



# **YOUNG LIVES SCHOOL SURVEY**

## **DATA USER GUIDE**

**ELISABETTA AURINO  
& ZOE JAMES**

**ETHIOPIA SCHOOL SURVEY  
(2012- 2013)**

**June 2014**

## **Background**

This note relates to the second round of the Ethiopia School Survey, conducted in 2012-2013 with a subsample of the Young Lives (YL) Younger Cohort children and their peers. The survey aims to offer information about the relationship between primary school children's backgrounds and their learning progress, and the data should be used in conjunction with this document, and accompanying technical documentation.

## **Sample Design**

The survey was conducted in two 'Waves', the first at the beginning of the 2012-2013 school year (Wave 1), and the second towards the end of the same school year (Wave 2). Wave 1 included all children (both Young Lives younger cohort children and non-Young Lives children) studying in all Grade 4 and Grade 5 classes in all the schools (both main school and satellite) located within the geographic boundaries of each Young Lives sentinel site. This included all school types: government, private, faith-based, NGO, community, etc. found within the sites.

Since absenteeism from school is common in Ethiopia, many children enrolled in each class were not present at the time of fieldwork. The survey sample therefore comprises all enrolled children in grade 4 and grade 5 who were present on the first day of fieldwork on which a pupil-level survey instrument was administered. Wave 2 of the survey followed up those children who had been surveyed in Wave 1.

In addition to the core Young Lives sentinel sites, the survey was extended to include ten sites in the Somali and Afar regions of Ethiopia. These sites were selected according to the same criteria laid out for selection of the original Young Lives sites (see Outes-Leon & Dercon, 2008 for more details), with the additional concern that they be safe and secure for fieldworkers to travel to and work in.

### *Additional considerations:*

- Those Young Lives children studying in sampled schools but not enrolled in Grades 4 or 5 were not included in either Wave 1 or Wave 2, since the survey was designed to be specific to Grades 4 and 5.
- Schools which contained Young Lives children studying in Grade 4 or 5, but which fell outside the geographic boundaries of the sentinel site, were not included in the survey.
- If children who had been surveyed at Wave 1 (i.e. were present on the first day of survey fieldwork in that school), were not present at the time of the Wave 2 fieldwork, they were simply not included, and no substitute children were added.
- Children who moved school between Wave 1 and Wave 2 were not tracked to their new school.

This design aimed to achieve a balanced sample of pupils at the class-level suitable for school and class-level analysis, as well as adding another layer of data to the longitudinal cohort data.

### **Achieved sample**

The wave 1 sample includes 11,983 pupils (of whom 493 are YL younger cohort children) studying in all Grade 4 and Grade 5 classes in all schools within the geographic boundaries of each sentinel site. It comprises 92 schools or 94 school sites (where satellite sites are considered separately), 280 classes and 355 teachers. The sample of classes is composed of 142 Grade 4 and 134 Grade 5 classes, respectively<sup>1</sup>.

Sixteen percent of children included in Wave 1 were not present at the time of Wave 2, producing a sample of 10,090 pupils surveyed at both the beginning and end of the school year.

### **Survey content**

In each surveyed school, the sample included:

- All pupils in Grade 4 and 5 attending on the first day on which a pupil instrument was administered
- All Grade 4 and 5 teachers of Maths;
- All Grade 4 and 5 teachers of literacy (in the relevant language of instruction)
- The Principal of the school

The instruments included in each Wave of the survey are detailed below in Table 1, together with short notes on their administration and broad scope. All survey questionnaires are available for download at [www.younglives.org.uk/our-themes/education](http://www.younglives.org.uk/our-themes/education).

---

<sup>1</sup> In the Somali region some schools had transitioned from 'ABE' to 'normal' school in the surveyed academic year. They therefore did not have a grade 4 class, most likely because those in G3 the previous year had been promoted to G5 as per the ABE structure. In affected schools, grade 3 classes were surveyed instead.

