fhi360
THE SCIENCE OF IMPROVING LIVES

EDUCATION POLICY AND DATA CENTER
Making sense of data to improve education

Household Survey Guidelines on Education

for use in the context of the IHSN Question Bank

EPDC Working Paper No. EPDC-09-04

December 2009

CONTENTS

APPENDICES

Click here to download appendices [1]

FIGURES

[1] http://epdc.org/policyanalysis/static/HouseholdSurveyGuidelineOnEducationAppencies1to4.pdf

ABBREVIATIONS

CWIQ	Core Welfare Indicators Questionnaire
DHS	Demographic and Health Surveys
EAM	*Enquête Auprès* des M*énages*
ECCE	Early Childhood Care and Education
EdData	Education Data for Decision-making
EdStats	Education Statistics Database (World Bank)
EFA	Education for All
EGRA	Early Grade Reading Assessment
EPDC	Education Policy and Data Center
FTI	Fast Track Initiative
GAR	Gross Attendance Rate
GER	Gross Enrollment Rate
GIR	Gross Intake Rate
GMR	Global Monitoring Report
HHS	Household Survey
HIES	Household Income and Expenditure Survey
IALS	International Adult Literacy Survey
IHS	Integrated Household Survey
IHSN	International Household Survey Network
ISCED	International Standard Classification of Education
LAMP	Literacy Assessment and Monitoring Programme
LSMS	Living Standards Measurement Survey
LSS	Living Standards Survey
MCC	Millennium Challenge Corporation
MDG	Millennium Development Goals
MICS	Multiple Indicator Cluster Survey
NAR	Net Attendance Rate
NER	Net Enrollment Rate
NIR	Net Intake Rate
OECD	Organisation for Economic Co-operation and Development
QUIBB	*Questionnaire des Indicateurs de Base du Bien-être*
SISEE	Statistical Information Systems Measuring Expenditure on Education
SLE	School Life-expectancy
UIS	UNESCO Institute for Statistics
UNESCO	United Nations Educational, Scientific and Cultural Organization
UNSD	United Nations Statistics Division
UOE	UNESCO/OECD/EUROSTAT
USAID	United States Agency for International Development
WMS	Welfare Monitoring Survey

INTRODUCTION

Household surveys are an important source of education data for policymakers at the national and international levels. The information gathered from household surveys are especially valuable because they can be cross-referenced with demographic, health, employment, income, expenditure, and other topics covered in the surveys, making it possible to correlate information and perform detailed analyses that would not be possible with information from any other source.

Despite the rich potential of household surveys as a source of education data for policymakers, education topics are treated unevenly in household surveys. The most broadly used series of household surveys, DHS and MICS, use an efficient set of 6 or 9 education questions that yield quality data, but cover only the most fundamental education topics; other surveys ask more than 40 education questions touching on a variety of topics, but lack rigorous design and wording. On a similar note, inconsistencies in the sequence or wording of questions across surveys can compromise the comparability of data.

The purpose of this report is to evaluate the education questions and modules used in a cross-section of household surveys, distill their best practices, and organize these practices into a coherent and coordinated suite of education questionnaire modules. The overall effort is indicator driven in that we begin by identifying the specific education indicators to be collected, and then evaluated best practices for questions and modules needed to calculate these indicators. The education indicators covered have been selected on the basis of their policy relevance, the attention they are given in international frameworks and agreements (such as the Millennium Development Goals), and the degree to which household surveys are an appropriate instrument for accurately and reliably measuring them. A detailed account of the process used to select indicators is given in the Developing the List of Indicators section of this report.

After identifying and prioritizing the education indicators to be recommended, we grouped them according to the sequence of questions needed for their calculation. These sequences of questions are referred to as modules because they can be added or removed in modular fashion depending on the policy priorities of each survey. The wording and sequence of questions and instructions within each module, and well as the sequence of modules themselves, are based on what we identified as best practices for each module. The overall process of developing and prioritizing modules is discussed in the Developing Questionnaire Modules and Questions section of the report; each module is treated individually in the Core Modules and Auxiliary Modules sections of the report.

The report is complemented by three more technically-oriented appendices. **Error! Reference source not found.** diagrams the flow of modules and of questions within each module. **Error! Reference source not found.** illustrates questionnaires for each module. **Error! Reference source not found.** provides metadata for each module and indicator, as well as instructions for calculating intermediate variables, and **Error! Reference source not found.** displays recommended tabulations for each indicator.

The overall project is funded through the International Household Survey Network (IHSN) and will feed into the IHSN's *Question Bank* (www.ihsn.org), a project developed to foster harmonization and improvement of data collection methodologies by providing a central repository of indicators, classifications, concepts and questions, and reference materials, and by providing international guidelines

on survey design and indicator calculation.

Developing the List of Indicators

In order to develop as inclusive a list of indicators as possible, EPDC took the approach of developing a broad list of education indicators from a variety of sources and contexts, and then removing indicators from the list based on three criteria: 1) indicators that cannot be obtained from a household survey, 2) indicators that can be obtained more effectively from another source, and 3) indicators with such narrow policy relevance that they should not be broadly recommended

Building the list

EPDC compiled a list of education indicators gathered, utilized, published by a variety of organizations:

- Indicators specifically referenced in the indicative frameworks of international compacts, and international monitoring institutions, specifically:
    - United Nationals Millennium Development Goals (MDG)
    - Education for All (EFA) Goals
    - Education for All Fast Track Initiative (FTI) Indicative Framework
    - Millennium Challenge Corporation (MCC) selection Criteria
- Indicators maintained and published in a selection of internationally recognized databases
    - UNESCO Institute for statistics (UIS) Education Database
    - World Bank EdStats Education Database
- Indicators that EPDC currently extracts from DHS, MICS, and other household surveys
- Indicators that could be extracted from the various household surveys that EPDC analyzed in a comparative study for IHSN/OECD in 2008:
    - Integrated Household Survey (IHS)
    - Living Standard Survey (LSMS)
    - Core Welfare Indicators Questionnaire (CWIQ; QUIBB in French)
    - Household Income and Expenditure Survey (HIES)

    - Ten country-specific surveys

- Indicators that could be extracted from EdData surveys

- Indicators that are not currently collected through household surveys but which EPDC believes could be of high use to the education policy community. This category includes several indicators that we believe to be impractical to collect through a household survey, but wanted to include in the initial conversation in order to generate ideas.

Narrowing the list

Once a master list of education indicators was compiled, EPDC consolidated the list to remove variant disaggregations of the same basic indicator (for example, 'Grade 1 Repetition Rate', 'Primary Female Repetition Rate', and 'Rural Secondary Repetition Rate' can all be categorized as 'Repetition Rate.' EPDC also removed indicators that are based on data that cannot be collected at the household level, indicators describing government expenditure on education, indicators describing the education system, and indicators describing specific characteristics of schools and teachers, for example. The remaining indicators were organized into a preliminary set of modules to explore in further detail.

EPDC shared this preliminary list of indicators and manuals with a selection of education policymakers and solicited feedback and input in a meeting and through continued communication via email. Based on feedback from the group, EPDC eliminated several modules that it had proposed as modules targeting indicators that are not currently collected but could be of use to the policy. Participants included:

Kristi Fair	United States Agency for International Development
Luis Crouch	Research Triangle International
Kurt Moses	FHI 360 - Systems Services
Mamadou Thiam	Education For All Fast Track Initiative Secretariat
Friedrich Huebler	UNESCO Institute for Statistics
George Ingram	Education Policy and Data Center
Annababette Wils	Education Policy and Data Center
Ania Chaluda	Education Policy and Data Center

Proposed Indicators

A list of the education indicators selected for inclusion in the report is given in Figure 1 on page 10. A brief explanation of education indicator groups that were not selected for inclusion in the question bank is given in Figure 2 on page 11. Technical definitions of each indicator can be found in Error! Reference source not found.

The Indicators listed in Figure 1 are organized according to the module that facilitates their extraction rather than the module with the name that most closely fits the indicator category that the indicator is commonly grouped into. This is the most practical way to group indicators for the purposes of this report, though the result is some indicator/module that may appear strange to education policymakers

    - Ten country-specific surveys

- Indicators that could be extracted from EdData surveys

- Indicators that are not currently collected through household surveys but which EPDC believes could be of high use to the education policy community. This category includes several indicators that we believe to be impractical to collect through a household survey, but wanted to include in the initial conversation in order to generate ideas.

Narrowing the list

Once a master list of education indicators was compiled, EPDC consolidated the list to remove variant disaggregations of the same basic indicator (for example, 'Grade 1 Repetition Rate', 'Primary Female Repetition Rate', and 'Rural Secondary Repetition Rate' can all be categorized as 'Repetition Rate.' EPDC also removed indicators that are based on data that cannot be collected at the household level – indicators describing government expenditure on education, indicators describing the education system, and indicators describing specific characteristics of schools and teachers, for example. The remaining indicators were organized into a preliminary set of modules to explore in further detail.

EPDC shared this preliminary list of indicators and manuals with a selection of education policymakers and solicited feedback and input in a meeting and through continued communication via email. Based on feedback from the group, EPDC eliminated several modules that it had proposed as modules targeting indicators that are not currently collected but could be of use to the policy. Participants included:

Kristi Fair	United States Agency for International Development
Luis Crouch	Research Triangle International
Kurt Moses	Academy for Educational Development - Systems Services
Mamadou Thiam	Education For All Fast Track Initiative Secretariat
Friedrich Huebler	UNESCO Institute for Statistics
George Ingram	Education Policy and Data Center
Annababette Wils	Education Policy and Data Center
Ania Chaluda	Education Policy and Data Center

Proposed Indicators

A list of the education indicators selected for inclusion in the report is given in Figure 1 on page 10. A brief explanation of education indicator groups that were not selected for inclusion in the question bank is given inFigure 2 on page 11. Technical definitions of each indicator can be found in **Error! Reference source not found.**

The Indicators listed in Figure 1 are organized according to the module that facilitates their extraction rather than the module with the name that most closely fits the indicator category that the indicator is commonly grouped into. This is the most practical way to group indicators for the purposes of this report, though the result is some indicator/module that may appear strange to education policymakers

accustomed to seeing categories in different groupings.

As an example, ‘Gross Intake Rate to the First Grade of Primary’ (GIR) is usually thought of as a measure of access and by this criteria one might expect it to be grouped with the participation indicators. The questions in the Participation module, however, are not sufficient to calculate GIR because they do not convey information about the grade or level attended during the previous year and, to properly calculate GIR, one must be able be able to distinguish between grade 1 pupils who are attending primary for the first time and those who are repeating the grade.

Information from the Efficiency questions are needed to make this distinction, so the indicator is grouped under the Efficiency module heading. Similarly, ‘School Life Expectancy,’ which one might typically think of as a measure of efficiency, is classified under the participation module because it does not require information from the efficiency module questions for calculation.

Figure 1: Indicators Selected for Inclusion in IHSN Question Bank

Priority	# Quest-ions	Module / Indicator[2]	EFA Goals	MDG Goals	FTI Ind. Frmwrk	MCC Criteria	UIS, EdStats, EPDC
Univrsally Relevant, Should be included in all Surveys (CORE)		**Learning**	●	●			●
	1	Literacy Rate	●	●			●
	1	Numeracy Rate		●			
		Attainment					●
	3	Educational Attainment		●			●
		% Ever Attended School					●
		Participation	●	●	●		●
	3	Net Attendance Rate	●	●			●
		Gross Attendance Rate	●				●
		Parity Measures for Attendance	●	●			●
		% Pupils Underage					●
		% Pupils Overage					●
		% Pupils On-time					●
		% Children out of School					●
		Total Net Attendance Rate					●
		Voc/Tech as a % of secondary enrollment					●
		School Life Expectancy (Primary to Secondary)					●
		Efficiency	●	●	●	●	●
	3	Primary Completion Rate (GIR to last grade of Primary)	●	●	●	●	●
		Net Intake Rate to the first grade of primary	●		●		●
		Gross Intake Rate to the first grade of primary	●				●
		Survival Rate (Grade 1 – Grade 5)	●	●			●
		% Repeaters			●		●
		Repetition Rate	●				●
		Dropout Rate					●
		Promotion Rate					●
		Primary to Secondary Transition Rate					●
		Graduation Rate					●
		New Entrants to G1 with Pre-Primary experience in the previous year	●				●
Auxiliary - Recommended		**School Characteristics**	●		●		●
	1	% Pupils attending private schools			●		●
		Decisions on Education					
	2	Reason for not attending school					
		Reasons for leaving school					
	2	Transfer Rate					
		Reasons for switching schools					
		Household Expenditure on Education	●				
	7 - 36	% Pupils whose households spent money on their education					
		Average non-zero per-pupil household expenditure on education					
Auxiliary - Optional		**Participation in Scholarship Program**					
	2 - 3	% pupils using NNN scholarship					
		Apprenticeship, Literacy Training & Out of School Education	●				
	1-6	Educational Attainment by category of non-formal education	●				
		Attendance rate by category of non-formal education					
		Opportunity Cost of Education					
	4	Total time commitment of education					
	1	Amount of time it takes to travel to school					
	1	# hours child typically spends at school					
	1	Amount of time spent on school-related activities on a typical day					
	1	Hours/week pupil spends on homework					
		Auxiliary Attainment					
	1	Highest Diploma earned					

Figure 2 Education Indicator Families not selected for inclusion in the IHSN Question Bank

[2] Technical definitions of each indicator can be found in **Error! Reference source not found.**.

Government Education Expenditures
These data are important but cannot be obtained through household surveys. They are gathered through government records.

Education Foreign Assistance
These data are important but cannot be obtained through household surveys. They are gathered through government records.

Teachers
Because teachers represent a small subset of the overall population surveyed, sampling is not an effective way to gather information about them. Administrative data maintained by agencies such as the ministries of education or finance may be a better source. Data gathered through an annual school census may also be an excellent source of information.

Tertiary Education
Sampling techniques such as those used in household surveys are not a good source of information on tertiary students because, with rare exceptions, students living dormitories or other institutional settings are not included in survey samples. Because tertiary students are usually concentrated in a small area rather than distributed across the country. Tertiary institutions generally have good record-keeping practices, it is best to collected these data directly from the institutions. Tertiary education is included as a level in attainment and attendance modules of the household survey.

Advanced measures of Numeracy and literacy (such as LAMP or EGRA)
These are too complex and costly to administer in a household survey.

Life Skills
A complex measure of 'Life Skills' would be too complex and costly to develop and implement in a household survey.

Sources of financial support for education received by the household
International remittances and similar cash flows from outside of the household unit are of policy interest because may help to defray the costs of education in some households. However, because inflows are not likely 'earmarked' for education both rather flow into a general household cash reserve, it is difficult for respondents to pinpoint the amount of financial support that goes to education expenditures. This issue might be better addressed in survey modules designed to measure household income and expenditure.

Classroom time use & pedagogical practices, Primary Learning Language at School
These indicators cannot be gathered through a household survey because survey respondents are unlikely to possess knowledge of specific classroom activities. They could be gathered through a school survey or school census.

% Pupils who attend school as boarding students
The great majority of household surveys do not cover individuals living in dormitories or other institutional settings.

Perceptions of shortcomings of the pupil's school
Questions that ask the respondent about the shortcomings of a pupil's school make the unrealistic assumptions that the respondent is familiar with-, and able to make an objective assessment of conditions of the school, and that the respondent knows what ideal school conditions should be.

Attitudes towards education
Household surveys may not be the best source of information for these types of questions because survey respondents may be household members who are not closely related to children in school and whose attitudes towards education are not relevant to decisions on a child's education. Overall, questions about attitudes towards education should be tailored to the specific context of the survey, and we cannot recommend them in broad terms.

% Parents who help children with their homework
Frequency of assistance with homework
% parents who read to/with their children
% parents who have visited in their child's school in the past year

These indicators are not commonly gathered or catalogued. Because the interview respondent may not be the parent of the pupil to whom the questions apply, the respondent may not be able to provide accurate responses.

% households that made a non-financial contribution to the school/teacher
It is difficult to define and measure non-financial contributions.

 # books in the home
There is no precedence for the collection of this data in household surveys.

Developing Questionnaire Modules and Questions

After developing a list of indicators to collect through household surveys, we identified the component pieces of information needed to calculate each indicator. We then looked to a cross section of education questionnaires currently in use to identify best practices for collecting these pieces of information. This component of the study builds off of a 2008[3] report in which EPDC evaluated education modules for 30 household surveys and developed recommendations for improving the structure and wording of education questionnaires.

Identifying best practices

The best practices for the questionnaire modules and questions recommended in this report are drawn from the 30 questionnaires evaluated in the 2008 report, a standard DHS IV questionnaire, a standard MICS3 questionnaire, and the 2002 Zambia EdData questionnaire. Figure 3 lists the questionnaires considered in this report.

Best practices were identified using criteria and suggestions gathered from a variety of sources, including feedback and suggestions from the participants in our information-gathering meeting, EPDC's 2007 Report, the 2008 'UN Principles and Recommendations for Population and Housing Censuses'[4], and chapter 2 of 'A Manual for Planning and Implementing the Living Standards Measurement Study Survey.'[5]

Figure 3 List of Household Survey Questionnaires Evaluated

Survey	Country	Year
Demographic and Health Surveys (DHS)	Malawi	2004
Education Data for Decision-making (EdData)	Zambia	2002
Enquete Aupres des Menages	Madagascar	2001
Enquête Camerounaise auprès des Ménages (ECAM)	Cameroon	2001
Enquête Djiboutienne auprès des Ménages	Djibouti	1996
Etude Nationale sur les Conditions de Vie des Populations	Burundi	1998
Expenditure and Consumption Survey (ECS)	Lao PDR	2003
Household Budget Survey (HBS)	Tanzania	2000
Household Income and Expenditure Survey (HIES)	Bangladesh	2000
Household Income and Expenditure Survey (HIES)	Sri Lanka	2002
Household Socio-Economic Survey (HSES)	Cambodia	2003
Inquerito as despesas e receitas familiares (IDRF)	Cape Verde	2001
Integrated Household Survey (IHS)	Malawi	2004
Integrated Household Survey (IHS)	Pakistan	2001
Integrated Household Survey (IHS)	Sierra Leone	2003
Living Standard Survey (LSS)	Bhutan	2003
Living Standard Survey (LSS)	Cote d'Ivoire	2002
Living Standard Survey (LSS)	Ghana	1998
Living Standard Survey (LSS)	Nepal	2003
Living Standard Survey (LSS)	Nigeria	2003
Living Standard Survey (LSS)	Vietnam	2004
Multiple Indicators Cluster Survey (MICS)		
National Household Survey	Uganda	2002
Questionário de Indicadores Básicos de Bem-Estar (QUIBB)	Mozambique	2002
Questionnaire des Indicateurs de Base du Bien-etre (QUIBB)	Benin	2003
Questionnaire des Indicateurs de Base du Bien-etre (QUIBB)	Burkina Faso	2003
Questionnaire des Indicateurs de Base du Bien-etre (QUIBB)	Gabon	2005
Questionnaire des Indicateurs de Base du Bien-etre (QUIBB)	Niger	2005
Socio-Economic Survey (SES)	India	2004
Socio-Economic Survey (SES)	Indonesia	2002
Socio-Economic Survey (SES)	Thailand	2002
Vulnerability and Poverty Survey (VPAS)	Maldives	2004
Welfare Monitoring Survey (WMS)	Ethiopia	2000

[3] How (Well) is Education Measured in Household Surveys? A Comparative Analysis of the Education Modules in 30 Household Surveys from 1996-2005. EPDC No. 0801, Spring 2008. http://epdc.org/static/HowIsEducationMeasuredInHouseholdSurveys.pdf

[4] Principles and Recommendations for Population and Housing Censuses. Revision 2. United Nations Statistics Division. New York, 2008.

[5] A Manual for Planning and Implementing the Living Standards Measurement Study Survey. World Bank. LSMS #126, 1996.

The following criteria were considered for identifying best practices:

- Questions gather information that accurately reflect the information the indicator is intended to represent
- Questions gather information that is as close to an internationally-defined standard for the indicator as possible
- Questions are worded in a way that is easy for the respondent to understand
- Questions gather information that the respondent knows the answer to
- Question filtering does not rely on any assumptions about relationships between pieces of information gathered
- The highest priority information is gathered earlier in the questionnaire, and lower priority information is gathered later in the questionnaire
- Instructions for both the respondent and the interviewer are clear and easy to understand
- For each question, timeframe, subject, unit of measurement, and other frames of reference are clearly stated and optimized for clarity and ease of response
- When there is a possibility that different respondents may prefer to respond to the same question using different units of measurement, the questionnaire should allow the respondent to specify the unit of measurement used in their response.[6]
- Shifts in question timeframe, subject, unit of measurement, and other frames of reference are minimized
- Where shifts in question timeframe, subject, unit of measurement, and other frames of reference cannot be avoided, guiding language is used to smooth the transition from one frame of reference to the next[7]
- Questionnaire flow is simple, minimizes filtering, and makes sense

[6] Specifying the unit of measurement makes it possible to postpone converting the data until a post-enumeration stage when a computer can be used to make the conversion. This practice saves time and eliminates opportunities for error over unit conversions performed at the time of the interview.

