing the work of each pair, and for facilitating access to sample schools. The supervisor facilitated discussion amongst the team each day to clarify any issues that arose during fieldwork. It was the role of this supervisor to check each completed instrument, clarify doubts, and report back to the Principal Investigator based in Hyderabad should they encounter difficulty.

In addition to this day-to-day supervision, members of both the India and Oxford teams conducted visits to each fieldwork team throughout the fieldwork period. These visits involved spot checks of completed protocols to ensure uniformity of completion and understanding. Where problems were identified they were investigated, and the conclusions relayed to all field teams.

For the purposes of the Classroom Observation protocol, inter-observer reliability checks were conducted in the middle of the fieldwork period (January 2011) whereby two fieldworkers observed the same children and/or the same teachers, and their scripts were returned to Oxford for comparison. For the child observation, 87.0% of observations matched and for the teacher observation, 88.5% matched, demonstrating a high level of mutual understanding in the interpretation of the protocol between fieldworkers.

### **The data**

Since each questionnaire collected data at different levels (e.g. class-level information in the teacher questionnaire) the data does not correspond to the questionnaires in terms of structure. Instead, data files are organised thematically, with variables from different parts of the survey pulled into certain levelled subfiles. The tables available with the archived data detail in full the way in which the survey instruments correspond to the data files.