**Table 1. Survey Instruments**

Survey wave	Instrument	Content	Administration details
Wave 1	<b>School Booklet 1a</b>	Young Lives Child Site Roster	<ul style="list-style-type: none"> <li>- Pre-completed roster with information about the YL children expected to be tracked in the site</li> <li>- YL sites only</li> </ul>
	<b>School Booklet 1b</b>	Teacher and Class Roster	<ul style="list-style-type: none"> <li>- The teacher roster records the names of every teacher who teaches maths and every teacher who teaches literacy to Grades 4 and 5 in the school</li> <li>- The class roster records the details of every Grade 4 and 5 class in the school, the shift in which they are taught, and their teacher(s).</li> <li>- One per school</li> </ul>
	<b>School Booklet 1c</b>	Individual Class Roster	<ul style="list-style-type: none"> <li>- This roster records the children attending each of the Grade 4 and 5 classes in the school, as well as identifying the YL children attending the class</li> <li>- One per class section</li> </ul>
	<b>School Booklet 2</b>	School Facilities Observation	<ul style="list-style-type: none"> <li>- Collects data on school infrastructure</li> <li>- Fieldworker completed through observation of school site during their time in the school.</li> </ul>
	<b>School Booklet 2a</b>	Satellite Site Facilities Observation	<ul style="list-style-type: none"> <li>- Collects data on school infrastructure</li> <li>- Fieldworker completed through observation of school site during their time in the school.</li> </ul>
	<b>School Booklet 3</b>	Principal Questionnaire	<ul style="list-style-type: none"> <li>- Collects background data on the Principal, as well as school-level information</li> <li>- Self-administered in language of their choice, fieldworker checked and supervised</li> <li>- Sometimes principals were in charge of main sites and their satellites.</li> </ul>
	<b>Teacher Booklet 1</b>	Teacher Questionnaire	<ul style="list-style-type: none"> <li>- Collects background data on the Teacher, as well as class-level information</li> <li>Self-administered in language of</li> </ul>

			their choice, fieldworker checked and supervised
	<b>Teacher Booklet 2</b>	Teacher Mathematics Knowledge	<ul style="list-style-type: none"> <li>- Assessment of 'Pedagogical Content Knowledge' in Maths</li> <li>- Self-administered in language of their choice, fieldworker supervised</li> </ul>
	<b>Pupil Booklet 1</b>	Background & Non-cognitive Questionnaire	<ul style="list-style-type: none"> <li>- Collects background data on the pupils, as well as information on attitudes to school</li> <li>- Administered to the whole class in the language of instruction</li> <li>- Self-completed, fieldworker-led and supervised</li> </ul>
	<b>Pupil Maths Test</b>		<ul style="list-style-type: none"> <li>- Curriculum-related assessment</li> <li>- Administered to the whole class in the language of instruction of G4</li> <li>- Self-completed, fieldworker-led and supervised</li> </ul>
	<b>Pupil Literacy Test</b>		<ul style="list-style-type: none"> <li>- Curriculum-related assessment</li> <li>- Administered to the whole class in the language of instruction of G4</li> <li>- Self-completed, fieldworker-led and supervised</li> </ul>
<b>Wave 2</b>	<b>School Roster</b>		<ul style="list-style-type: none"> <li>- Collects information on teacher attendance</li> <li>- One per school</li> <li>- Fieldworker completed</li> </ul>
	<b>Class Roster</b>		<ul style="list-style-type: none"> <li>- Collects information on pupil attendance and progress</li> <li>- One per class section</li> <li>- Fieldworker completed</li> </ul>
	<b>Pupil Maths Test</b>		<ul style="list-style-type: none"> <li>- Curriculum-related assessment</li> <li>- Administered to the whole class in the language of instruction of G4</li> <li>- Self-completed, fieldworker-led and supervised</li> </ul>
	<b>Pupil Literacy Test</b>		<ul style="list-style-type: none"> <li>- Curriculum-related assessment</li> <li>- Administered to the whole class in the language of instruction of G4</li> <li>- Self-completed, fieldworker-led and supervised</li> </ul>

## **Survey Development**

The survey was developed following a series of stakeholder meetings, in which key education priorities and research questions were identified, and the potential contribution of a Young Lives school survey in relation to existing research activities was clarified.

### *1. Development and Piloting of Wave 1*

Survey questionnaires and pupil assessment instruments were developed during the Spring and Summer of 2012. Since pupil assessments of maths and literacy were to be administered in seven different languages of instruction, specific language consultants were employed to work on the development of test items during this period (see associated technical documentation on the development of the assessments for more details). All other instruments were translated into the languages of instruction used across survey schools.

Pilot work then took place in 6 regions (Amhara, Oromia, Tigray, SNNP, Somali and Afar) and 5 languages of instruction (Amharic, Oromiffa, Tigrigna, Sidama, and Somali) at the start of October 2012. Staff members from the Ethiopian Development Research Institute (EDRI) and from Oxford University were involved in the piloting. The pilot led to the refinement of the questionnaires and tests, and all changes were reflected across all language versions of the instruments.

### *2. Development and Piloting of Wave 2*

Since Wave 2 of the survey was made up of only a revised pupil assessment and additional school and class rosters, further pilot work was not conducted. Instead, the new assessment items were selected from maths and literacy item banks piloted as part of the Young Lives' household survey, which was being developed at the same time. The training for wave 2 included within it a pilot of the roster instruments, which provided sufficient opportunity to test and refine the instruments prior to the final fieldwork.