[7] For example, when shifting from questions about the current year of schooling to the previous year of schooling, leading language would be a statement read by the interviewer along the lines of: "Up until now we have been talking about NAME and his/her schooling during this current year. Now I would like to talk about NAME and his/her schooling one year ago, during the previous year of school."

- Data is gathered using as few questions as possible
- Compound questions and leading questions are avoided or broken down into simpler questions
- Whenever responses to a question can be anticipated, pre-coding is used in order to minimize opportunities for coding error in responses
- 'Don't know' is always a response option

Recommending Practices

After we identified what we believed to be the best practice for a module or a question, we recommended that same practice word for word. However it is often the case that the best possible practice for a particular module or question could be achieved by bringing together a combination of best practice elements found in two or more questionnaires. When this was the case, we identify our recommended best practice and explain the genesis of this recommendation.

When EPDC was not able to identify a best practice for gathering particular information, EPDC used principles exemplified in existing best practices for other questions to develop vest practice questions. Because these questions have not been tested in a field setting, it is important that they be vetted and tested by outside experts before they are included in the IHSN question bank. These unproven questions are clearly marked as 'Recommended for *Consideration*' for inclusion in the question bank, whereas proven questions are marked as 'Recommended.'

Recommendations that apply to all modules

Guidelines for Developing Response Categories

Many of the modules proposed in this report use questions that are paired with response option menus. The response option menus try to anticipate types of responses or information that the question will elicit, and provide guidance to the interviewer on how the responses should be categorized and coded for processing. Most response option menus must be modified by the questionnaire developer to reflect the context where the survey will be conducted. For example, response categories to the question 'What grade and level of school did NAME attended during the most recent school year,' must be modified because some countries have four grades of primary and other countries have eight grades of primary. Wherever there is a response option menu that needs to be modified by the questionnaire developer, it is noted in this report and in the metadata for the module that contains the question.

While actual response categories will vary from survey to survey, there are general guidelines that should be followed to ensure that response categories are well crafted. These guidelines are outlined below. If additional guidance is required for developing a particular response option menu, this guidance is provided in the report discussion for the corresponding module.

Questionnaire developers should develop *and test* a set of response categories that are appropriate to the context where the survey will be conducted. Illustrative examples of response option menus are a good starting point for this, and are provided in the report discussion of each module. When devising a custom

set of response categories, the following considerations should be kept in mind:

1. Response categories must use words and terms that are meaningful to household respondents and that household respondents will be familiar enough to choose from. For example 'government subsidized private NGO school' should be avoided if the typical household member is unlikely to know whether a particular religious school receives government subsidies, but 'BRAC school' could be used if the typical respondent is likely to be able to identify a school in this way.

2. Response categories must be exhaustive. This means that the response options should cover all (or very nearly all) possible responses to the question. For example, if a question asks 'What type of school does NAME attend?' and the response options are 'Government-run' and 'Church-run', but a large number of children attend schools that are neither government run nor church run, then the question response options are *not* exhaustive; at least one more response option needs to be added to cover this additional category of responses.

3. Response options must be mutually exclusive. This means that categories should be devised so that it is impossible for there to be more than one correct response option to the question. For example, if a question asks 'What type of school does NAME attend?' and the response options are 'Government-run,' 'Church-run,' 'English-medium,' and 'Non-English-medium' the response options are not mutually exclusive because it is possible to attend a Government-run English-medium school.

4. The survey documentation should include a clear definition of each school type response category, as well as an explanation of how the categories can be distilled into the broader categories needed by data analysts. For example, if the purpose of the question is to determine the proportion of pupils attending public and private schools but the response option categories are something other than 'public' and 'private,' then guidance needs to be given on how the response option categories can be mapped to 'public' and 'private'

5. Response Categories should not be redundant with information gathered elsewhere in the survey. For example, do not distinguish between 'Primary Government' and 'Secondary Government' since the questionnaire already asks about school levels in other modules.

6. 'Other' should always be given as an option since it is inevitable that response options will not be completely exhaustive.

7. 'Don't Know' should always be an option since there will always be some respondents who don't know the answer to a question.

8. Questionnaire response options should always be tested before they are fully implemented. Questionnaire designers should check to ensure that the conditions outlined here are being met. As a rule of thumb if, the proportion of responses given as "don't know" plus the proportion of responses given as "other" exceeds 5%, then the categories need to be revised.

Questionnaire Case Conventions

In the questionnaire, instructions for the interviewer and other pieces of text that are not to be read out loud are printed all in UPPER CASE letter. Text that is to be read out loud by the interviewer is printed in regular sentence case. Also listed in UPPER CASE are elements of text that have been left generic in this report because they need to be made specific to the context of each survey.

The following text elements are generic in the questionnaire in Appendix 2, and must be customized to the context of the survey. When these elements are completed with context-specific information, the upper case space-holder should be replaced with a sentence-case word.

> '**YYYY' (current school year)**: Replace this with the year-range of the current or most recently completed school year.
>
> '**YYYY-1' (previous) school year**': Replace this with the year-range of the school year prior to the current or most recent school year.
>
> **'NNNN' (Scholarship Program)**: Replace this with the name of the scholarship program.

School years and how they relate to the structure and wording of the questionnaire are discussed in greater detail in the

School Participation module.

The following text elements is generic (and upper case) in the questionnaire in Appendix 2, and must remain generic (and upper case) in the context-specific version of the questionnaire. It can only filled in by the interviewer at the time of the interview:

> **'NAME' (of household member about whom)**: The interviewer will get this information off of the household roster. The interviewer need not, of course, actually print the name on the questionnaire script since the script will be used with reference to numerous household members. Instead, the interviewer should substitute in the name of the household member when they read the script aloud.

Guiding Language

Some of the modules recommended in this report make abrupt conceptual shifts that may overlooked by questionnaire respondents unless the survey interviewer explicitly points out these conceptual changes. For example, one question may ask about attendance during the current school year and another may ask about attendance during the previous school year; one question may ask about participation in formal-sector education and the next may ask about participation in non-formal-sector education. These shifts may appear to questionnaire developers, and may be missed altogether by survey respondents. If respondents miss these shifts, information gathered through the questions may be compromised; for example, respondents may respond to questions about the previous school thinking they are still being asking about the current school year. Clearly worded questions, in themselves, may not be enough to alert respondents to these conceptual shifts.

In order to ensure that respondents have a clear understanding of what they are being asked about at all times, we recommend adding guiding language to the questionnaire at points where conceptual changes occur. The EdData questionnaires follow this practice, and our recommendations for guiding language follow the example set by EdData. For example, between the participation and efficiency modules, where the questionnaire shifts from asking about the current school year to asking about the previous school year, the following language is added to the questionnaire:

> Up until now we have been talking about NAME and his/her education during the YYYY [CURRENT] school year. Now I would like to ask about NAME's schooling one year ago. In other words, I would like to ask about NAME's schooling during the YYYY-1 school year.

The recommendations in this report integrate guiding language wherever needed. However, because much guiding language is inserted at transitions between modules, and because it is anticipated that questionnaire developers may remove or alter modules or questionnaires in order to accommodate their particular needs, questionnaire developers should also pay attention to this issue and ensure that proper guidance is provided wherever there is the potential that survey respondents may become confused or overlook a transition.

MODULES

Modules are organized into three broad categories based on the relative importance of the indicators generated by each. This hierarchy is intended as rough guidance for questionnaire designers trying to decide which education modules to include in their questionnaire.

EPDC organized the modules into three categories:

Core modules: We strongly recommend that these four modules – Learning, Attainment, Participation, and Efficiency – be included in all household surveys. The core modules have an elegant design that makes it possible to calculate a large number of internationally relevant indicators using a relatively small number of questions. Learning, Attainment, and Participation modules are highlighted by the United Nations as core topics for censuses.[8] Though efficiency topics are not highlighted by the UN, we believe they should be considered a core topic for household surveys because of the large number of internationally-monitored indicators that can be calculated with the addition when our recommended efficiency module is included. The indicators produced provide a fundamental understanding of the education situation in a country. The core modules are consistent with a number of past DHS and MICS surveys, and would build on a body of data that is relatively comparable over time and across countries.

Recommended Auxiliary Modules: We recommend that these modules be included in every household survey if resources permit. These modules build on the core modules and address

[8] Principles and Recommendations for Population and Housing Censuses. Revision 2. United Nations Statistics Division. New York, 2008. (p113)

fundamental issues such as reasons for non-attendance among school-aged children, and the household financial burden of sending a child to school. These modules are not foundational in the same way that the core modules are, but are of great potential interest to education policymakers because they build an understanding of the social and economic forces driving the literacy, attainment, attendance, and efficiency indicators measured through the core modules.

Optional Auxiliary Modules: These modules may or may not be of interest to policymakers depending on the policy priorities of the country where the survey takes place. Questionnaire designers should only consider including these modules in their questionnaire if the issues covered in the modules are of particular policy relevance in their policy context.

CORE MODULES

It is highly recommended that the four education modules discussed in this section –

Learning, Educational Attainment,

School Participation, and Educational Efficiency – be treated as a core set of education questions that should be included in every household survey questionnaire. The simple structure of these questions reduces the likelihood of errors or misunderstanding during the interview process, and taken as a whole, the questions included in these modules are an elegant and efficient way to gather the information needed to calculate more than half of the indicators recommended through this report, many of which are indicators commonly used to track progress towards international goals education goals. Moreover, the education indicators generated through these modules are some of those that are most often used by national and international policymakers in the education policy realm.

There is strong precedence for the use of these questions in the education modules of well established lines household surveys: the exact wording is used in the widely used MICS and DHS questionnaires, and questions with closely similar wording can be found in the LSMS, CWIQ and other series. An advantage of this is that data gathered through this core module will be compatible with data collected through MICS and DHS, and to a lesser extent, LSMS and CWIQ. The core modules also serve as an excellent foundation for the Auxiliary Modules suggested for use in household surveys.

It is worth noting that these core modules are an excellent way to gather information on educational attainment, school participation, and educational efficiency as they relate to the formal education sector. The core modules are not appropriate for gathering information related to the informal sector. A discussion of gathering informal sector data can be found in the Non Formal Education module.

Learning

Learning indicators are used as measures of the extent to which the population members have obtained some of the basic functional skills that can be gained through education. Two skills that can be measured through household surveys, Literacy and Numeracy are discussed here. These indicators are of value to policy makers both as a measure of the human capital accumulation within the population, and as a tool for targeting education programs. Literacy rates in particular are of high interest to the international community and are specifically cited as indications of progress towards the Education for All and Millennium Development Goals.

Literacy

Definition

The United Nations defines literacy as follows:

A literate person is one who can both read and write a short, simple statement on his or her everyday life. An illiterate person is one who cannot, with understanding, both read and write such a statement. Hence, a person capable of reading and writing only figures and his or her own name should be considered illiterate, as should a person who can read but not write as well as one who can read and write only a ritual phrase that has been memorized. However, new understanding referring to a range of levels, of domains of application, and of functionality is now widely accepted.[9]

Measuring Literacy

For our purposes in household surveys, household

[9] Principles and Recommendations for Population and Hous
Statistics Division. New York, 2008. Pg 147.

Figure 4 Literacy results from six household surveys using both self-assessment and direct-assessment instruments.*

The Morocco Literacy Measurement Survey 1991-1992 asked respondents to report their literacy and then assessed literacy through direct assessment. Found that 45.5% of respondents reported literacy while 33.1% were assessed to have "at least basic competence" and 23.8% demonstrated "full independent competence.'

An unnamed 1998 survey in conducted in rural areas of Bangladesh found that among respondents who self-reported that they were literate, 83% demonstrated "minimal competence" in a direct assessment.

In the 2001 DHS Ethiopia survey, Women aged 15-49 had a literacy rate of 17.7% by self-assessment, and a rate of 10.5% by direct assessment. Men of the same ages were self-assessed at 44% and direct-assessed at 29.2%.

In the 2001 Nicaragua DHS, women aged 15-49 self-assessed literacy rate of 73.7%, and a direct assessment literacy rate of 67%. Men's literacy was not measured.

In the 1993 Tanzania Human Resource Development Survey, 98% of children aged 7-15 with some education self-reported that they were literate, but 72% were able to correctly answer 2 of 6 multiple-choice questions about a passage and 1.7% were able to correctly answer 5 of the 6 questions.

In the Ghana GSS/OED 2003 Household and School Survey, respondents underestimated their literacy rates in self assessments: Females self-reported at 41.6% and tested at 47.9%; males came in at 56% and 62% respectively.

surveys can gather information on literacy in one of two ways: Self-Reporting refers to a practice of asking the survey respondent whether they and/or others in the household are literate; Direct Assessment refers to the practice of testing household members in order to determine whether (and the degree to which) they are literate. Some complex direct assessments can involve hours of testing to gather a nuanced understanding of various literacy skills, but when direct assessments are administered in household survey setting, most simply ask each household member to read a simple sentence out loud.[10]

For the designer of a household survey, there are considerable trade-offs to consider when deciding whether to measure literacy through self-assessment or direct assessment. Some literacy measurement specialists believe self-assessed literacy to be inaccurate and a poor measure of a person's actual ability to read and write. They contend that respondents may feel pressures to exaggerate their literacy ability, and argue that, when one household member is asked to report the literacy of other household members, respondents may have a poor idea of this information.

There is evidence that self-assessed literacy rates tend to be higher than direct-assessed literacy rates. A 2006 EFA Global Monitoring Report background paper looks at six household surveys that measured both self-assessed literacy and direct-assessed literacy for the same household respondents and finds that, self-assessed rates can be as much as 15 percentage points higher than directly-assessed rates (Figure 4). It is worth noting that in a Ghana 2003 survey, self-assessment literacy rates were actually *lower* than direct-assessment literacy rates. The study concludes that the household assessments used in most household surveys are largely inappropriate:

> The large discrepancies between household assessments and direct assessments … underline the great potential for bias in household assessments and the need to move away from household assessments and towards direct assessments in developing country household surveys (Schaffner, 44)[11].

The 2008 Global Monitoring Report (UNESCO 2007) compares reported literacy rates against tested literacy rates for Kenya and notes that reported literacy rates are 13-15 percentage points higher than tested rates. The report goes on to caution that "conventional literacy data tend in fact to over-estimate literacy levels and should be interpreted with caution."

Direct-assessments of literacy also have notable shortcomings. Simple direct-assessments such as those administered through a household survey may give a reasonably accurate measure of low-level reading skills (the ability to de-code a text), but they miss the higher-level reading skills such as comprehension

[10] Much more information on the measurement of literacy can be found at the UIS website (http://uis.unesco.org).

[11] Schaffner, Julie. Measuring literacy in developing country household surveys: issues and evidence Background paper for the Education for All Global Monitoring Report 2006: Literacy for Life; Publ: 2005. http://unesdoc.unesco.org/images/0014/001462/146285e.pdf

and interpretation that could be captured through a complex assessment (Schaffner, 2). Perhaps more importantly, direct-assessments are much more challenging and costly to administer. Whereas self-assessment involves little more than asking a household member whether or not others are literate, the direct assessment involves tracking down and testing household members who may be away working, at school, or travelling; this exercise may require a considerable outlay of additional resources, especially since the rest of the education modules do not require direct interaction with each household member. Moreover, direct-assessment involves researching and testing appropriate test-phrases for each survey, producing additional survey materials, and training interviewers on how to administer the test. Another shortcoming of the direct assessment is that it does not address the full definition of literacy, which includes both the ability to read *and* the ability to write.

Despite the concerns that have been raised about the veracity of self-reported literacy data, the UNSD, in its recommendations for 2010 censuses, voices a preference for the direct-assessment of literacy, but validates the use of self-assessment when direct assessment is not practical:

> UNESCO recommends that literacy tests should be administered, in order to verify, as well as improve, the quality of literacy data. Nevertheless, administering a literacy test to all household members in the course of enumeration may prove impractical and affect participation, therefore limiting the utility of the results. Countries have regularly used simple self-assessment questions within a census to provide an indication of literacy rates at the small area level. An evaluation of the quality of statistics should be provided with census statistics on literacy.

Language of Literacy

In its principles and recommendations, the UN notes that:

> The notion of literacy applies to any language insofar as it exists in written form. In multilingual countries, the census questionnaire may query the languages in which a person can read and write. Such information can be essential for the determination of educational policy. This item would, therefore, be a useful additional subject of inquiry.[12]

In order to fully measure literacy, survey developers should be prepared to gather information on it in *any* language. This is relatively easy when using self-assessment, since survey developers can simply specify that they are asking about literacy in any language. Questionnaire developers may wish to go a step further as ask which languages the household member is able read and write in. Addressing the issue of language is more challenging in a direct-assessment since it requires interviewers to be ready to assess literacy in any number of languages that a household member may be able to read. This, in turn requires preparing comparable assessment materials in a number of languages and scripts, and may require testing in languages that the interviewers themselves are not able to read.

[12] Principles and Recommendations for Population and Housing Censuses. Revision 2. United Nations Statistics Division. New York, 2008. Pg 148.

Precedence

Of the 30 surveys analyzed in our 2008 report, 26 gathered information about literacy, and all that did so gathered self-assessed literacy. Recent DHS and MICS surveys test for reading ability using flashcards with a simple statement printed on them. EdData surveys test for both reading ability and a capacity for simple math. Though the definition of literacy includes the ability to both read and write, we have not seen an example of a survey that tests writing ability .

Out of the 26 surveys that gathered self-assessed literacy data, variations on the basic question are as follows:

Are you literate?	1 survey
Can you read or write?	2 surveys
Can you read and write?	15 surveys
Can you read? Can you write? (as separate questions)	8 surveys

The first question should not be used because it relies on the respondent to define literacy. The second should not be used because it misses the UN definition of literacy. The third and fourth options come closest to addressing this definition.

A subset of these questions operationalize literacy a little more specifically by asking questions such as "Can you read a one page letter in English?" (Malawi IHS), or "Can NAME read and write a simple phrase in French or English?" (Cameroon ECAM). This is an idea that makes sense since it helps communicate the basic level of literacy ability that is being sought out.

Out of the 26 questionnaires that queried self-assessed literacy, 19 made no mention of the specific language of literacy; 7 asked specifically about one of more languages. We recommend that survey designers consider asking about literacy "in any language" in order to clarify the information sought.

Filtering

DHS, MICS, and four of the 26 surveys from our 2008 study that queried literacy include some sort of filter so that household members above a certain educational attainment (usually more than a complete primary education) are not queried or tested on their literacy status. The rationale for this is an assumption that individuals who have reached a particular attainment level would not have been able to do so without having learned to read and therefore can be assumed to be literate. However, the UN[13] cautions (and we agree) that we should not assume a relationship between attainment, attendance, and literacy. We recommend filtering these questions by nothing but the age of the household member.

[13] IBID.

The UN definition of adult literacy applies to ages 15+, and youth literacy applies to ages 15-24. Based on these definitions, literacy should be queried for all individuals aged 15 or higher. It would be ideal if literacy were queried for ages 10+ because literacy rates for active pupils could be applied as a measure of progress in the school system.

DHS and MICS surveys query literacy for female household members between the ages of 15 and 49. Some DHS surveys query literacy for male household members between the ages of 15 and 54 or 59, depending on the survey. These are ranges are determined by the age range for the sub-sample of male and female household members who are asked to respond directly to the *individual* questionnaires that include the literacy tests. Among the surveys studied in by EPDC 2008, 20 used the same age filter they had used for the attainment question (usually ages 5+), and 6 used a different set of age filters (3 queried ages 10+; 3 queried ages 15+); all of these surveys covered 'reported' literacy rather than 'tested' literacy.

For Self-Assessed literacy, we recommend asking about literacy rates for all household members ages 3+. It is true that it is highly unlikely that many 3 year olds are able to read, but we suggest this age filter because we are recommending the same age filter for the attainment questions that will follow, and keeping the same filtering simplifies the flow of the questionnaire. Moreover, beginning with this low age will ensure that we capture information on learning for young children as they progress into and through the school system.

For direct-assessed literacy tests, we expect that many directly administered literacy test will be added to some other module that involves direct interviews of household members. In cases like this, survey may simply follow the filtering recommendations for this other module rather than the preferred filtering for a literacy module. Designers should bear in mind that if the test is not administered to ages 15+, then the data collected will not conform to the UN definition for literacy. Designers unable to test a sample representative of ages 15 + might choose to test for ages 15-49, 54, or 59 so that their results are comparable with the body of results from DHS and MICS literacy testing. To the extent that survey planners *do* have control over the filtering for the literacy tests, we recommend:

> Preferred standard: Ages 10+ *, do not filter by attendance or attainment*

Recommendations

Use direct assessment measures of literacy and numeracy

Evidence shows that direct assessments of literacy are more reliable than self-assessments. We strongly recommend that survey designers make every effort to use direct assessment rather than self-assessment to measure literacy and numeracy. When direct assessment is not possible, self-assessments can be accepted as a less-desirable substitute.

Direct Assessment of Literacy

The prompt for a direct-assessment of literacy should be as follows:

> *Now I would like you to read out loud as much of this sentence as you can.*

Developing the precise sentences and/or math problems to be used in each survey requires some understanding of local norms as well as expertise in developing test questions. Due to these constraints, specific sentences will not be recommended in this report.

Response options should be:

1 Cannot read at all
2 Able to read only parts of the sentence
3 Able to read the whole sentence
4 No card with the required language
5 Blind/Visually impaired

Self-Assessment of Literacy

When a direct assessment is not possible and self-assessment must be used instead, we recommend the following:

> *Can* NAME *read and write a simple phrase in any language?*

If survey designers are interested in gathering data on literacy in various languages, a follow-up question is recommended:

> *In what languages?* [select more than one if applicable]

The response menu to this question should list all likely languages, as well as 'other'.

Numeracy

Definition

We were unable to locate a single, standard, international definition of numeracy. In the glossary of the 2006 EFA Global Monitoring Report, numeracy is defined as:

> Usually, the ability to add, subtract, multiply and divide. More broadly, it means the knowledge and skills required to effectively manage and respond to mathematical demands posed by diverse situations, involving objects, pictures, numbers, symbols, formulas, diagrams, maps, graphs, tables, and text. Encompassing the ability to order and sort, count, estimate, compute, measure, and follow a model, it involves responding to information about mathematical ideas that may be represented in a range of ways.[14]

In an ALL paper, numeracy is defined as follows:

> Numeracy is the knowledge and skills required to effectively manage and respond to the

[14] 2006 EFA Global Monitoring Report: Literacy for Life. 2005. UNESCO. Paris.

> mathematical demands of diverse situations…Numerate behavior is observed when people manage a situation or solve a problem in a real context; it involves responding to information about mathematical ideas that may be represented in a range of ways; it requires the activation of a range of enabling knowledge, factors, and processes.[15]

Out of this, we can take a narrow definition of numeracy as "the ability to perform the calculations of addition, subtraction, multiplication and division," and a broader definition of numeracy as "the knowledge and skills required to effectively manage and respond to mathematical demands posed by diverse situations." The broader definition is ultimately more meaningful, but cannot be measured without the use of complex instruments that are inappropriate for household surveys. The simple definition could be measured through a simple direct assessment (such as asking each test respondent to perform one of each type of calculation), or through self assessment (such as asking each survey respondent "Is NAME able to add, subtract, multiply and divide"). The tradeoffs associated with direct-assessment and self-assessment are discussed in the Literacy section of this report.

A key difference between measuring literacy and measuring numeracy is that while literacy skills can only be exercised with reference to written text (i.e. one cannot 'decode' text that has not been written), numeracy skills need not be exercised with reference to written numbers (i.e. one can perform written calculations *or* one can perform mental arithmetic without writing anything down) (2006 Global Monitoring Report, Box 6.2, pg 149.) We interpret the GMR glossary definition of numeracy to acknowledge this in the phrase "it involves responding to information about mathematical idea that may be represented in *a range of ways*." Based on this point, we believe it is important that, in contexts where literacy is not universal, any assessment of numeracy is designed to measure an individual's ability to reason mathematically without regard for their ability to read or write mathematical symbols.

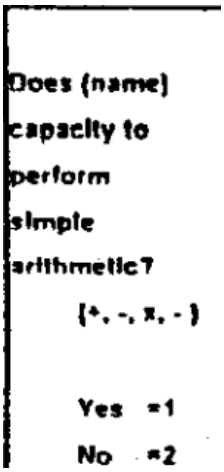
16
[NOM] sait-il/elle effectuer des opérations de calcul par écrit ?
OUI.....1 **NON .....2** *Inscrivez le code "1" seulement si la personne est capable d'effectuer les quatre types d'opérations (additions, soustractions, divisions et, multiplications).*

Precedence

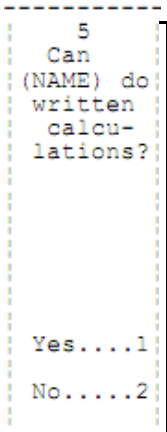
There is some precedence for the measurement of numeracy in household surveys. Six of the 30 questionnaires investigated for EPDC's 2007 report include a self-assessment query on numeracy. The EdData questionnaire uses a direct-assessment of numeracy.

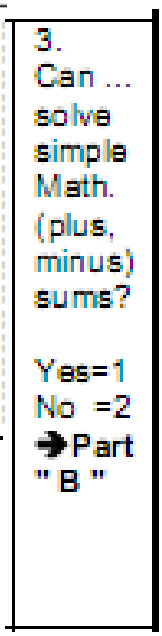
← EAM Djibouti 1996 asks about the ability to perform written calculations and specifies that a household member is numerate if they are able to perform four types of calculations: addition, subtraction, multiplication, and division. The question is asked regarding all household members age 5 and older.

[15] Measuring Adult Numeracy and Life Skills: New Frameworks for Assessment. 2005. Statistics Canada. Ottawa.



← Ethiopia WMS 2000 does not specify whether it is asking about written calculations; it implies that the household member must be able to perform all four calculation types.

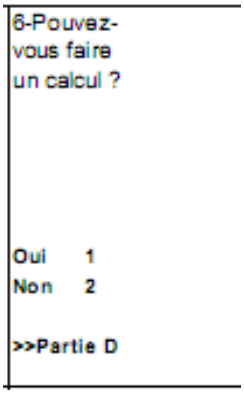


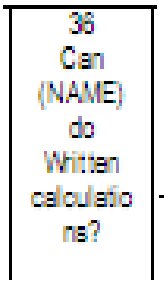


← Ghana LSS 1998 asks about written calculations but does not specify calculation types.

← Pakistan without

IHS 2001 asks about addition and subtraction specifying whether or not they must be written.





← Madagascar 2001 EAM does not specify whether or not calculations are written, or what functions should be considered.

← Sierra Leone IHS 2003 asks about written calculations, but does not specify the type of calculation.

In the Zambia 2002 EdData questionnaire, the interviewer is instructed to say "Now I would like you to add these numbers for me" and then hand the respondent a card with a math equation on it. Because the math equation is written on paper, this test measures a respondents able to decode written text of the equation and then their ability to perform the actual math calculation.

Filtering

None of the questionnaires filter by any criteria other than age. Four of the self-assessment questions ask about all household members ages 5 and higher; one each asks about ages 4 and higher and 10 and higher. The EdData direct assessment addresses only ages 6-11.

Recommendations for consideration:

Recommend for consideration: Establish a threshold of ability above which a respondent can be

considered 'numerate' and below which a respondent can be considered 'innumerate.'

For reasons discussed in greater detail in our treatment of literacy, direct-assessments of numeracy are highly recommended over self-assessment wherever possible. In anticipation that some questionnaires will use direct-assessment and others will use self-assessment to measure numeracy, we strongly recommend the establishment of a simple threshold for a binomial ('Numerate' or 'Innumerate') measure of numeracy. As discussed in the Literacy portion of this report, the UNSD established the threshold for literacy as the ability to read and write "*a short, simple statement on his or her everyday life*" (threshold language is presented in italics). A similar threshold for numeracy would be useful for establishing a uniform measure of numeracy.

Recommended for consideration: Direct Assessment

Only one of the questionnaires investigated in this study included a direct-assessment of numeracy, but this instrument leaves room for improvement because it requires respondents to be able to read and decode a written mathematical equation in order to perform the test. Using this instrument, individuals who possess math skills but are illiterate would likely be incorrectly recorded a innumerate.

We recommend for consideration that, instead of a text-based direct assessment, numeracy skills be tested though a non-text based assessment similar to the math problem presented in Figure 5. For this math problem, which is a sample of the numeracy assessment questions used in the Adult Literacy and Lifestyles (ALL) Survey, the question "In total, how many bottles are in the two full cases?" can be read aloud by the interviewer, and the respondent can give their answer orally. A set of non-text-based numeracy assessment questions with a comparable level of difficulty would need to be developed so that in the same household multiple people could be tested. The specific problems used would need to be tested in each context.

Figure 5 Illustrative example from the Adult Literacy and Lifeskills Survey (ALL) of a numeracy assessment that does not require the ability to read.*

Question:

"In total, how many bottles are in the two full cases?"**

* http://nces.ed.gov/surveys/all/Items.asp

** Answer: 48

Recommended for consideration: Self-Assessment

For surveys where the direct assessment of numeracy is not possible, questionnaire designers may use a self-assessment question instead. We recommend for consideration that, rather than specify the narrow conception of numeracy as the ability to perform calculations on paper, this question be worded to ask broadly about the ability to perform calculation on paper *or* mentally. We also recommend for consideration that this question be coordinated with the recommended threshold for numeracy and with

the direct-assessment of numeracy so that all three are calibrated to measure the same calculation skills. A self assessment question that is coordinated with the threshold definition and direct-assessment instrument proposed here could be worded as follows:

> *Is NAME able to solve everyday problems that involve adding or multiplying numbers either on paper or in his/her head?*

Educational Attainment

Educational attainment indicators provide a measure of the accumulation of human capital in a population by communicating information about the highest level of formal education attained by household members. Measures of educational attainment are of interest to policymakers because attainment has been found to be correlated with many measures of development and quality of life, such as child and adult health and mortality, income and occupation, urban migration, democracy, and economic productivity. The questions recommended in this module are widely used and should be considered a core component of any education module.

Educational Attainment indicators do not convey information about other potentially important sources of human capital, such as apprenticeships and literacy training, that are obtained outside of the formal school system. A separate Non Formal Education module can be used to gather information on human capital accumulated from certain sources outside of the formal school system.

Definition

Educational Attainment is currently defined by the United Nations as "the highest grade completed within the most advanced level attended in the educational system of the country where the education was received."[16]

As part of an ongoing review of ISCED, a definition of "educational attainment" is currently being drafted and it is anticipated that the new definition will focus on the highest level completed rather than the highest level. In order to accommodate this new definition when it is made public, the Educational Attainment module presented here may need to be revised to include a question similar to that which is discussed in the Auxiliary Attainment: Highest Diploma Earned section of this report.

Precedence

MICS & DHS: Three questions:

> *Everybody above age 5*
>
> Has NAME ever attended school?

[16] Principles and Recommendations for Population and Housing Censuses. Revision 2. United Nations Statistics Division. New York, 2008. Pg 150.

If have ever attended school:

What is the highest level NAME has attended?

What is the highest grade NAME completed at that level?

The wording of these questions directly addresses the UN definition of attainment. Highest level attended and highest grade completed are recorded as two separate variables. An advantage of this practice is that it makes it possible to identify respondents who, for example, attended the first grade of secondary but never completed that grade (Highest level attended: *Secondary*; highest grade completed, *None*), though this information introduces an element of uncertainty because we are forced to assume that such a respondent did actually complete the last grade of the previous level. The questions do not specify what it means to 'complete' a grade. This approach is not suited for school levels that have parallel grade sequences (for example technical secondary and academic secondary tracks) because it does not allow for differentiation between tracks. Moreover, because grade/level combinations are not pre-coded, there is no structure in place to prevent the recording of grade/level combinations that do not exist in the school system.

One possible justification for the coding system used by DHS and MICS is that it allows for flexibility in recording grades/level combinations that may exist in some areas or types of schools, but not in others. In practice however, it is impossible for the data processor to distinguish between valid aberrations from the norm and erroneous responses.

LSMS: Two questions for one group, two questions for another group

Everybody above age 5

Have you ever attended school?

Are you currently enrolled in school?

If has ever attended and not currently attending school:

What is the highest grade you have completed in school?

If has ever attended and currently attending school:

In what grade are you currently enrolled in school?

This setup reduces the number of questions asked by the interviewer because the attainment question is addressed to only those who have attended school but are not currently attending. Attainment data for those currently attending school must be derived from an assumed relationship between attainment and current attendance. We can probably assume that an attendee's most recently completed grade is the one immediately preceding the grade they are currently attending, but that assumption does not hold in all cases. The UN recommends that we assume no relationship between Attainment, Attendance, Literacy.[17] Unlike DHS/MICS, LSMS surveys record attainment data in a single variable that lists all level/grade combinations that are (presumably) valid in the education system. This reduces the likelihood that non-

[17] IBID

existent grade/level combinations will be recorded, and makes it easier to record school levels that do not have well defined grade sequences. While we lose information about respondents who attended secondary without completing a grade at that level, we do have a more precise understanding of the last grade completed at a specific level.

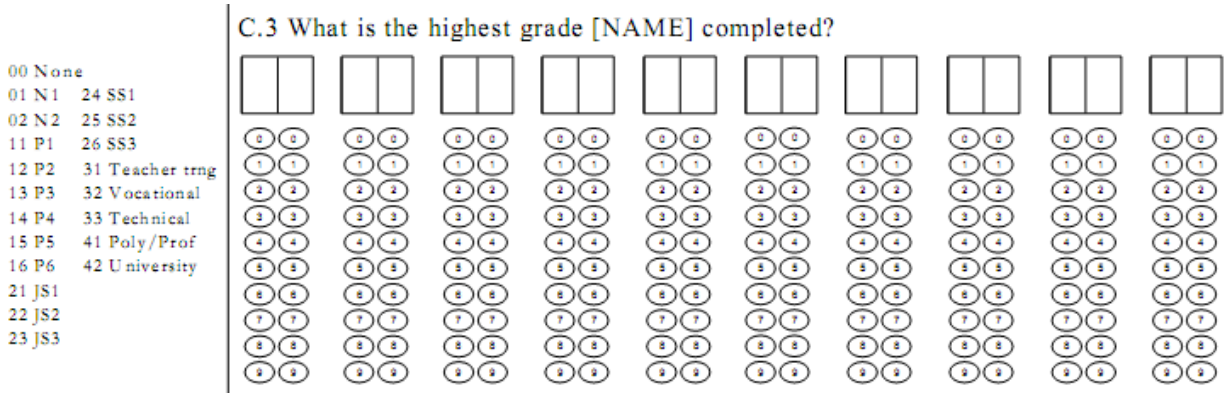
CWIQ: Two Questions

All household members

Has NAME ever attended school?

If have ever attended:

What is the highest grade that NAME has completed?



The attainment question is asked of all household members regardless of age or any other criteria. Grade and level are recorded in a single variable from a menu of options. If the menu of response options is well crafted and documented so that it is clear how each menu item corresponds to a particular school level or grade within each school level that can be broken into grades, we believe this approach will reduce opportunities for coding or response error. It also makes it easier to accommodate levels with grades and levels without grades within a single question structure.

The straightforward questions directly correspond to the indicator in question, so there is no need for assumptions when determining a respondents attainment. Because the same questions are asked of all respondents, there is little need to worry that the question was asked of the proper sub-population.

Recommendations regarding filtering:

Among the surveys investigated in EPDC's 2008 report, two surveys asked questions for ages 3+, three for 4+, 15 for 5+, and one for 6+. Some surveys did not filter by age at all. The general standard has been to ask questions of people ages 5+, but in order to better accommodate questions concerning pre-primary attendance, we strongly recommend a filter for ages 3+ instead. According to UIS, the official starting

age for pre-primary education is either 3 or 4 in most countries, so including ages 3+ would make it possible to calculate NAR for pre-primary.

Attainment questions should *not* be filtered by any criteria other than age because we cannot safely assume any relationship between attainment and literacy, attendance, or any other variable.

Recommendations regarding School Level & Grade Choices:

We recommend pre-coding the list of response options for the question 'What is the highest grade that NAME has completed?' Response items on the list should include specific grades with level where the sequence of grades is clearly defined (in most countries, for example, the number of grades in primary school is uniform across the country), but avoid listing specific grades for levels where grades are not meaningful (for example, a child might attend two years of pre-primary, but the contents of the first and second years are not differentiated; or in tertiary education, course requirements for the first year of college are not differentiated from requirements for the second year).

Figure 6: Illustrative School Level / Grade coding sequence

School Level Codes (first digit):

0	Pre-primary
1	Primary
2	Secondary (or lower secondary if applicable)
3	Upper Secondary (if applicable)
4	Post-secondary Non-tertiary
5	Tertiary

School Grade Codes (second digit):

0	For school levels where it is not possible to differentiate between grades
1-9	For school levels where it *is* possible to differentiate between grades, or post-secondary levels where it is not necessary to distinguish between years.

Sample Combined School Level & Grade Code sequence using the above recommendations:
(In a hypothetical school system with Lower Secondary and two parallel upper-secondary tracks)

01	School-based pre-primary center
02	Informal pre-school
11	Primary, Grade 1
12	Primary, Grade 2
13	Primary, Grade 3
14	Primary, Grade 4
15	Primary, Grade 5
16	Primary, Grade 6
21	Lower Secondary, Grade 1
22	Lower Secondary, Grade 2
23	Lower Secondary, Grade 3
31	Upper Secondary – General, Grade 1
32	Upper Secondary – General, Grade 2
33	Upper Secondary – General, Grade 3
34	Upper Secondary – Technical, Grade 1
35	Upper Secondary – Technical, Grade 2
41	Post-Secondary – Teacher Training
42	Post-Secondary – Technical
43	Post-Secondary – Vocational
50	University
98	Don't know
99	None

It is imperative that grades and levels are listed using names and references that will be easily understood by questionnaire respondents rather than names and references that might be more familiar to policymakers. For example, if secondary grades are commonly referred to as forms, the word 'form' should be used in the questionnaire. Grades and levels can always be converted to other formats during data processing after interviews have been concluded.

The two-digit pre-coding sequence used in CWIQ and LSMS questionnaires is an elegant way to indicate school level and school grade in one number. In this illustrative sequence, the first digit corresponds to the school level. For the Primary and Secondary levels, specific grades should be measured along with levels;[18] these grades are represented by the second digit in the coding sequence for those levels. The second digit of this two-digit coding sequence should proceed sequentially when possible, with exceptions for parallel tracks being carefully documented and labeled. In the sample code presented in Figure 6, for example, Upper Secondary is divided into two parallel tracks: the three grades of General Upper Secondary correspond to codes 31, 32, 33 respectively; the two grades of Technical Upper Secondary correspond to codes 34 and 35 respectively. At the post-secondary non-tertiary level, the sample codes correspond to education sequences that may last longer than one year.

For the Pre-Primary, Post-Secondary, and Tertiary school levels, individual grades need not be recorded since there is not always a clearly defined progression of grades at these levels.[19] At these levels, second-digits could be used to differentiate between fields of study if relevant. This is illustrated through the codes for post-secondary education (41-43) in the sample code presented in Figure 6.

Because countries' education systems vary in organization and duration, the grade/level codes used in most questionnaires will not correspond precisely with what is presented in Figure 6. It is especially important that School Level/Grade codes are carefully documented so they can be properly analyzed by data users who were not involved in the survey design. Because nationally-defined school levels may not correspond with UN-defined (ISCED) definitions of school levels, it is also important that the survey documentation include a conversion table explaining how nationally-defined school levels can be converted to ISCED levels.

In countries where the education system has been reformed over time, the educational attainment levels of some respondents may not correspond with contemporary levels and grades. In these cases, the questionnaire should include a conversion table similar to that presented in Figure 7 for interviewers to use to convert older attainment levels to equivalent contemporary attainment levels.

[18] UOE Data Collection Manual Vol 1.: Concepts, definitions and classifications. 2009. Montreal/Paris/Luxembourg, UNESCO-UIS/OECD/EUROSTAT., pg 36.

[19] UOE Data Collection Manual Vol 1.: Concepts, definitions and classifications. 2009. Montreal/Paris/Luxembourg, UNESCO-UIS/OECD/EUROSTAT.

Pre-Primary, ECCE or both?