### *3. Fieldworker training for Wave 1 and Wave 2*

Supervisor training for Wave 1 of the survey happened at the end of October 2012 in Addis Ababa and was led by staff from EDRI with supervision by staff from the University of Oxford. Each survey module was introduced and discussed during training and fieldworkers were asked to complete training exercises, after which teams visited schools to practice administering different survey components. Following this, supervisors and EDRI staff members led regional fieldworker training.

Each fieldworker and supervisor was issued with a manual<sup>2</sup> explaining the general purpose of the survey, together with a more detailed breakdown of each component and its key considerations.

---

<sup>2</sup> The manual is available for download on the Young Lives website at [www.younglives.org.uk/our-themes/education](http://www.younglives.org.uk/our-themes/education)

A separate training exercise was undertaken for Wave 2 of the survey, at the end of April 2013. Supervisor training was once again held in Addis Ababa and was led by staff from the University of Oxford and EDRI. Supervisors visited schools to practice administering the rosters, which also offered the opportunity to refine the instruments. Regional fieldworker training then took place, led by Supervisors and EDRI staff. A separate manual for this second wave was also developed and issued to each fieldworker and supervisor.

### **Fieldwork**

Wave 1 fieldwork took place in October and November 2012 and was conducted by fifteen supervisors and forty enumerators split into teams working over a period of 4-6 weeks. Wave 2 fieldwork took place in April and May 2013 and was conducted by 12 supervisors and 37 enumerators working in teams over 3-4 weeks.

### **The data**

The data are hierarchically structured, at the pupil, teacher, class, school site and principal level. The files contain information on both Young Lives children and their peers, and these can be clearly identified via the 'childid' and 'pupildid' identifiers. The below table details the identifiers present at each level of the data, together with the combination of IDs needed to merge files across levels.<sup>3</sup>

As the classes may have either a single teacher who teaches all subjects, or separate subject teachers, the combination of the school id and class id does not uniquely identify the teacher (as in the case of 1:1 teacher-class ratios). The link between teacher-level (e.g. School Booklet 1b) and class-level (e.g. School Booklet 1.c) data is accordingly done separately for the maths and literacy teachers. For this reason, there are different unique identifiers for the analysis of teacher/class level data related to the teacher being identified as either the maths or the literacy teacher. In those cases where the same teacher teaches both subjects, the teacher level information is duplicated.

Figure 1 below clarifies this point. The teacher roster (e.g. School Booklet 1b) is used to record the IDs of the teachers who teach maths and language of instruction in all the Grade 4 or 5 classes in the school, while the class roster (e.g. School Booklet 1.c) identifies who teaches maths and literacy in each class. For example, in class 5A (with class id C02) Ashmelash Haile teaches both maths and literacy. In order to link the teacher-level data with the class-level data, it is necessary to refer back to the teacher roster to match the Teacher ID of Ashmelash Haile (T01) and separately link this ID with the in the column 'maths teacher ID' and the column "language/literacy teacher ID".

---

<sup>3</sup> NB not all identifiers are present in all data files, since the data are organised at different levels.

**Figure 1. Linking Teacher-level with class-level data**

1	2	3
1. Teacher ID	2. Teacher's Name	3. Number of Days Absent Since 21 Maskarram 2005 (E.C.) (01 Oct 2012 G.C.) (N.B. this includes any day of absence, regardless of reason e.g. Including for sickness)
T01	Ashmelash Haile	6
T02	Nardos Tesfay	5
T03	Abraham Yadete	5
T04	Worheneh Elemu	4

1. Class ID	2. Class Name in School e.g. 4A	3. Grade (4 or 5)	4. Shift (01=Morning/ 02=Afternoon/ 03=Whole Day)	5. Maths Teacher ID (from teacher roster)	6. Language Teacher ID (from teacher roster)
C01	4A	4	01	T02	T02
C02	4B	4	02	T01	T01
C03	5A	5	01	T03	T04
C04	5B	5	02	T03	T04

Table 2 provides the combination of unique identifiers.

**Table 2. Combinations of IDs**

Level	Unique ID Combination
Pupil level (All children)	schlid classid pupilid
YL Child level	childid
Class level data	schlid classid
School level data	schlid
Maths teacher level data	Schlid mtchid
Language teacher level data	Schlid lngtchid
Principal level data	schlid

## References

Outes-Leon, I. and Sanchez, A. (2008) An Assessment of the Young Lives Sampling Approach in Ethiopia, *Technical Note 1*, Oxford: Young Lives.