While most of the grades and levels of education mentioned above should be familiar to most survey planers, it is worth noting the distinction between **Pre-Primary** education and **Early Childhood Care and Education (ECCE)**[20]. Pre-primary education is defined as:

Programmes at the initial stage of organized instruction, primarily designed to introduce very young children, usually from age 3, to a school-type environment, and provide a bridge between the home and a school. Upon completion of these programmes, children continue their education at ISCED 1 (primary education). (UIS online Glossary)

UIS defines ECCE as a broader term that includes both Pre-primary and other non-formal types of education:

> Programmes that, in addition to providing children with care, offer a structured and purposeful set of learning activities either in a formal institution (pre-primary or ISCED 0) or as part of a non-formal child development programme. ECCE programmes are normally designed for children from age three and include organized learning activities that constitute on average the equivalent of at least two hours per day and 100 days per year.

Following the logic of these UIS definitions, ECCE is a broader term that encompasses both formal, school-based Pre-primary education, and non-formal non-school-based forms of organized early childhood care. It would be nice to able to distinguish between these types of early-childhood care (the Global Monitoring report publishes attendance rates for both pre-primary and ECCE), so we recommend distinguishing between formal and non-formal types of early childhood care in the list of response categories. The 2006 Nepal DHS questionnaire sets a precedent by listing these categories as response options in their questionnaire,

- School based pre-primary centers
- Informal pre-school

Guiding Language

The EdData questionnaires include a short statement that is read by the interviewer before the education questions are started:

> "Now I would like to ask you some questions about NAME and his or her schooling. When we talk about schooling, it includes, preschool, primary school, secondary school, and higher levels of schooling."

[20] Early Childhood Care and Education (ECCE) is synonymous with Early Childhood Care and Development (ECCD).

This statement is helpful because it helps the respondent understand what they are being asked

Figure 7: Sample Educational Attainment Level Conversion Table

about. Without the guiding statement, respondents may wrongly assume, for example, that the interviewer is not interested in pre-school.

General education system for conversion		equivalent General educational levels								
		System under the French time	From 1945 until 1954			Complementary education (CE)system	Educational system in Northern Viet Nam			Current national educational system
			Free region		Temporarily occupied region		Prior to 1981	From Qu¶ng B×nh northward		
Level	Grade		1945-1950	1950-1954				1981-1986	1986-1989	
Primary School	1	Grade 5 (Cours enfantin)			Grade 5 primary school		Pre-school	Grade 1	Grade 1	Grade 1
	2	Grade 4 (Cours prÐparatoire)	Grade 4	Grade 1	Grade 4 primary school	Grade 1 CE	Grade 1	Grade 2	Grade 2	Grade 2
	3	Grade 3 (Cours elementaire)	Grade 3	Grade 2	Grade 3 primary school	Grade 2 CE	Grade 2	Grade 3	Grade 3	Grade 3
	4	Intermediate 1 (Moyen1) Intermediate 2 (Moyen2)	Grade 2	Grade 3	Grade 2 primary school	Grade 3 CE	Grade 3	Grade 4	Grade 4	Grade 4
	5	Upper intermediate (SupÐrieur) Certificate (Certificat)	Grade 1	Grade 4	Grade 1 primary school	Grade 4 CE	Grade 4	Grade 5	Grade 5	Grade 5
Lower Secondary School	6	First year (Premiłre annÐe)	First year		7th class Secondary school	Grade 5 CE			Grade 6	Grade 6
	7	Second year (Deuxiłme annÐe)	Second year	Grade 5	6th class Secondary school	Grade 6 CE	Grade 5	Grade 6	Grade 7	Grade 7
	8	Third year (Troisiłme annÐe)	Third year	Grade 6	5th class Secondary school	Grade 7 CE	Grade 6	Grade 7	Grade 8	Grade 8
	9	Fourth year-Diploma (Quatriłme annÐe - Dipl«me)	Fourth year	Grade 7	4th class Secondary school	Grade 7B CE	Grade 7			Grade 9
Upper Secondary School	10	First year	First year Specialisation	Grade 8	3rd class	Grade 8 CE	Grade 8	Grade 10	Grade 10	Grade 10
	11	First part of secondary school degree (BaccalaurÐat premiłre partie)	Second year Specialisation	Grade 9	2nd class Baccalaureate I	Grade 9 CE Grade 10A CE	Grade 9	Grade 11	Grade 11	Grade 11
	12	Second part, secondary school degre (BaccalaurÐat deuxiłme partie)	Third year Specialisation		1st class 2nd education degree	Grade10B CE	Grade 10	Grade 12	Grade 12	Grade 12

Recommendation regarding wording and sequence of questions:

The example set by the CWIQ questionnaire is simple, straightforward, and minimizes opportunities for error. By asking about grade *completed*, the question accurately addresses the UN definition of attainment because, by definition, the individual must have attended the level in question in order to have completed a grade at that level. Using a single variable to code grade and level attained (rather than a pair for grade and level) makes it easier to handle levels with individual grades and levels without individual grades.

Now I would like to ask you some questions about NAME and his or her schooling. When we talk about schooling, it includes, preschool, primary school, secondary school, and higher levels of schooling.

Household members Ages 3+

Has NAME ever attended school?

If has ever attended:

What is the highest grade that NAME has completed?

Recommended for consideration:

Though the pair of questions recommended above have a well-established precedent for inquiring about school levels at the primary level and higher, it seems that they may result in an under-counting of ECCE attainment and attendance if questionnaire respondents do not think of ECCE as part of 'school.' We recommend for consideration that the question 'Has NAME ever attended school?' be modified slightly to read 'Has NAME ever attended school (or any type of pre-school)?'

School Participation

School participation indicators provide measures of the extent to which segments of the population are participating in formal school programs. These indicators are especially useful to policymakers interested in gauging school participation and developing strategies to increase school participation.

Definition

According to the United Nations:

> School attendance is defined as attendance at any regular accredited educational institution or programme, public or private, for organized learning at any level of education at the time of the census or, if the census is taken during the vacation period at the end of the school year or during the last school year. According to the International Standard Classification of Education (ISCED) education is taken to comprise all deliberate and systematic activities designed to meet learning needs. Instruction in particular skills which is not part of the recognized educational structure of the country (for example in-service training in factories) is not normally considered "school attendance".[21]

Attendance or Enrollment?

> The 30 surveys EPDC analyzed in 2008 got addressed issue of attendance using eight variant wordings of a similar question:

- Are you currently attending school? (15 surveys)
- Are you currently enrolled? (5 surveys)
- Do you attend school this year? (3 surveys)

[21] Principles and Recommendations for Population and Housing Censuses. Revision 2. United Nations Statistics Division. New York, 2008. 2.150

- Did you attend school in the past 12 months? (2 surveys)
- Are you currently registered? (1 survey)
- Are you currently in school system? (1 survey)
- Do you attend any education institutes now? (1 survey)
- What is your usual activity? (includes: student. 2 surveys)

These questions would suggest that terms 'Attend' and 'Enroll' are interchangeable, but they are not. 'Enrollment' describes a pupils status on an official register or list of participants at a school 'Attendance' describes a pupils status as a regular participant or visitor at the school. We can expect household members to know whether a child regularly visits or participated in activities at the school, but cannot necessarily expect a household member to know whether the child's name is written on an official roll at the school. So it is better to consistently use the word *attend* when referring to school participation data gathered from a household survey.

Timeframe

Data collection for a census or survey can take place over a period of several months. As a result, there is a good chance that if a survey is planned without regard for the academic calendar, some households might be interviewed during the school year, some over a break, and in some cases, some over the following school year. Because education data that are collected over more than one school year can cause confusion in interpreting the results, it is highly recommended that survey planners do their utmost to time enumeration so it occurs during a single school year.

In cases when it is impossible to avoid having the enumeration period span more than one school year, questions can be worded to try to work around this issue. In this case the best way to avoid confusion is to identify the specific school year in question. For example "Did NAME attend school during the 2006-2007 school year". This is the practice with ECAM, MICS, and DHS. Questions that refer relative periods such as 'the most recently completed school year' or 'the current school year' can cause confusion when a survey spans more than one school year.

Frequency of Attendance

The internationally accepted definition of school attendance does not stipulate a proportion of time that a child must spend in school in order to be considered attending. Thus, measures of pupil absenteeism are not relevant for calculating attendance rates. Absenteeism can be an important policy issue, however, and is discussed in further detail in the Pupil Absenteeism module.

Filtering

The surveys analyzed used a variety of age filters for attendance questions – most of these ranged from a beginning age of 3-7 and an ending age of either 24 or 29. The United Nations recommends that

attendance data be collected for ages 5-29 at a minimum.[22] We recommend expanding bottom limit of the age filter range to age 3 in order to include capture information on pre-primary attendance and underage primary attendance by 3-4 year-olds. We also recommend increasing the upper-limit of the range to include respondents of age 30 in order to accommodate the collection of data for household members who were age 29 at the beginning of the school year but age 30 at the time that their age is reported in the interview (see explanation of adjusted ages below).

Adjusted Age for calculating attendance rates

Several participation indicators are calculated based on counts of the proportion of pupils attending school who fall within a particular age range (eg: the Net Intake Rate is the proportion of pupils attending grade 1 who are of the official age for grade 1). Special care must be used when calculating these indicators from household survey data because surveys are usually conducted over several months and the age of the same child may be recorded differently if their household was surveyed at the beginning or the end of the enumeration period. This can result in a distortion of age-sensitive indicators, especially in cases where the enumeration period ends many months after the beginning of the school year.[23]

There is no standard approach to dealing with this problem, though several alternatives exist. MICS tabulations adjust for age distortion in their participation rates by subtracting one year from all ages in surveys where the enumeration period took place a certain number of months after the beginning of the school year. The UOE data collection on education manual recommends adjusting all ages to a common reference point (usually January 1)[24]. Both of these methodologies are unsatisfactory because they adjust ages somewhat arbitrarily and the resulting adjusted ages not necessarily much more accurate than the unadjusted ages.

Most DHS surveys do not adjust for potential age distortion effects when calculating participation rates. For the 2005-2006 India DHS survey, attendance rates *were* calculated using ages adjusted to reflect the household members would have been expected to have at the month of the beginning of the 2005 school year. This was achieved by calculating the number of whole years that had elapsed from the year and month of the household members' birth to the year and month of the beginning of the 2005 school year; for household members whose month of birth was not known, a month was randomly imputed; household members whose year of birth could not be known were excluded from the calculations.

---

[22] Principles and Recommendations for Population and Housing Censuses. Revision 2. United Nations Statistics Division. New York, 2008.

[23] In an extreme example: If 100% of grade 1 pupils started school at the appropriate age, but their ages were recorded 11 months after the beginning of the school year, approximately 92% of pupil would have had a birthday in the mean time and would have their age recorded in the survey as one year higher than it was at the beginning of the year. As a result, the NIR would be falsely calculated to be 8% when it should have been calculated to be 100%.

[24] UOE Data Collection Manual Vol 1.: Concepts, definitions and classifications. 2009. Montreal/Paris/Luxembourg, UNESCO-UIS/OECD/EUROSTAT.

EPDC has developed a technique for adjusting ages to reflect the ages household members would have been expected to have at the beginning of the school year. This technique is based on fact that for each additional month that elapses between the beginning of the school year and month of the interview when a household member's age is recorded, there is an additional 1/12 probability that the household member has had a birthday. To adjust ages, we identified all of the household members whose ages were recorded in the first month following the beginning of the school year and subtracted one year from the age of a random selection of 1/12 of them; for household members whose ages were recorder during the second month after the beginning of the school year, we subtracted one year from the ages of a random selection of 2/12 of them, and so on.

We strongly recommend that, since participation rates are calculated with reference to the official entrance age for school, household members' ages be adjusted to reflect their age at the beginning of the school year when participation indicators are calculated (adjusted ages need not be used for literacy, attainment, or other non-participation indicators). We recommend using the approach used in adjust ages in the 2005-2006 India DHS:

Adjusted ages should be calculated using household members' month-of-birth information is available along with their age information. Only month-of-birth and year-of-birth are needed because, in most cases, the beginning of the school year is not accurate down the day of the month (in other words, not all school in the country open on the exact same day). Because of this we try to be accurate only to the month. The age adjustment takes place during data-processing, not at the time of the interview:

> If household members' date-of-birth information is available, adjusted age is calculated as number of whole years elapsed between the household members date of birth and the beginning of the school year. In cases where household members ages are not known, a month of birth value (1-12) can be randomly imputed and used for an adjusted age calculation. It is preferable to use actual month-of-birth information whenever possible; in surveys where month-of-birth is available for some household members but not for others, actual data should be used when it is available, and imputed values used for the remainder.[25]

School Level & Grade Choices:
As discussed in the Educational Attainment section, we recommend a pre-coded menu of response options. For each household survey, the response options offered in the attainment, participation, and efficiency modules should be identical.

Guiding Language

Now I would like to

Recommendation:

[25] This random-month imputation ultimately has the same effect as EPDC's probabilistic approach, with the added benefit that it is easier to explain.

We strongly advise a two question sequence used in DHS/MICS questionnaires, and, in keeping with the discussion on attainment, we recommend the CWIQ inspired single-variable menu-option response to the grade and level question:

> *If Yes to 'Has NAME ever attended school' and 3<=Age<=30*
>
> *Did NAME attend school at any time during the YYYY [current or most recent] school year?*
>
> *IF Yes*
>
> *What level and grade did NAME attend during this school year?*

Recommended for consideration:

As discussed in the Educational Attainment section, we believe there is a potential for ECCE to be under-reported if questionnaire respondents assume that these questions are asking about schooling at the primary level and higher. We recommend for consideration a slightly modified question that reads as follows:

> Did NAME attend school or pre-school at any time during the YYYY [current or most recent] school year?

Educational Efficiency

Educational efficiency indicators provide measures of pupils' flow through the formal school system. Among other things, efficiency indicators measure the extent to which pupils are progressing from grade to grade and from school level to school level, the extent to which pupils are repeating grades or dropping out early, and the extent to which pupils are reaching benchmarks in the school system. Some of the ways that educational efficiency indicators are used by policymakers include: to identify problems with the internal efficiency of the school system; to project pupil flows through the school system and allocate school resources accordingly; and as proxy measures for the quality of education provided to students.

Indicator Definitions

There are a great number of efficiency indicators used in the international community. Many of these indicators are related to three fundamental efficiency indicators that are listed here. Efficiency indicators are not defined the in UN document, but definitions are provided in the UIS glossary of education indicators.

> *Repetition Rate:* Proportion of pupils from a cohort enrolled in a given grade at a given school year who study in the same grade in the following school year. To calculate it, divide the number of repeaters in a given grade in school year t+1 by the number of pupils from the same cohort enrolled in the same grade in the previous school year t.

Promotion Rate: Proportion of pupils from a cohort enrolled in a given grade at a given school year who study in the next grade in the following school year. To calculate it, divide the number of new enrolments in a given grade in school year t+1 by the number of pupils from the same cohort enrolled in the preceding grade in the previous school year t.

Dropout Rate: Proportion of pupils from a cohort enrolled in a given grade at a given school year who are no longer enrolled in the following school year. It is calculated by subtracting the sum of promotion rate and repetition rate from 100 in the given school year. For cumulative dropout rate in primary education, it is calculated by subtracting the survival rate from 100 at a given grade (see survival rate).[26]

Precedence:

From EPDC 2008 report:

> Education flow ratios can be derived from data on grade level and attendance over two consecutive years, or from questions that directly ask about repetition, promotion, and dropout. Half (15) of the surveys ask some questions on flows, but the detail of the information gathered collected varies: only four surveys ask about grade and level attended during the previous school year – details enough to calculate promotion, dropout, and repetition; one survey asks about grade last year – to calculate repetition rate; six surveys ask about attendance last year but not grade – enough to calculate dropout rates only; and another four surveys ask about the number of years needed to complete a school level, and the number of times any grades were repeated over this period – from which proxies of repetition can be calculated.
>
> Table 15 shows the types of questions employed in the 30 surveys covered. The question sequence that covers attendance this year, last year, and grades in both years is presently used by the DHS and MICS surveys and leads to the most information in the most parsimonious way. All of the formulations that we found in the other surveys are inferior, in terms of efficiency of information collection, to these formulations.

[26] UIS Education Indicators Glossary

Table 15. Types of questions on school flows and which surveys employed them.

Type of coverage	Countries	Indicators that can be calculated
Attendance last year and grade last year	Cote d'Ivoire, Madagascar, Malawi, Ethiopia, DHS, MICS	Promotion, dropout, repetition rate
Attendance last year, in same grade?	Cameroon	Repetition rate
Attendance last year but not grade	Benin, Burkina Faso, Djibouti, Gabon, Mozambique, Niger	Dropout rate
Time needed to go through school – primary or secondary	Nepal, Malawi, Pakistan	Proxy for repetition
Ever repeated and frequency of repetition – primary or secondary	Sierra Leone,	Proxy for repetition
How many times repeated a grade in primary	Nigeria	Proxy for repetition

School Level & Grade Choices:
As discussed in the Educational Attainment section, we recommend a pre-coded menu of response options. The response options should be the same as those offered for the attainment and participation questions.

Recommend for Consideration: Guiding Language:

It is important to make an effort to ensure that interview respondents understand that this pair of questions is inquiring about the *previous* school year and not asking (redundantly) about the current school year a second time. The EdData questionnaires do this.

> Up until now, we have been talking about NAME and his/her schooling during the YYYY current school year. Now I would like to ask about NAME and his/her schooling one year ago – In other words, I would like to ask about NAME's schooling during the YYYY-1 school year.

Recommended:

> The sequence of questions used in DHS/MICS surveys is an efficient way to gather the information needed to precisely calculate all efficiency rates. An additional advantage of this wording is that it makes use of the recommended question sequence for Attendance.
>
> *If Yes to 'Has NAME ever attended school' and 3<=Age<=30*
>
> - Did NAME attend school at any time during the YYYY-1 [previous] school year?
>
> *IF Yes*
>
> - What level and grade did NAME attend during this school year?
> - Did NAME attend school at any time during the YYYY-1 [previous] school year?

IF Yes

What level and grade did NAME attend during that school year?

Brought together, the Educational Attainment, Attendance, and Efficiency modules would comprise a compact 6-question CORE module that would allow for the calculation of more than 25 of the most commonly collected and requested education indicators.

Auxiliary Modules

Auxiliary modules are modules that are considered to be of good potential value for inclusion in a household survey, but not so fundamentally necessary as to merit inclusion in every survey. Questionnaire designers may choose to include or omit these modules at their discretion.

School Characteristics

School Characteristics indicators are intended to provide policymakers with a richer understanding of the kind of education pupils are receiving. Indicators covered in this module might include: the percentages of pupils attending public, private, or religious schools, and the percentages of pupils attending schools that do or do not charge a particular fee. This indicators all communicate information about the education system, but the unit of observation is the pupil.

When considering questions in this category, questionnaire developers must keep in mind that questions are usually answered by some household member other than the child attending school. Thus, questions should focus on issues of general knowledge within the household (such as the number of hours that the child is absent from the home on a school day) and avoid questions that only the pupil could know the answer to (such as how often the teacher does not arrive on time).

The indicators covered in this module are not consistently treated as a module in other questionnaires, so we will treat each question or grouping of questions separately. We will then suggest a sequence for grouping the full set of questions together as a module.

% Attendance, by school administrative body

Depending on the characteristics of the school system, this indicator set could include '% Attendance, Public Education', '% Attendance, Private Education', '% Attendance, Religious education,' and/ or other categories.

Definition

The 2009 UOE data collection manual includes a lengthy discussion of the distinctions between public and private education institutions:

Educational institutions are classified as either ***public*** or ***private.*** Private institutions are further classified between ***government dependent private*** and ***independent private institutions***.

The classification between ***public*** and ***private*** is made according to whether a public agency or a private entity has the ultimate control over the institution. *Ultimate control* is decided with reference to who has the power to determine the general policies and activities of the institution and to appoint the officers managing the school. Ultimate control will usually also extend to the decision to open or close the institution. As many institutions are under the operational control of a governing body, the constitution of that body will also have a bearing on the classification. The public/private classification is not determined by source of funding, ownership of buildings, or regulating body.

> The terms "***government-dependent***" and "***independent***" refer only to the degree of a private institution's dependence on funding from government sources; they do not refer to the degree of government direction or regulation.
>
> A **government-dependent private** institution is one that either receives 50 per cent or more of its core funding from government agencies or one whose teaching personnel are paid by a government agency – either directly or through government.
>
> An **independent private** institution is one that receives less than 50 per cent of its core funding from government agencies and whose teaching personnel are not paid by a government agency.

Precedence

Of the 30 surveys analyzed in the EPDC 2008 Report, most (21) included a question on the type of school attended. The operative portions of the question used in each survey can be grouped into three as follows:

LSMS, HSES	"Is the school… Public or Private?"
CWIQ	"Who runs the school…?"
HIES, IHS, WMS	"What type of school is…?"
DHS, MICS	These surveys do not gather this information

The "Is the school… Public or Private?" questions appear to cut straight to the information international analysts are interested in, but they place the burden of determining whether a school is public or private on the shoulders of the respondent. Given the sophisticated distinction between public and private institutions and the likelihood that survey respondents may not be aware of the governance structure at a school, there is a danger that these questions allow leeway for incorrect responses.

The "Who runs the school…" and questions do not directly address the question of public vs. private attendance, but they have the advantage that they are worded to collect information that a respondent is more likely to be able to report correctly. These questions lend themselves to a response-menu type format that would allow respondents to select their response from a list of the school-type categories with

names that are meaningful to potential household respondents. These response categories would need to be explained in the survey documentation so they could be categorized into public/private categories during post-survey enumeration.

The "What type of school is…" questions have the same advantages as the "Who runs the school…" questions, but because they are worded more broadly, they could be used to gather more detailed information on the characteristics on the schools that may be of interest to policy planners. As an example, the Bangladesh HIES survey distinguishes between 'Private Bengali Medium' and 'Private English Medium' among other categories -- distinction that makes more sense when the question is worded in this way.

Recommendation

If currently attending school==YES

During the YYYY [CURRENT] school year, What type of school did NAME attend?

Recommend for consideration:

Three of the thirty surveys analyzed in 2008 include questions on the type of school attended during the current year *and* the type of school attended during the previous year. We believe this question can be valuable under certain circumstances and should be included if policymakers are interested in gathering information about pupils moving from one school type to another between school years.

In the Decisions on Education module**,** we recommend considering the addition of a set of questions to address the issue of student mobility (transfers from one school to another). If the student mobility questions are included in a questionnaire, then we recommend adding a question targeting household members who did not attend the same school both years in order to determine whether they also switched school types. The question would be structured similarly to the one above, but would ask about the school attended during previous year, and would be filtered to address only those household members who reported that they did not attend the same school both years.

When analyzed in conjunction with the information gathered through the Decisions on Education module, this question would reveal potentially useful information on the factors that lead households to select one school system over another.

During the YYYY-1 [PREVIOUS] school year, What type of school did NAME attend?

Response categories

Questionnaire developers should develop and test a set of response categories that are appropriate to their specific education system. Illustrative examples of school type categories used in existing surveys are provided on the following page. When devising a set of response categories, the following considerations should be kept in mind:

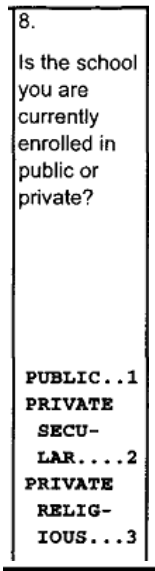
1. Response categories must use words and terms that are meaningful to household respondents and that

household respondents will be familiar enough to choose from. For example 'government subsidized private NGO school' should be avoided if the typical household member is unlikely to know whether a particular religious school receives government subsidies, but 'BRAC school' could be used if the typical respondent is likely to be able to identify a school in this way.

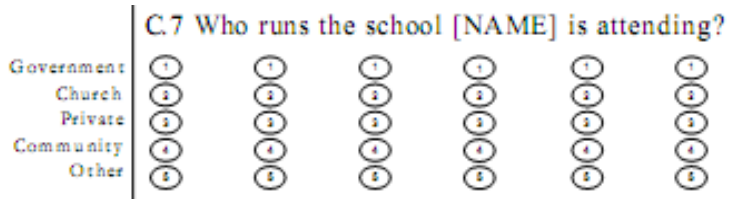
2. Response categories should be exhaustive and mutually exclusive – meaning that categories should be devised to that no school in the country that does not fall into any of the categories, and no school in the country could possibly fall into more than one of the categories.

3. The survey documentation should include a clear definition of each school type response category, as well as an explanation of how the categories can be distilled into the broader categories of 'public' and 'private' (and further divided into the UOE defined 'government dependent private' and 'independent private' if applicable).

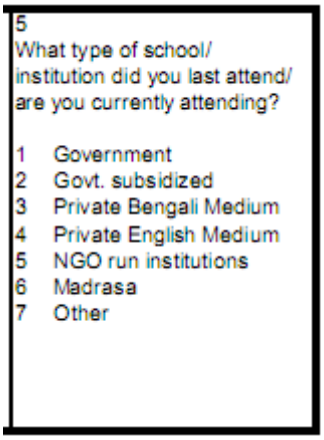
4. Response Categories should not be redundant with information gathered elsewhere in the survey. For example, do not distinguish between 'Primary Government' and 'Secondary Government' since the questionnaire already asks about school levels in other modules.



←The long and short LSMS questionnaires distinguish between 'private secular' and 'private religious.' The long version of the questionnaire repeats this question for each level of school.

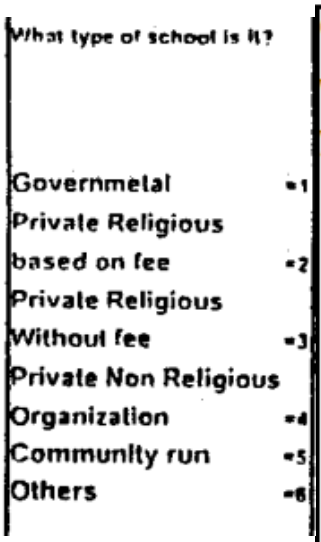


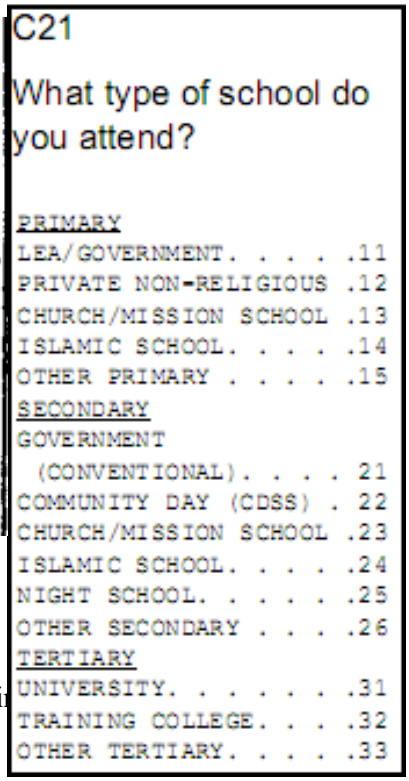
The CWIQ questionnaires provide four response ←options plus 'other.'



IHS differentiates between response options that are available for each level of education. →

←The categories used in this HIES combine information about the organization running the school and the medium of instruction.





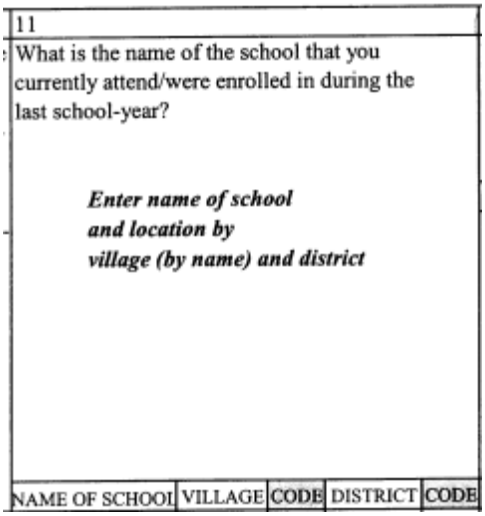
This WMS differentiates between private → schools according to whether or not they are religious *and* whether or not they are fee-based

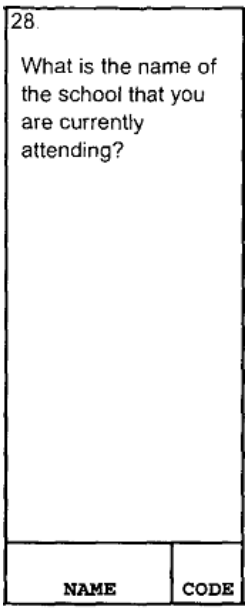
Specific School Attended

An alternative way to gather detailed information on the schools household members are attending is to use the questionnaire to gather identifying information on a school and then to use that to information to match the household member with detailed school information from an outside source. Outside sources of detailed school information might include a school survey or community resources survey conducted in concert with the household survey, or an annual Ministry of Education school census. This approach makes it possible to merge household survey data with detailed school information that could not otherwise be obtained through a household survey.

Precedence:

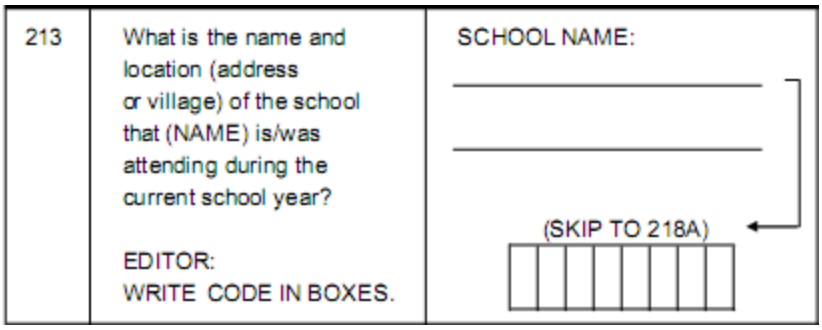
The model LSMS questionnaire includes a question asking for the name of the school attended. This information is to be matched to school information that was gathered through a related school questionnaire that was administered at all nearby schools. None of the LSS surveys analyzed in the 2008 report included this question. →





← This question in the Lao PDR questionnaire gathers unique identifying information about the school being attended by the household member.

EdData questionnaires also gather this information. →



Recommendation

If feasible, this approach is a potentially efficient way to link household members to a wealth of detailed school information. We recommend including a question asking for unique identifying information for the school, but are not able to recommend the specific information needed to match the response (besides the name of the school), since this information would be specific to the source of the school data.

Decisions on Education

Data gathered through this module can be used to analyze the reasoning behind specific household decisions related to educational career of a child. Questions in this modules build off of the core module to identify school-aged children who are of policy interest because they are not attending school, recently withdrew for school, or recently transferred from one school to another, and gather information about factors causing this change. For example, the respondent is asked for the reason that a child is not in school. These indicators are not defined or collected at the international level, but can be an invaluable resource for policymakers seeking to address attrition rates and low attendance rates.

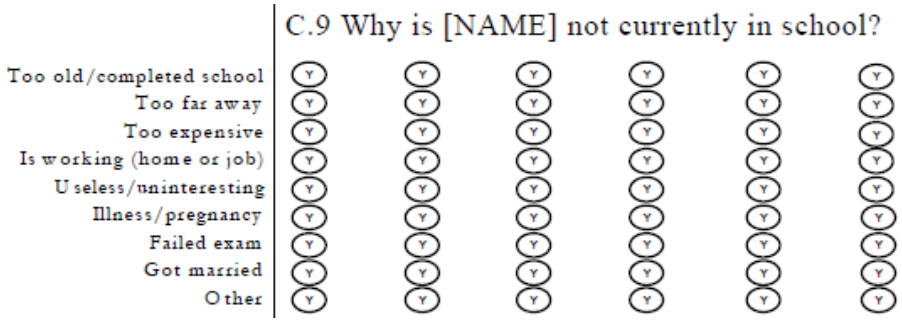
The 'reasons for leaving school' indicator can be used to improve the accuracy of efficiency measures gathered in the core module. It would, for example, allow a policymaker to draw the following distinctions in the population of pupils who attended the last grade of school in year one, and did not attend school in year two: Pupils who dropped out without completing the grade; pupils who completed the grade and did not pass the graduation exam (if it exists)[a non-proxy completion rate]; pupils who completed the grade and passed the graduation exam [graduation rate].

Precedence

DHS and **MICS** do not cover these topics.

CWIQ asks

If Ever Attended == YES and Currently Attending==NO



Respondents are advised that they can select more than one answer.

This question is well worded. Because it is asked only of household members who had attended and are not currently attending, it keeps the flow of the questionnaire simple and reduces the possibility of questionnaire flow errors. On the other hand, an opportunity to gather crucial policy relevant information is missed because the question does not apply to household members who have never attended school.

Along similar lines, because there is no filtering by age, questionnaire flow seems simple. However, there is little need to ask why newborns and 82 year-olds are not currently in school. Filtering the question to household members aged 3-30 makes sense.

ECAM Cameroon asks a single question to two different groups:

Q9
Pourquoi *(Nom)* n'a-t-il/elle pas fréquenté en 2000/2001 ou n'a-t-il/elle jamais fréquenté une école ? *CF CODES*

Codes Q9
1 = Coût trop élevé
2 = Emploi /Apprentissage
3 = Echec scolaire
4 = Maladie /Handicap
5 = Mariage /Grossesse
6 = Eloignement
7= trop jeune
8 = Autre(à préciser.)

Household members aged 5+ who have never attended school, and Household members aged 5+ who have ever attended school but are not currently attending school

It is useful that the question captures information about the two population groups we are interested in. The wording of the question is awkward and may be confusing to both respondents and interviewers. Additionally, this approach makes it awkward to include options in the menu of reasons that might be relevant to household members who had attended but not relevant to household members who had never attended (eg: 'Failed Promotion exam').

IHS Malawi asks two questions:

C11

What was the reason you never attended school? CAN GIVE UP TO 2 REASONS.

```
STILL TOO YOUNG TO ATTEND
 SCHOOL. . . . . .1
NO MONEY FOR FEES, UNIFORM .2
POOR QUALITY OF SCHOOLS. . .3
ILLNESS OR DISABILITY. . . .4
NOT INTERESTED, LAZY . . . .5
PARENTS DID NOT LET ME . . .6
HAD TO WORK OR HELP AT HOME.7
SCHOOL TOO FAR FROM HOME . .8
SCHOOL CONFLICT WITH
 BELIEFS . . . . . .9
OTHER (SPECIFY). . . . . 10
          (»NEXT MODULE)
```

1st reason	2nd reason

←One is directed at household members aged 5+ who have never attended school.

The other question is directed at household members aged 5+ who have attended in the past and are not currently attending school →

This two-question approach is attractive because it captures information on the two groups of policy interest while avoiding the problems encountered using a single-question approach. This two-question approach also helps to simplify the flow of the questionnaire.

C19

Why did you not continue your education?

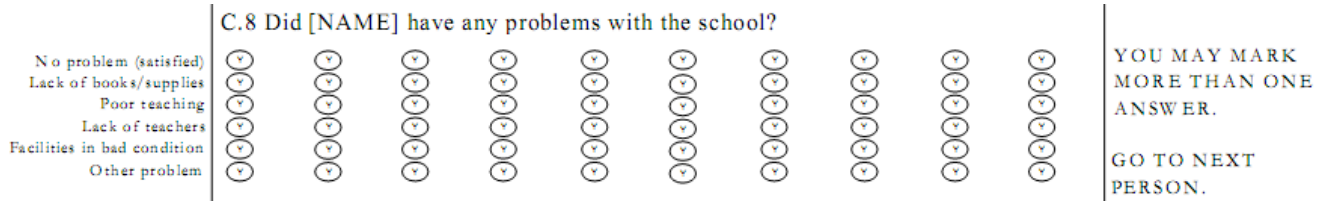
```
ACQUIRED ALL EDUCATION WANTED . 1
NO MONEY FOR FEES OR UNIFORM. . 2
TOO OLD TO CONTINUE . . . . . 3
MARRIED / BECAME PREGNANT . . . 4
ILLNESS OR DISABILITY . . . . . 5
FOUND WORK. . . . . 6
NOT INTERESTED, LAZY. . . . . 7
PARENTS TOLD ME TO STOP . . . . 8
HAD TO WORK OR HELP AT HOME . . 9
POOR/CROWDED SCHOOL FACILITIES.10
POOR QUALITY INSTRUCTION. . . .11
TEACHERS OFTEN ABSENT . . . . .12
SCHOOL TOO DANGEROUS FOR GIRLS.13
SCHOOL TOO FAR FROM HOME. . . .14
SCHOOL CONFLICT W/ BELIEFS. . .15
FAILED PROMOTION EXAM . . . . .16
DISMISSED / EXPELLED. . . . . .17
OTHER (SPECIFY) . . . . . .18
```

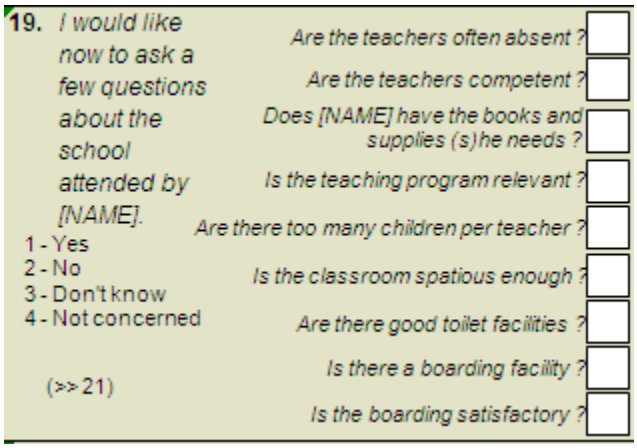
1st reason	2nd reason

Perceived quality of the school currently attended

Seven surveys included questions designed to gather information on the perceived quality (or lack of quality) of resources available at the school that a household member is currently attending. All five of the CWIQ questionnaires included an identical question on problems at the school. The Bhutan LSS questionnaire and Madagascar questionnaire also included relevant questions:

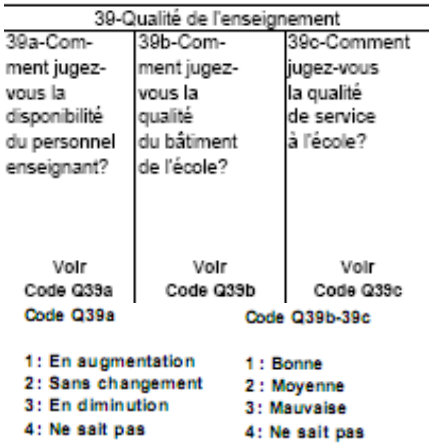
All of the CWIQ questionnaires include a question that is asked of all household members 'currently' in school. Respondents are allowed to select more than one response. The same set of response options is used in every survey:↓





←The Bhutan LSS questionnaire is structured differently, with the respondent being asked to respond to a series of questions concerning the quality of the school. The questions are somewhat indiscriminate, asking about questions such as toilet facilities that the household member may not have information on, and the quality of boarding facilities that may not exist at the school.

The Madagascar questionnaire includes a series of three questions about the availability of teaching personnel, the quality of the school building, and the quality of service at the school. Oddly, response options for the latter two questions include 'Good,' 'Average,' 'Bad,' and 'Don't Know,' while response options for the first question include 'Improving,' 'Not changing,' 'Worsening' and 'Don't know.'→



Recommendations

Do not include the question 'Why did NNNN never attend school?'

From a policy perspective, there is little value in a retrospective question that asks household members 'Why did NNNN never attend school?' When asked of older household members, the question invites foggy memories and speculation about a decision that may have been made decades ago in a policy context that has since changed. To a lesser extent, the same is true when asked of school-aged (ages 3-30) household members who have never attended school. For school-aged members, it would be more productive to ask 'Why is NNNN not attending school' since responses to this question would relate more directly to policy options that could put the household member in school today. An additional problem with this question is that it does not query a single discreet decision since the reason a household member did not attend school at age 4 may have been different that the reason the same person did not attend at age 9 and did not attend at age 20.

Do not include the question 'Did NAME have a problem with the school they attended during the YYYY [CURRENT] school year?'

Responses to these questions may be unreliable and/or biased for several reasons. First, the questions assume that the respondent has a basic knowledge of conditions of the school, even though that often may not be the case – for example, to answer the question 'Are the teachers competent' the respondent would

need to have some idea of combined the qualifications, previous experience, teaching instinct, and work ethic of the teacher in question, information that the respondent is unlikely to possess unless they spend considerable time at the school. Second, the questions require the respondent to draw a comparison between conditions at the school and the conditions desired for an ideal school, but the respondent is unlikely to have an informed concept of that an ideal school should be like. Third, there is no objective way to calibrate the magnitude of the problem reported by the respondent though we can assume that since the household member is attending school despite the problems reporting, the problems are not large enough to compel the household member to withdraw or transfer to a less troubled school.

Do use a two-question structure that queries two distinct groups separately:

- *school-age household members who have never attended school*
- *school-age household members who have attended school but are not currently attending*

Include the question 'Why is NNNN not currently attending school?' for household members aged 3-30 who have never attended school.

From a policy perspective, this question looks forward towards what can be done in order to get this population into school, rather than looking backwards at what may have prevented the member from attending sometime in the past.

Include the question 'Why did NNNN not continue his/her education?' for household members aged 3-30 who have attended school in the past but are not currently attending school.

This question will capture information about household members who left school in the most recent year as well as household members who left school some time ago. It will be possible to use the question about attendance during the previous year to differentiate between the two groups.

Response Categories

Response options for these categories will need to be tailored to the socio-economic context of the area where the survey is being conducted. Questionnaire designers should carefully test responses to make sure they are meaningful to both respondents and policymakers, and to ensure that they represent the range of experience in the enumeration area.**Recommended for Consideration: Student Mobility**

To our knowledge, data on school transfers, or the extent to which students move from one school to another school over the course of their education careers, are not collected or analyzed by the education policy community. School transfer information could potentially be of interest to policymakers because (one would imagine) households tend not to bother transferring their pupils from one school to another unless 1) they believe the new school could better serve their needs, 2) the old school is no longer available to them. By collecting data on the extent to which school transfers occur and the reasoning behind the decision to transfer, policymakers could uncover valuable information about insufficiencies in the school system and considerations that prompt a household to move their child to a new school.

Household surveys are the best potential source of data on school transfers. School transfer data are difficult to collect through school surveys because it is difficult for a school administrator in either the departing school or the receiving school to verify details of a pupil's transfer, and it is impossible for a school administrator to report the reason for a pupil's transfer. In some cases, data from school surveys overestimate the number of dropouts and new entrants in the school system because transfer students are recorded as such. Because of this, collecting 'transfers' data in household surveys would improve the reliability of pupil data from school surveys by helping data analysts differentiate between dropouts and transfers.[27]

We propose that IHSN consider adding a pair of questions regarding transfers:

If pupil attended during the current year and pupil attended during the previous year:

> Did NNNN attend the same school during the YYYY-1 [previous] school year and the YYYY [current] school year?

If 'No':

> 'What was the main reason that NNNN changed schools?

Response Categories

Since there is no direct precedent to the question on reasons for transferring from one school to another, we cannot recommend a best practice in this area. Questionnaire developers might consider the following list of reasons for possible inclusion in their response menu. As always, response menus should be adapted to local conditions and carefully tested before a survey is implemented.

- Previous school closed or moved to a new location
- Previous school did not offer instruction at the grade/level required
- Household moved to a new location
- New school easier to get to
- New school is offered at more convenient times
- New School is less expensive

[27] This section makes the assumption that the majority of student transfers occur during the break between one school year and the next – something we might call an 'inter-annual' transfer rate. However, in some school contexts, student mobility patterns may occur on an 'intra-annual' basis – for example students might attend a community school for the three months of each school year that free lunches are offered at that school, but spend the rest of each year at the government school where quality is perceived to be higher. The questions would need to be modified to address this issue.

- New school offers better instruction
- New school is a better match for our cultural/religious beliefs
- New school is safer
- New school offers additional benefits [such as a lunch program]
- New school will create more opportunities for advancement
- New school has better facilities
- Placement of Questions

Because the filter for this question requires information about whether the household member attended on both years, it is difficult to identify a way to make these questions flow smoothly with the rest of the module. One option would be to place the sequence at the end of the module, so that they are treated almost as a separate set of questions with the filter *if ages 3-30 and attended school both years.* Another slightly better option is to have the sequence immediately follow the question 'Did NNNN attend school during the YYYY-1 School year.' We believe this option is preferable because it is in line with at least one of the questions that set the condition for the sequence being asked.

Cost of Education

Overall, expenditures on education can be divided into two categories – public expenditures (expenditures by the government) and private expenditures (expenditures by households, religious organizations, charitable organizations, and other private entities). While public expenditures are often well documented, data on private expenditures are more difficult to find[28]. Household surveys can be a valuable source of information on the portion of private expenditures that is borne by households. The three modules that follow – Household Expenditure on Education,

Participation in Scholarship Programs, and

Opportunity Cost of Education – all relate to measuring the private cost of education to households.

According to the UNESCO Institute for statistics, private expenditure on education includes "direct private costs (such as tuition and other education related fees and the costs of textbooks, uniforms and transport)" and "indirect private costs (lost output when employees participate in on-the-job training)." Private expenditure on education is accounted 'net of subsidies received

[28] Education Counts – Benchmarking Progress in 19 WEI Countries. World *Education* Indicators. 2007. UNESCO Institute for Statistics. Montreal, 2007. (pg 35)

from public sources,' meaning that government-funded scholarships or subsidies that are awarded to the household to help pay for the cost of education are *not* included in the measure of private education expenditure.[29] The Household Expenditure on Education module measures the direct private costs of education. The

Participation in Scholarship Programs module measures some of the subsidies that should not be counted towards direct private costs. The

Opportunity Cost of Education module can be used to approximate the indirect private costs of education.

Household Expenditure on Education

Definition:

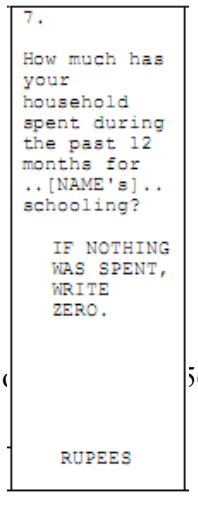
We are not aware of a formal international definition of Household Expenditure on Education so we will treat it as private expenditures on education made by the household of a pupil – the household portion of "direct private expenditures" discussed above under Cost of Education. Household Expenditure on Education is not synonymous with overall private expenditure on education because household expenditure excludes funding originating from private sources such as religious institutions and charitable foundations.

Ideally, household expenditure on education would be calculated as the sum of direct and indirect education expenses paid by the household, minus funds received by the household for the purpose of paying education expenses from outside sources (eg: international remittances or support from a charitable organization). It is beyond the scope of an education module to reliably measure the extent to which inflows such as international remittances or charitable donations contribute to education since some inflows may be 'earmarked' for education expenses and others may flow into a general household account, of which some unknown portion may contribute to education. A more complex undertaking that should be addressed through a Household Income and Expenditure type survey.

Precedence

Household surveys can be used to obtain a good measure of the direct costs of education through a set of questions directed at direct expenditures per pupil:

MICS, DHS, CWIQ, and ECAM do not include questions on these topics.



[29] IBID

In the Nepal LSS questionnaire, there is a single question on costs, which asks the respondent to give the total amount spent on education over the year. LSMS recommends avoiding this practice because it is likely to result in an underestimation of the full cost of education (presumably because it does not lead the respondent through considering the range of possible education costs).[30] →

The standard short version of the LSMS module includes a single multi-component question on expenditures. The number of cost components listed in the question varies from questionnaire to questionnaire:
→

9. How much has your household spent on your education in the last 12 months for:						
A. Tuition and other required fees?	B. Parent Association fees?	C. Uniforms and other clothing?	D. Text-books?	E. Other educational materials (exercise books, pens, etc.)?	F. Meals, transportation and/or lodging?	G. Other expenses (extra classes, optional fees)?

2. How much has your household spent on your education in the last 12 months for:							
A. Tuition and other required fees?	B. Parents association fees?	C. Uniforms and other clothing?	D. Text-books?	E. Other educational materials (pens, exercise books, etc.)?	F. Meals, transportation and/or lodging?	G. Fees for tutoring or extra classes	H. Other expenses (optional fees, etc.)?

3. Did any people who are not members of your household, such as relatives or friends, pay any of your educational expenses during the past 12 months?
YES..1 NO...2 (»6)

4. Who has paid for some or all of your educational expenses during the past 12 months?			
FATHER.....1 MOTHER.....2 GRANDPARENTS...3 UNCLE/AUNT.....4 BROTHER/SISTER.5 OTHER RELATIVE.....6 NEIGHBOR/ FRIEND.....7 OTHER (SPECIFY___)...8	1ST	2ND	3RD

5. How much money did this person (these people) pay for your educational expenses during the past 12 months?
DOLLARS

↑ In the standard long version of the LSMS questionnaire, the multi-component question is complemented with three additional questions looking at contributions to education by non-household members.

Not all LSMS surveys are worded in the same way. The 2004 Vietnam survey, for example, asks 'What is the cost for [NAME]'s attendance for the last 12 months according to the school's regulations'. This wording is not recommended: school regulations should be collected first-hand from schools rather than through a household survey. Moreover, when actual costs diverge from 'costs according to the regulations' it is much more valuable to know actual costs.

The IHS and HIES modules have a single multi-component question on expenditures. A key difference is that they include a column for total expenses in case the respondent is not able to break disaggregate the sum into component pieces. →

C30

How much was spent on your education in the last 12 months by the household, family, and friends for:

A. Tuition, including any extra tuition fees	B. School books & other materials	C. School uniform clothing	D. Boarding school fees	E. Contribution for school building or maintenance	F. Parent association & other school related fees	G. Other expenses	TOTAL (»NEXT MODULE)
MK	MK	MK	MK	MK	MK	MK	MK

521	Now I would like you to think about the last school year again (2004-2005). I am interested in learning more about what kinds of things your household spent money on for (NAME)'s schooling last year and how much was spent on each thing. In the last school year, did your household spend any money for (NAME) to get to school and home from school?	YES ..... 1 NO ..... 2 PRIVATE VEHICLE ..... 3 DON'T KNOW ..... 8 (SKIP TO 523)
522	Thinking about the last school year, how much did your household spend for (NAME) to get to and from school? Was this amount spent daily, weekly, monthly, per term, or yearly?	TRANSPORT COST LE . PT INCLUDED IN LUMP SUM .... 9999996 DON'T KNOW ..... 9999998 DAILY ..... 1 WEEKLY ..... 2 MONTHLY ..... 3 EACH TERM ..... 4 YEARLY ..... 5 DON'T KNOW 8

← EdData surveys take a more structured approach to gathering data on education expenditures. For each category of possible expenses, the respondent is first asked to consider whether the household incurred any expenses under the category, and if yes, what the costs were; next, the respondent is asked to report the periodicity of the expense payment they are describing. Respondents are also asked whether each reported expense also falls into a reported lump sum of education expenses.

Though the EdData approach is considerably more question-intensive than the LSS/IHS/HIES model, the benefits from this structure most likely outweigh thecosts: An LSMS manual on questionnaire development points out that "This formulation makes the questionnaire longer

in terms of printed pages, but probably does not increase the interview time since some sort of probing would probably have been used frequently. Most importantly, it makes the interpretation of data much clearer."[31]

By making it possible for the respondent to report expenses in the units in which they normally occur, this structure also reduces the likelihood of math errors in annualizing non-annual costs at the time of the interview. It is relatively trivial for a computer to do the math once the data have been collected.[32]

N°	During the last 12 months how much have you spent on : .... ?							
i d e n	School fees	Contributions to PTA	Uniforms and sports gear	Books and school supplies	School transport	Food, accom. and boarding	School-related activities	Other expenditure

↑ In a document outlining guidelines for statistical information systems measuring expenditure on education (SISEE), UNESCO and IREDU (University of France at Lyon) describe the minimum requirements for a household survey module that would be compatible with their guidelines.[33] According to SISEE, the module should: cover education-related expenses outside of the school, classify purchases by function rather than type, determine expenditure "throughout the complete school year" [despite the fact that their questionnaire references 'the past 12 months'], should make it possible to measure the 'sum total' expenditure, and should allow for the measurement of transfers (assistance) received from outside the household, an issue this report will address in the section entitled

Participation in Scholarship Programs.

Response Categories

[31] A Manual for Planning and Implementing the Living Standards Measurement Study Survey (Pg 44)

[32] IBID, 48.

[33] Technical Reference Manual: Statistical Information System on Expenditure on Education (SISEE). UNESCO Institute for statistics, Paris, 1998. (pg 53, 95) http://www.uis.unesco.org/ev.php?URL_ID=5455&URL_DO=DO_TOPIC&URL_SECTION=201

Many questionnaires include similar expense categories, even if those categories are grouped differently in each questionnaire expense matrix. Figure 8 includes illustrative examples of the set of categories included in three household expenditure matrices. We recommend against lumping together expense categories (ie: we recommend recording 'transportation expenses' and 'accommodations expenses' as separate expenses rather than lumped together as 'transportation and accommodations expenses'). This is in keeping with the philosophy as outlined in the previous page. We recommend the following expense categories as a starting point for consideration:

- School fees
- Other specifically named fees
- Parent Association fees
- Other Fees
- Private tutors
- Uniform
- Sports clothing and other required clothing
- Text books
- Other learning materials
- Meals
- Transportation
- Boarding or lodging
- Other Expenses

It is crucial that expense categories be adjusted and tested to reflect the context where the survey will be conducted. If schools do not use uniforms, then uniforms should be dropped from the list. If schools charge an attendance fee and a graduation fee, then each fee should listed individually. If, in a pre-test a category is never used, then consider dropping or changing it. If 'Other expenditures' is consistently large, probe to find out what 'other' includes and add it to the list of categories.

Figure 8: Household Expenditure Expense Categories used by LSMS, IHS/HIES, and SISSEE

Category	LSMS	HIS/HIES	SISSEE
School Fees	Tuition	Tuition	School Fees
	other required fees	any extra fees	
Parent Association Fees	Parent Association Fees	Parent association	Contributions to PTA
		and other related fees	
Clothing	Uniforms	School Uniform	Uniforms
	other clothing	clothing	
			sports gear
Materials	Text books	School Books	Books
	Other educational materials (exercise books, pens, etc)	other materials	school supplies
	Meals	Boarding school fees	Food
Logistics	Transportation		School transportation
	Other expenses (extra classes, optional fees)	Other expenses	
	Lodging		Accommodations and boarding
		Contribution to school for building maintenance	
Misc			School Related Activities
			Other Expenditure

Recommendations

Timeframe: Of the 14 surveys investigated in EPDC's 2008 report that included questions about household expenditures on education, eleven referenced the past 12 months (or past year) as the timeframe for which respondents should report costs. Two reference the current school year, one references the previous school year, and another references

'this year or last.' The EdData questionnaire, which was not analyzed by EPDC in 2008, specifically names the previously completed school year.

From an education policy perspective, it is most useful to use the previously completed school year as the time frame of reference for these questions. It is better to tie costs to a specific school year because this makes is possible to measure the costs associated with particular grades; costs that are simply reported over 'the previous 12 months' are likely to span two academic years and thus confuse the costs associated with the grade attended during each year. Though respondents may find it challenging to remember costs associated with the previous school year rather than the current school year, it is important to ask about the previous school year in order to capture information about important examination and promotion fees that are concentrated at the end of the school year; these data would be lost in questions about the "current" school year.

Payer: Of the 14 surveys investigated in 2008, eight were worded so they queried expenses paid by the household; two used the word 'you', which is a bit ambiguous because it might refer either to the respondent or to the pupil; three were written an a passive voice so that no payer was named in the question; and one asked about expenses paid by the 'household, family, or friends.' We recommend *against* asking about donations from outside the household because it can be difficult to determine the influence that contributions to a general household fund have on expenditures to education. We recommend asking about all 'household' expenditures to education without any reference to how the money came into the possession of the household.

Currency: All of the questionnaires investigated appear to make the assumption that there will be no variation in the unit of currency used in responses. Some questionnaires avoid the topic of currency altogether, which is a mistake since respondents may be predisposed to think in terms of a currency other than the national currency, or to speak about thousands, millions, crore, or lakh as if they were single units. In our estimation, for surveys that cover an area where all respondents are likely to be comfortable using a single currency, it is sufficient to specify the currency and number of units in which responses should be given. For surveys that cover an area where more than one currency is commonly used, it may be wise to go a step further and give respondents the option of selecting the currency costs are reported in (as per our discussion of timeframe above). This makes it possible to convert currencies at the time of data processing rather than at the time of the interview.

Filtering: These questions should apply to all household members who we know attended school during the *previous* complete school year. This group will, by default, be limited to household members ages 4-30.

Structure: We recommend keeping the more probing question structure established by EdData, but presenting it in a grid format that is consistent with the format used throughout the education modules.

Guiding Language

Follow the example set by the EdData questionnaire:

I would like to continue asking about NAME and his/her schooling during the YYYY-1 [previous] school year. I am interested in learning more about what kinds of things your household spent money on for

Figure 9: Sample proposed layout for Household Expenditures on Education module

NAME's schooling last year, and how much money was spent on each thing.

I would like to continue asking about NAME and his/her schooling during the YYYY-1 [previous] school year. I am interested in learning more about what kinds of things your household spent money on for NAME's schooling last year, and how much money was spent on each thing.
During the full YYYY-1 school year, did this household spend money on XXXX for NNNN's education?
How much?
Was this amount spent Daily, Weekly, Monthly, Each Term, School Year?

Response Codes: Yes.....1 No.....2 Don't Know.....3

Time Codes: Daily.....1 Weekly....2 Monthly..3 Each term....4 Each year.....5 Don't know..6

ID	Tuition and required fees			Parents Association Fees			School uniform and other required clothing			Text Books			Other Educational Materials			Meals			Transportation			Boarding			Other Educational Expenses		
	Y/N	Kwacha	Time Unit	Y/N	Kwacha	Time Unit	Y/N	Kwacha	Time Unit	Y/N	Kwacha	Time Unit	Y/N	Kwacha	Time Unit	Y/N	Kwacha	Time Unit	Y/N	Kwacha	Time Unit	Y/N	Kwacha	Time Unit	Y/N	Kwacha	Time Unit

Participation in Scholarship Programs

As discussed in the Cost of Education section, the value of subsidies (such as scholarships) is difficult to address unambiguously in an education questionnaire for several reasons: it is difficult to determine whether a scholarship contributes to a specific pupil's education or to a general household fund and how the scholarship funds should be accounted for in subsequent analysis. Moreover, it is not certain how often respondents are aware of the actual value offset by any subsidy or scholarship they are the beneficiaries of. Rather than attempt to disentangle the ambiguities related to the financial value of scholarships and other subsidies, our advisory panel of education policymakers recommends that questionnaire designers focus on gathering information about participation in *specific* subsidy/scholarship programs. With a measure of participation in-hand post-enumeration analysts can use public information about each scholarship program to estimate the true value of any subsidies.

A question about participation in a subsidy program should be preceded by a question used to determine if the respondent is aware that this specific subsidy program exists at all. If the respondent has never heard of the program, then they should not be asked the following question because they are not qualified to determine whether the household member in question is participating in the program.

It is key that the question sequence names and asks about a specific subsidy program; if questionnaire developers are interested in learning about more than one subsidy program, then the second subsidy should be asked about in a separate additional question series.

Precedence

Among the 30 surveys analyzed by EPDC in our previous report, eight covered the issues of scholarships or other education subsidies in some way. Of these eight surveys, six were LSS surveys, one was an HES

survey, and one was an IHS.

8.	9.
Did ..[NAME].. receive a scholarship to help pay for your educational expenses? YES.....1 NO.....2 (➔NEXT PERSON)	How much did ..[NAME].. receive for scholarship over the past 12 months? RUPEES

← The LSS and IHS questionnaires address subsidies trough a pair of questions – the first asking whether each household member who attended school in the past 12 months received a scholarship and the second asking what the value of the scholarship was. It is unclear what should be done if a household member benefitted from more than one subsidy.

The HES questionnaire takes a more targeted → approach, filtering the questions to apply only to female students attending secondary school. The survey asks whether the household member is receiving a specific subsidy; it then asks about the value of the subsidy, and ends with a question about a tuition waiver, but does not ask about the value of the waiver – one might presume that this is out of recognition of the reality that respondents are less likely to be aware of the value of a waiver.

ASK ONLY SECONDARY FEMALE STUDENTS		
5 Are you recieving the female secondary stipend? **ASK ONLY FEMALE STUDENTS** 1 Yes 2 No **>>8**	6 How much did you receive in total from this stipend in the past 12 months? Taka	7 Do you benefit from the tuition waiver? 1 Yes 2 No

Timeframe

The LSS/IHS questionnaires filter for household members who attended school in the past 12 months. As we discuss in the

School Participation section, however, it is strongly recommended that references to attendance are filtered with reference to a specific academic year (eg: "The 2009/2010 school year"). All subsidy questions should be filtered according to whether the household member attended school during a specifically named school year. Careful consideration should be given, however, as to whether the school year named should correspond to the 'Current/Most Recent school year' or the 'previous school year' (see discussion of timeframe from the Educational Efficiency module).

Overall, it is preferable that the timeframe referenced in the filter and the subsidy question refer to the

'Current/Most Recent' school year because: 1) this will produce the most timely data, and 2) respondents will have a more accurate memory of the most recent year than they would have of the previous school year. A notable drawback of this approach, however, is that school subsidy information will not be synchronized with educational expenditure information because the expenditure information is always collected with reference to the previous complete school year (see discussion of timeframe from the Household Expenditure on Education module). The structure of the scholarship program should also be taken into consideration – if the scholarship is an ongoing process or benefit that is initiated with the beginning of the school year, then it is possible to set a timeframe to refer to the current school year. If, however, the scholarship is a one-time benefit that may take place late in the school year (ie: an exam fee waiver), then it would be better to ask about the previous school year in order to ensure that the entire academic year is covered by the question. The questionnaire designer will need to weigh these considerations when determining the timeframe for this question.

If scholarship questions use a different timeframe than the previous set of questions, then it is essential that guiding language is included to alert the survey respondent to the school year they are now being asked to consider (note that this will vary depending on what timeframe is selected for this sequence of questions).

Filtering

The LSS/IHS questionnaire filter to include all household members who are considered to be attending school. The HES questionnaire filters to include only those household members who are eligible to participate in the specific subsidy – in this case females attending secondary school.

It is recommended that, at a minimum, the sequence of subsidy/scholarship questions should be filtered to include only household members who are attending school. However, if the specific subsidy in question targets a more specific subpopulation, the filtering should be adjusted correspondingly. If this module is expanded to address more than one subsidy program, and these subsidy programs are intended to target *different* sub-populations, then the question sequences for each subsidy program should be ordered so they progress from the most broadly-defined group to the most narrowly-defined group. In any case, it is imperative that filtering instructions are prominent on the questionnaire and interviewers are clear on what should be asked of whom. If more than one question sequence is used, both questions sequences should refer to the same time frame.

Wording

Because the filters and questions must be modified to fit specific scholarship programs, example questions are given with reference to two hypothetical programs, one that provides free lunches to primary and secondary pupils of a certain income level that cannot be screened for, and a second which provides two services, scholarships and mentoring, to females attending secondary school:

If household member attended Primary or Secondary school during the YYYY current school year

Have you heard about a program called 'Free Lunch at School' program?

If YES

During the YYYY school year, Did NAME get free lunches through the 'Free Lunch at School' program?

If household member is female and attended Secondary school during the YYYY current school year

Have you heard about a program called 'Smart Girls Scholarship Program (AGSP)' ?

If YES

During the YYYY school year, Did NAME receive a scholarship through ''Smart Girls Scholarship Program (AGSP)'

During the YYYY school year, Did NAME have a mentor through ''Smart Girls Scholarship Program (AGSP)'

Opportunity Cost of Education

The financial burden to households for enrolling children in school can be divided into two subcategories: Direct Costs and Opportunity Costs (a.k.a. indirect costs). Direct Costs, which include direct payments for education related expenses, are investigated in the 'Household Expenditure on Education and

Participation in Scholarship Programs modules. Opportunity Costs, also referred to as indirect costs, refer the foregone value of the best use of a person's time. In the context of education policy, the opportunity cost of attending school is the foregone value of whatever the pupil would have been doing if they were not attending school. In developing countries, the opportunity costs of education could be high since it is not uncommon for children to help with household chores, farm work, or earn wages.

The Opportunity Cost of education is calculated using two pieces of information: the amount of time spent on school and school-related activities, and the value of the next best use of the time spent on school and school-related activities. The education module can be used to obtain the first of these pieces of information – the amount of time spent on school and school-related activities, the financial value of the next best use of the student's time must come from another module of the household survey or from an outside source[34]. The time component of the opportunity cost of education includes time spent attending

[34] Two possible sources of valuations are: 1) the standard local wage rate for child labor, if such a

school or classes, time spent traveling to and from school, and time spent on homework. Even if analysts are not able to obtain a value for the next best use of a child's time, the amount of time, in itself, is a valuable proxy measure for opportunity cost.

Policymakers thinking about the opportunity cost of education might also look outside of the education modules for information about the value to the household of a child's time. For example, household questionnaires often include a series of questions on the time needed to travel to the household water-source and who is responsible for gathering the water, as well as the time needed to travel to sources of fuel for heating/cooking and who is responsible for gathering the fuel. If these responsibilities rest on children and the time needed to gather water or fuel are significant, then the opportunity cost of a child's time can be presumed to be high. An advantage to this approach is that it assists with the formulation of policy alternatives to respond to specific causes of high opportunity cost (a nearby well could drastically reduce the time needed to gather water).

Students boarding at a boarding school

Definition

We are not aware of an internationally-accepted definition of pupils boarding at a boarding school, but consider it sufficient to define students boarding at school as students who are attending a school where they also live and eat with a community of their peers. Because most schools that enroll boarding students also enroll non-boarding students it is important to distinguish between students merely attending a boarding school and pupils boarding at a boarding school.

Discussion

It is essential to note that, under most circumstances, household surveys are not an appropriate tool for gathering information on students who are boarding at a boarding school. This is because, based on the way 'household' is defined in many household surveys, these students may be either systematically excluded from household rosters or be inconsistently included in household rosters.

Survey planners and analysts should consult the definition of the household to confirm whether or not boarding students attending a boarding school would be expected to be included in a household survey, but the general expectation should be that they are not. For the purposes of calculating many education indicators, accidentally including data from a few boarding students would have little impact on an indicator value. Other indicators may be are more likely to be skewed by boarder data, and when this is the case, questionnaire planners should consider including a question on students boarding at a boarding school as a way to protect against bias.

Data on boarders can be helpful as a quality control measure for information gathered in

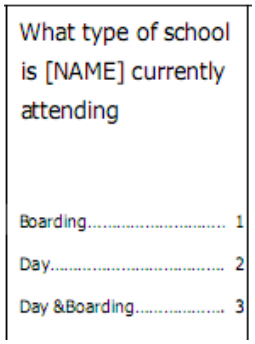
s ts, obtained from the community questionnaire, or 2) the profit function of a farm or
l discussed on page 159 of LSMS 2000 Vol 1.

the indicators 'Distance or time traveled to school' or '# hours a child typically spends at school.' If boarding students, who may travel exceptional distances to reach a school, and typically spend 24 hours a day at the school, are not screened out of the calculation of these indicators, indicator results may have biased results.

Information about pupils boarding at school is not a priority in the international community, though it may be of interest to national or local planners who see boarding schools as a way to facilitate access to schooling to children who live in remote areas where a school cannot be accessed on a regular basis. The questions proposed in this section could possibly be used to gather information on students boarding at a boarding school if the definition of a household were modified to ensure that the survey respondent included in the household roster included former household members who are currently away at a boarding school. The Malawi IHS survey does this. It is beyond the scope of this report to recommend practices on modifying the definition of a household.

Precedence

Among the 30 surveys EPDC analyzed in 2008, two included questions on boarding students attending a boarding school.



← The Malawi IHS questionnaire asks "Are you a boarder or a day student at the school?" and allows two response options.

The Uganda National Household Survey asks "What type of school is NAME currently attending" with response options indicating that the question has to do with boarding or non-boarding

Recommendation:

We recommend a question similar to that asked in the Malawi IHS question:

If attending school during the current year:

Was NAME a day scholar or a boarding student during the YYYY [CURRENT] school year?

Response options:

Response options for this question should allow for a simple binary response. Options can be adjusted to reflect common usage in the country where the survey will be conducted, but we recommend these responses:

- Day Student
- Boarding Student

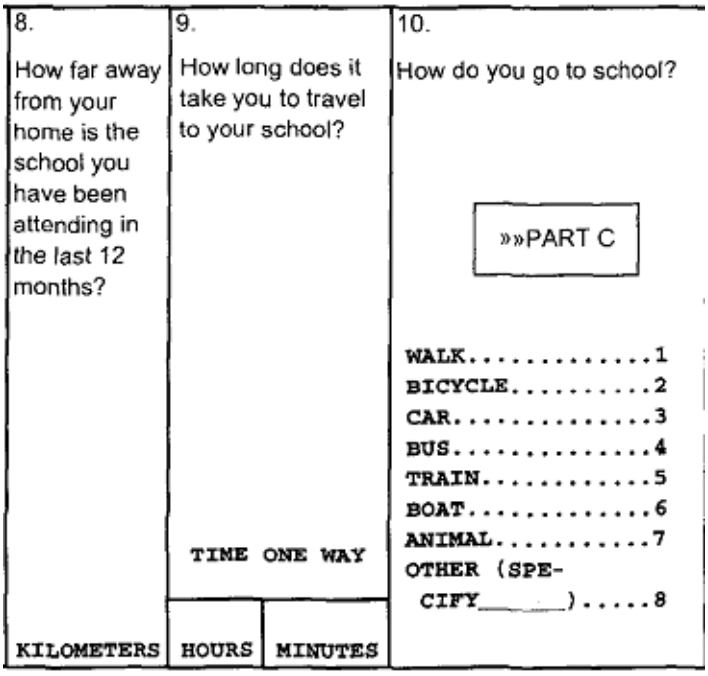
Distance or Time travelled to school

Definition

To our knowledge, there is no international standard definition of this indicator.

Precedence

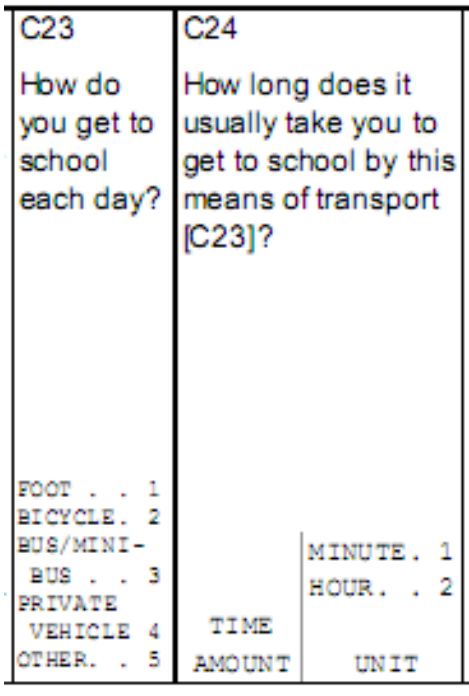
Among the 30 surveys analyzed for the 2008 report, eight questionnaires included one or more questions addressing either the distance, amount of time, or mode of transportation taken by a pupil to travel to school. Of these, five included a question on the amount of time it takes to travel to school, three asked about the distance as measured in kilometers and/or meters, and two asked about the mode of transportation used.



8. How far away from your home is the school you have been attending in the last 12 months?	9. How long does it take you to travel to your school?		10. How do you go to school?
	TIME ONE WAY		»»PART C WALK.....1 BICYCLE.....2 CAR.....3 BUS.....4 TRAIN.....5 BOAT.....6 ANIMAL.....7 OTHER (SPE-CIFY_____).....8
KILOMETERS	HOURS	MINUTES	

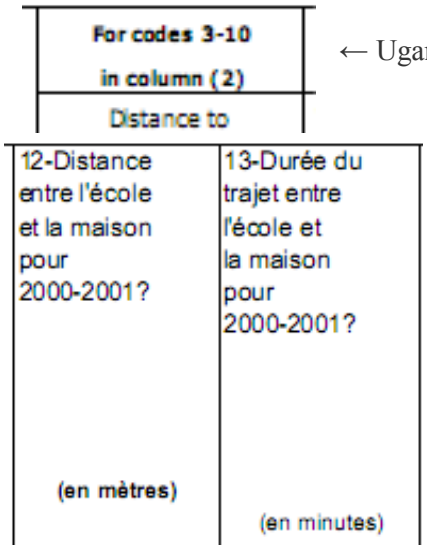
← The long version of the LSMS questionnaire includes examples of all three questions in sequence. Responses to the time question can be given in a combination of minutes and hours.

The Malawi questionnaire first determines what mode of transportation is used to reach school, and then asks the amount of time needed using that mode of transportation. The time questionnaire also notes whether values are given in minutes or hours. →

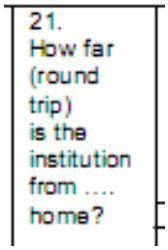


C23 How do you get to school each day?	C24 How long does it usually take you to get to school by this means of transport [C23]?	
FOOT . . 1 BICYCLE. 2 BUS/MINI-BUS . . 3 PRIVATE VEHICLE 4 OTHER. . 5	TIME AMOUNT	MINUTE. 1 HOUR. . 2 UNIT

← Uganda simply queries 'Distance to school' and specifies km.

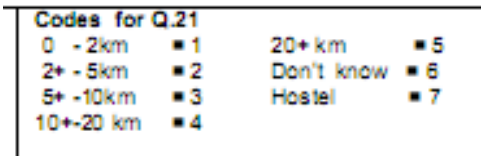


For codes 3-10 in column (2)	
Distance to	
12-Distance entre l'école et la maison pour 2000-2001? (en mètres)	13-Durée du trajet entre l'école et la maison pour 2000-2001? (en minutes)



←Madagascar asks both about the distance to the school from the home and time it takes to reach the school from the home, specifying the unit to be used in the response to each.

The Pakistan survey asks distance to school and specifies response codes →



Discussion

Any one of these questions or a combination of these questions may be appropriate for use in the questionnaire depending on national context and the policy issues of interest to questionnaire designers. Because these indicators are not collected in a systematic way at the international level, questionnaire designers have some leeway in selecting a question or combination of questions that meets their needs. Some discussion of the alternatives is offered below:

Distance measured as a function of time

Overall, it seems that there is more policy relevance in measuring distance to school in terms of the amount of time it takes the pupil to get there rather than the physical distance to the school. Because of the uneven distribution of quality infrastructure in developing countries, distance in itself is not a consistent measure of relative difficulty or ease of travelling to school. A 2 km trip to school, for example, presents a vastly different challenge for a child walking alone on a flooded, unimproved, road in the forest, as compared to a youth traveling via moped on an improved road. Measuring distance to school in terms of the time it takes to travel from home to the school is a closer approximation of the relative barrier presented by the trip.

A question could measure time in terms of travel time needed for "you" (meaning the respondent, not the child) or "a typical adult" to travel from the home to the school. Advantages of this approach would be that interview respondents may not know exactly how long it takes the child to reach school, but are more likely to be able to report how long it might take themselves or another typical adult to walk the same

distance. Disadvantages are that such a response would be less relevant to the challenges a *child* might face, and that even when asked to report the time it needed for the respondent to walk a distance that the respondent walks every day, the respondent is likely to respond with only a rough estimate. Given that either approach presents the potential for an unknown amount of ambiguity, we suggest the question that addresses the child's transit time.

Distance measured as geographic distance travelled

Measuring distance between the school and the home geographically (ie in kilometers or miles) has the advantage that measures such as kilometers or miles are standard units that are inherently comparable across time and countries. Much of that advantage is negated when these distances are obtained through a household survey interview because the distances are not being measured in any consistent manner, but rather as reported by a household member. If the distance between a home and a school is to be measured as part of a household survey, then the value of that distance should be obtained through some instrument other than the interview itself. Measures of distance might be obtained through the use of GIS or an odometer.

Mode of transportation

A question on the mode of transportation used to reach school may be of interest to policymakers under specific circumstances, but it is not of general interest to the international community. One might argue that data on the time it takes to reach school cannot be used without an understanding of the mode of transportation used by the child, but this is only partially true because information on the status of infrastructure between the home and the school would still be lacking. Infrastructure may be a more important consideration than mode of transportation in many scenarios, since, to continue the example above, traveling 2 km by moped on a flooded, unimproved, road in the forest may be more difficult than the same distance by moped on an improved road. The question on mode of transportation is not recommended for general use in household surveys, but may be included if policymakers are specifically interested in investigating this topic.

Recommendations:

We recommend that data on the distance between the home and the school be collected as a function of the *time* needed to make the trip and not the geographic distance between the home and the school. Data on the geographic distance between a school and a household can be relevant to policymakers, and if collected, we recommend that it is not collected through an interview question, but using a more reliable instrument such as GIS or an odometer.

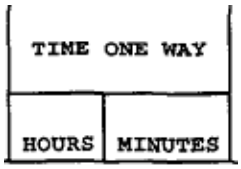
It is recommended that any question about the amount of time taken reach the school should be worded precisely in order to avoid any potential ambiguities:

> *On a typical day during the* YYYY *[CURRENT] school year, how much time did it take for NAME to travel directly (one way) from this household to the school NAME was attending?*

If used, the question on mode of transportation could be worded as:

On a typical day during the YYYY [CURRENT] school year, how did NAME to travel from this household to the school NAME was attending?

Response Categories:



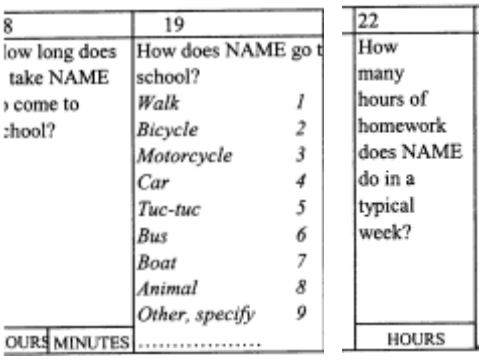
The response categories on time travelled used in the LSMS questionnaire allow for the response to be given in a combination of hours and minutes. This flexibility is desirable, but careful instructions should be given to ensure that times are recorded consistently.

If the question on mode of transportation to work is used, response options should be organized into a coded response menu. Response option categories must be developed and tested in country in order to ensure that categories are appropriate for the local context. As with all response option menus, the names of the categories must be meaningful to questionnaire respondents, and must be exhaustive and mutually exclusive.

Time Commitment for School

For the purposes of household surveys, we take this as a measure of the amount of time a pupil spends at school and engaged in school-related activities outside of the school.

Precedence

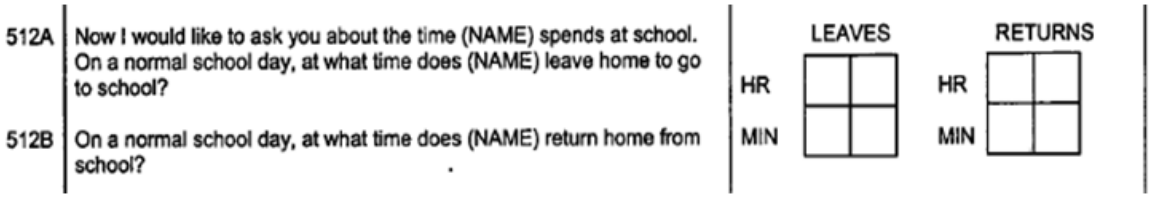


18	19		22
low long does take NAME come to chool?	How does NAME go t school?		How many hours of homework does NAME do in a typical week?
	Walk	*1*	
	Bicycle	*2*	
	Motorcycle	*3*	
	Car	*4*	
	Tuc-tuc	*5*	
	Bus	*6*	
	Boat	*7*	
	Animal	*8*	
	Other, specify	*9*	
OURS MINUTES	.....		HOURS

←The Lao PDR questionnaire includes a question on the amount of time used to travel to school (the full wording of the question is not visible in our copy of the questionnaire), and another question on the number of hours of homework performed per week.

The Malawi IHS questionnaire includes questions onthe amount of time needed to travel to school, the amount of time spent attending school, and the amount of time spent on homework. The transportation question asks about the usual amount of time needed whereas the attendance and homework questions ask specifically about the amount of time taken on the previous day. →

C23	C24		C25	C26
How do you get to school each day?	How long does it usually take you to get to school by this means of transport [C23]?		How many hours did you spend yesterday attending classes at school or university, or in some other training course?	How many hours did you spend yester-day doing home-work, that is, school work to be done at home?
FOOT . . 1 BICYCLE. 2 BUS/MINI-BUS . . 3 PRIVATE VEHICLE 4 OTHER. . 5	TIME AMOUNT	MINUTE. 1 HOUR. . 2 UNIT	HOURS	HOURS

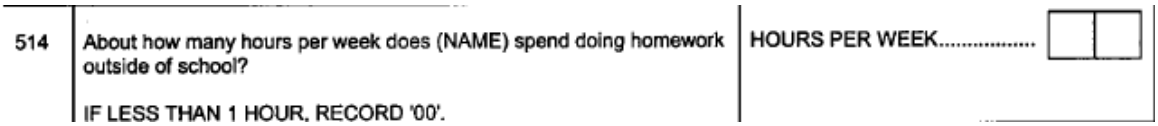


		LEAVES	RETURNS
512A	Now I would like to ask you about the time (NAME) spends at school. On a normal school day, at what time does (NAME) leave home to go to school?	HR	HR
512B	On a normal school day, at what time does (NAME) return home from school?	MIN	MIN

↑ The EdData questionnaire asking for the specific times that the pupil typically leaves for school and returns from school. Taken together, the responses to these questions are used to calculate the amount of time taken for transportation and attendance combined. This approach seems well-reasoned because by asking about the time of departure and arrival it focuses on gathering information that the survey respondent is most likely to know the answer to, as opposed to the questions about the amount of time needed to travel to school, which, as discussed in the Distance or Time travelled to school section, forces the respondent to engage in a certain amount of speculation.

The approach is not appropriate for our purposes, however because to use it, we would need to make the assumption that, from the time the child leaves home in the morning to the time the child returns home at night, they do nothing engage in school-related activities and school-related travel. This assumption may be untenable if there is a good chance children will spend time playing, engaged in sports, or enjoying other non-scholastic pursuits before they return home.

EdData uses a separate question to ask about the number of hours per week spent on homework.↓



514	About how many hours per week does (NAME) spend doing homework outside of school? IF LESS THAN 1 HOUR, RECORD '00'.	HOURS PER WEEK..... ☐☐

Several LSS questionnaires and the Sierra Leone IHS questionnaire ask about the amount of time travelled to school but do not ask about time spent at school or time spent on homework.

Recommendations

Each of the three components of indirect cost should be queried in a separate question:

> *On a typical day during the YYYY [CURRENT] school year, how much time did it take for NAME to travel directly (one way) from this household to the school NAME was attending?*
>
> *On a typical day during the YYYY [CURRENT] school year, how much time did NAME spend away from home for school?*
>
> *Does this include the travel time?*
>
> *On a typical day during the YYYY [CURRENT] school year, how much time did NAME spend on school work outside of school hours (for example, studying, doing homework, or working with a tutor)?*

Pupil Absenteeism

Definition

To our knowledge, there is no international standard definition of this indicator.

Precedence

Out of the 30 surveys investigated by EPDC in 2008, three included one or more questions intended to gather information about student absenteeism.

3	4
How many days were you absent from school/ educational institutions during the past 30 days? **IGNORE DAYS SCHOOL CLOSED OR PUBLIC HOLIDAYS** **IF NO ABSENCE, WRITE "00"**	What was the main reason for your longest absence from school/educational institutions over the past 30 days? 1 No absence 2 Sick 3 Household work 4 Farm work 5 Family business 6 Other work 7 Bad weather 8 Other reasons
No. of days	

← The Bangladesh HIES question sequence includes a question to ascertain a rate of absence for the pupil, and a second to get at the reason for any absence. The first question in this sequence is overly complex because it asks the respondent to answer two sub-questions simultaneously: 3a) How many days was the school open the past 30 days; 3b) How many of those days did the pupil miss? These should be broken down into separate questions.

The Malawi IHS questionnaire correctly asks separate questions about the number of days the school was open and the number of days the pupil attended. A third question is used the find the reason for any absence. →

24	25	26
How many days has NAME'S school been open in the past 7 days? ***If 0 days >> next person***	How many days has NAME attended school in the past 7 days?	If NAME was absent any days, what was the reason? *No absence = 1* *Agricultural work=2* *Work at house/home =3* *Other work =4* *Illness = 5* *Family illness/death=6* *Other (specify…) 7*
DAYS	DAYS	

← The questionnaire from the Lao PDR ECS also includes a pair of questions to determine the number of days in a recent period that the school was open and the number of those days that the pupil attended. The questionnaire adds a third question designed to gather information about any longer-term absences over the previous 12 months.

C27	C28	C29
How many days was your school in session over the past 2 weeks?	How many days did you attend school in the past 2 weeks?	At any time in the past 12 months, did you ever temporarily withdraw from school, so that you missed more than two consecutive weeks of instruction? YES . . 1 NO . . . 2
NUMBER	NUMBER	

505 Now I would like to ask you some questions about (NAME)'s school attendance last year. There are many reasons that children sometimes do not attend school, even though school is open and classes are in session.

In the last school year did (NAME) miss school for any of the following reasons? RECORD ANSWER FOR EACH REASON LISTED. IF YES, ASK AND RECORD NUMBER OF DAYS MISSED FOR THAT REASON. IF NUMBER OF DAYS IS MORE THAN 95, RECORD 95.		NUMBER OF DAYS
(NAME) was needed to care for sick relatives or household members.	SICK RELATIVE YES 1 → NO 2	
(NAME) was needed to help at home such as caring for younger children, cooking or cleaning, or fetching water or wood.	DOMESTIC WORK YES 1 → NO 2	
(NAME) was needed to tend animals, or work on the family (NAME) was needed to tend animals, or work on the family farm or in the family business.	FAM. FARM/ BUSINESS YES 1 → NO 2	
(NAME) was needed to work for an employer.	EMPLOYER YES 1 → NO 2	
School fees or other school costs were due, and the money was not available.	NO MONEY YES 1 → NO 2	
(NAME) was receiving private tutoring or was studying for exams at home.	TUTORED/ STUDYING YES 1 → NO 2	
(NAME) did not want to go to school.	DID NOT WANT YES 1 → NO 2	
Because of a wedding, funeral or other ceremony.	FUNERAL/ CEREMONY YES 1 → NO 2	
(NAME) was ill.	ILLNESS YES 1 → NO 2	
(NAME) did not have proper/suitable clothes	SUITABLE CLOTHES YES 1 → NO 2	
(NAME) woke up late.	WOKE UP LATE YES 1 → NO 2	
Because of bad weather.	BAD WEATHER YES 1 → NO 2	
(NAME) missed school for any other reasons.	OTHER YES 1 → NO 2	
	(SPECIFY)	

508	CHECK 505 AND CALCULATE THE TOTAL DAYS ABSENT. WRITE THE TOTAL NUMBER IN THE SPACE PROVIDED BELOW. In total, you have said that (NAME) missed ___ days of school last school year. Does that seem accurate? IF NO, PROBE AND MAKE CHANGES TO Q. 505. THEN REPEAT Q. 508. IF NUMBER OF DAYS IS MORE THAN 95, RECORD 95	YES ..... 1 NO ..... 2 NUMBER OF DAYS

← The EdData questionnaire uses a structured set of question to gather information on the number of days of school missed by reason for having been able to miss school. This is a more rigorous way to gather information on the topic since it forces the respondent to think more carefully about the different ways the pupil may have missed school, but it the tradeoff is that it invests time in gathering a lot of detailed information that otherwise may not be of great interest to researchers.

Discussion

Broadly speaking, household surveys are not an ideal source for detailed information on student absenteeism. There are two main reasons for this:

First, before one can determine the number of days of school that a pupil missed over a set period, it is necessary to have baseline information on the number of days that the school was open over the same period. In a household survey, this baseline information can only come from the survey respondent, but respondents can not necessarily be relied upon to be aware of, let alone to correctly remember and report, all holidays, teacher in-service days, days that a teacher may have been absent for one reason or another and so on.

Second and more importantly, because household surveys are *not* designed to take into account variations in activity related patterns in the academic calendar or the progression of the seasons, it would be difficult to obtain truly unbiased results for a measure of absenteeism. Household surveys are conducted over a matter of 2-6 months, and may conducted in a manner such that each geographic area represented in the survey may be canvassed separately and in sequence. In other words; the survey is completed in Region A before it is begun in region B and so on. Some regions may be canvassed while school is in session and others may be canvassed while school is out of session. One region may be canvassed during flu season, malaria season, rainy season, fishing season, or tourist season, and another region may not be. Alternatively, the whole country may be canvassed at a time when absenteeism is abnormally high or low. In short, many of the causes of student

absenteeism are cyclical in nature, but household surveys are not equipped to adjust for seasonal variation.

Overall, school surveys are a better source of student absenteeism information than household surveys. With a school survey, it is easier to gather reliable baseline data on when a particular school was open, and school surveys are easier to adjust so they account for possible seasonal variations in attendance.

Despite the shortcomings discussed here, household surveys can still be an interesting source of information about absenteeism, especially when care is taken to avoid problems associated with seasonal variation or attempting to establish a base number of days that school was open. To do this, we recommend a question that asks broadly about the number of days missed over an entire school year.

Recommended for consideration:

We recommend that household survey designers consider a new style of question designed to measure pupil absenteeism in a way that mitigates some of the challenges outlined above. Such a question might be worded as follows:

If NAME attended school during the YYYY-1 [PREVIOUS] School Year

There are many reasons that a children do not attend school, even though the school is open and classes are in session. Over the full *YYYY-1 [PREVIOUS]* School Year, how many days did NAME miss school even though school was open?

An advantage of this approach is that it helps to smooth out biases in absenteeism that may be caused by seasonal or regional patterns. This could be followed by a question on the primary reason for any absences. We recommend a variation on the Bangladesh question:

If NAME missed one or more days of school

What was the main reason for the NAME's longest absence from school during the *YYYY-1 [PREVIOUS]* school year?

Response options should be tested and tailored to the specific country context.

Non Formal Education

The core questionnaire modules proposed in the Core Modules section of this report are excellent for gathering information on educational attainment or school attendance and as they relate to the formal education sector, but are not effective for gathering data on attainment or attendance in non-formal education settings. Because non-formal education practices such as workplace training programs, apprenticeships, internships, and non-traditional education programs can be a significant source of human capital in some contexts, it would be an oversight to focus on formal education while ignoring non-formal

education. GMR goals related to adult literacy and life-skills learning, greater emphasis has been placed on the measurement of human capital obtained from non-formal sources.

Definition of non-formal education

We are not aware of a clear universal definition of non-formal education. For the purposes of household surveys, the distinction between formal and non-formal education is important though a precise taxonomy of the differences between the two categories is not necessary. This discussion taken from the ISCED97 documentation is sufficient:

> **Formal education**
>
> Education provided in the system of schools, colleges, universities and other formal educational institutions that normally constitutes a continuous 'ladder' of full-time education for children and young people, generally beginning at age five to seven and continuing up to 20 or 25 years old. In some countries, the upper parts of this 'ladder' are constituted by organized programmes of joint part-time employment and part-time participation in the regular school and university system: such programmes have come to be known as the 'dual system' or equivalent terms in these countries.
>
> **Non-formal education**
>
> Any organized and sustained educational activities that do not correspond exactly to the above definition of formal education. Non-formal education may therefore take place both within and outside educational institutions, and cater to persons of all ages. Depending on country contexts, it may cover educational programmes to impart adult
>
> literacy, basic education for out-of-school children, life-skills, work-skills, and general culture. Non-formal education programmes do not necessarily follow the 'ladder' system, and may have differing duration.

For the purposes of questionnaire design, the operative difference between formal and non-formal education is the fact that formal education consists of some variation of a continuous, sequential, 'ladder' of education, whereas non-formal education consist of one or multiple non-continuous or non-sequential education opportunities. With formal education, there is a clear sequence of school levels and grades, with lower grades and levels serving as prerequisites for each higher grade or level. It is this hierarchy of grades and levels that makes it possible to ask the structured questions such as 'What is the highest school level NAME has attended?' since a household member who has attended pre-primary, primary, and secondary school can easily determine that secondary was the highest of those levels and respond appropriately.

The structure of questions used in to gather information about formal education is not appropriate for inquiring about non-formal educational opportunities because when levels do not have a clear hierarchy a question about 'what is the highest level NAME has attended' cannot be answered well. As an example, if

a household member's educational career includes adult literacy training, three years experience as an apprentice to a mechanic, and an after-hours training program on entrepreneurship, there is no objectively correct way to rank these experiences and report the 'highest' of them. Similarly, if a household member had attended formal primary school as a child, but also attended a non-formal literacy training program as an adult, there is no objectively correct way to report one or the other as higher.

Definitions of subcategories of non-formal education:

While there are clearly different types of non-formal education, the specific terms used to describe these types do not appear to have clear universal definitions and sometimes seem to be used interchangeably or with definitional overlap. For example "internship, apprenticeship, workplace training, professional training, skilled training" can be used to describe similar training experiences, as could " adult education, literacy training, continuing education, post-literacy training."

At the international level, there is no compelling reason to force specific definitions on most of the terms discussed above. Instead of focusing on specific definitions, we suggest wording questions broadly in order to gather information on non-formal education based on the functional output of the education rather than any label that might be applied to the education type. In other words, ask something along the lines of:

"Has NAME ever participated in any type of training or class outside of the regular school system with the intention of learning how to read or write?"

or

"Has NAME ever participated in any time of training or apprenticeship outside of the regular school system with the intention of learning a skilled trade?"

And not ask questions such as:

"Has NAME ever been an apprentice"

or

"Has NAME ever received any informal training"

However, if national or local policy makers have a compelling reason to be able to identify household members who have participated in a specific type of non-formal training, and are able to properly name or define that type of training in a way that respondents will be able to understand, then it would be fine to include a specifically-worded question about the training program. For example:

"Has NAME ever participated in a *Bi-Alpha* class to learn reading and writing?"

Precedence

Generally, questionnaires seem to group together questions about apprenticeships and other out-of-the-school-system trainings, but ask about literacy training elsewhere.

Literacy

Eight of the 30 surveys analyzed in 2008 include a question on attending or having attended some sort of a literacy program. Several of the questionnaires include follow-on questions on topics such as the duration, cost, or provided of the training

6. Has (NAME) attended a literacy course?	**Yes.....1** **No.....2** **(>> 8)**

← Five surveys use questions that are a very close variation on this example taken from the Nigeria 2003 LSS

This more carefully worded question from the 2002 Uganda National Household Survey questionnaire is filtered to apply only to respondents with less than a secondary education. →

For those with codes 01 - 17 in column (5) & aged 18 year & above
Have you ever participated in a literacy program that involves learning to read or write (not including primary school)?
Yes..... 1
No..... 2

Has ..[NAME].. ever attended non-formal class?	Is ..[NAME].. currently attending non-formal classes?	What kind of non-formal class is ..[NAME].. currently attending?
1 = Yes 2 = No (=>> 14)	1 = Yes 2 = No (=>> 14)	1= Literacy programmes (6 months) 2 = Vocational training (Tailoring, motor repairing, Khmer classical music training, hairdressing...etc. 3= Post literacy programmes (Agricultural training includes such as planting vegetable, mushrooms, raising fish, animal.. 4 = Others (Specify)

←The Cambodia questionnaire includes literacy as one of four possible responses to a question about the kind of non-formal class a person is attending.

At least two other questionnaires include literacy training as a response category in their general question on educational attainment.

Apprenticeship and/or other training outside the school system

Four of the 30 surveys include a question on being or having been an apprentice. Three are close variations on this example from the Ghana LSS questionnaire. The fourth includes 'Apprenticeship' as a response category in their general question on educational attainment.

The 2003 Nigeria LSS and 2003 Sierra Leone IHS include full sections of specific follow-up questions for respondents who are or have attended an apprenticeship. Survey designers interested in gathering extensive data on apprenticeships may look to these for examples.

Nigeria LSS:

6. Has (NAME) attended a literacy course? Yes.....1 No.....2 (>> 8)	7 For how long has (NAME) attended this course? YEARS / MTHS	8. Has (NAME) ever been an appren-ice? Yes, in the Past.....1 No.....2 (>>14)	9. How long was (NAME) an apparent-ice? YEARS / MTHS	10. What is the main trade (NAME) learnt? Carpentry.....1 Masonry.....2 Tailoring.....3 Black Smithng.....4 Mechanical....5 Electronics/ Electrical.....6 Painting/ Spraying.....7 Trading.....8 Hairdressing/ barbing.....9 Catering.....10 Other.....11 (SPECIFY)	11 Did (NAME) pay a fee for this training? Yes, in kind.....1 In cash.....2 Both.....3 N0.....4. (>>14)	12. How much did (NAME) pay for the training? AMOUNT	13 Who paid for the Training? Self.....1 Parent.....2 Other Relative.....3 Employer....4 Govt.....5 NGO.....6 Community Assoc.....7 Private Organizat-ions.....8 Other.....9 (Specify)	14 Has (NAME) attended other short training course(s) lasting not more than 6 months? Yes.....1 No.....2 (>> NEXT MEMBER)	15. What is the total number of months (NAME) attended such course(s) in the last 5 years? (Since 1998/99)	16 What was the main subject of the most recent training? Clerical.....1 Managerial...2 Computer.....3 Marketing.....4 Teaching.....5 Leadership....6 Medicine.....7 Farming.....8 Other .....9 (SPECIFY)

Mozambique:

(4 ans et plus)						
44-Avez-vous suivi une formation visant à apprendre un métier ? Oui 1 Non 2>>Q49	45-Combien de temps avez-vous été en formation pour la dernière fois ? en mois	46-A quelle branche de formation avez-vous été formé pour la dernière fois ? Voir Code Q46	47-Quel type de formation avez-vous suivi pour Q44? Voir Code Q47	48-Indiquer la deuxième branche de formation éventuelle apprise Voir Code Q48	49-Avez-vous participé à des stages de perfection-nement pour l'une ou l'autre des formations apprises ? Oui 1 Non 2 >> LS	50- Pendant combien de semaines en tout avez-vous suivi un stage pour la dernière fois ?

Sierra Leone IHS:

Code Q46 - Q48

01 : Agriculture
02 : Pêche
03 : Elevage
04 : Exploitation forestière
05 : Autres activités primaires
06 : Industrie agro-alimentaire
07 : Industrie alimentaire
08 : Industrie extractive
09 : Industrie textile et du cuir
10 : Industrie du bois
11 : Industrie chimique
12 : Industrie des matériaux de construction
13 : Industries diverses
14 : Energie
15 : BTP
16 : Commerce général
17 : Commerce spécialisé
18 : Autres commerces
19 : Transport privé (camions, taxi, ...)
20 : Hôtellerie - Restauration
21 : Santé privée
22 : Enseignement privé
23 : Sécurité privée
24 : Banques - Assurances
25 : Administrations publiques et parapubliques
26 : PTT privée
27 : Non classé ailleurs

Code Q47

Apprentissage avec :
1 : un membre du ménage
2 : un apparenté non membre du ménage
3 : un non apparenté non membre du ménage
4 : Dans une entreprise publique ou privée

Formation professionnelle dans :
5 : Dans une entreprise oublique ou privée
6 : dans un organisme d'état

Dans un école de formation spécialisée :
7 : publique
8 : privée

37. Has (NAME) attended a literacy Course? 1 = Yes 2 = No (>> 39)	38 For how long have you attended this Course		39 Why haven't you attended any literacy Course? 1 = None Available 2=Unsuitable Time 3 = No Nanny for Children 4 =Too much Household Chores 5= Other. (SPECEIFY)	40. Are you or have you been an apprentice? 1 =Yes, Currentl 2 = Yes in the Past 3 = No (>>46)	41. How long were you an apprentice?		42. What is the main trade you learnt? 1 = Carpentry 2 = Masonry 3 = Tailoring 4 = Black Smiting 5 = Mechanical 6 = Electronics/Electrical 7 = Painting/ Spraying 8 = Hairdressing / Catering 9 = Other..... (SPECIFY)
	YR	MT			YR	MT	

43. Who owns the Trade Centre you last attended? 1=Central Government 2=Local Government 3=Mission/Reli gious Body 4=Non-Govt. Organisation 5=Private 6=Other (Specify)	44. Did you pay a fee for this training? 1 = Yes, In kind 2 = Yes, In cash 3 = Both 4 = N0	45 Did you receive free room or board? 1 = Ye s 2 = No	46 Have you attended other short training courses lasting not more than 6 months? 1 = Yes 2 = No **(>> NEXT MEMBER)**	47 What is the total number of months you ever attended such course in the last 7 years (Since 1995)	48 What was the main subject of the most recent training? 1 = Clerical/Prof. 2 = Managerial 3 = Computer 4 = Marketing 5 = Teaching 6 = Leadership 7 = Medicine 8 = Other..... (SPECIFY)	49 Who owns the Training Centre you last attended? 1= Central Government 2=Local Government 3=Mission/Reli gious Body 4=Non-Govt. Organisation 5=Private 6=Other (Specify) .....

54. Avez-vous suivi une formation hors du système scolaire? OUI.1 NON.2**»»58**	55. Dans quel domaine de compé-tence ? (CF CODES)	56. Pourquoi avez-vous suivi cette formation ? APPRENDRE UN METIER..1 ETRE PLUS PERFORMANT DS ACTIVITE ACTUELLE...2 ETRE PLUS COMPETITIF.3	57. Où cette formation a-t-elle eu lieu ? PERS PHYS.1 CAB PRIVE.2 PME/GRANDE ENTR PRIV.3 GARAGE,ATEL MENUISERIE, ETC.....4 AUTRES....5

← The Cote d'Ivoire LSS questionnaire asks about training outside of the school system and then follows up with a sequence of questions on the subject of the training, reason for obtaining the training, and source of the training.

Recommendation:

Wording

The most practical way to gather information on non-formal education is through a sequence of questions that is independent of the core module on formal education, and asks separate questions about each type of non-formal education that is of interest to the questionnaire designers. Placing questions about non-formal education in a separate module has the additional advantage that it makes it possible to extend these questions to individuals beyond the CORE module age range of 3-30 years in order to gather information on the adult learners who are often targeted for non-formal programs.

To ensure it is clear what type of educational experience is being asked about in each question, we recommend that the wording of the question include a description of the function of the educational opportunity being asked about. These questions are modeled after the literacy question from the 2002 Uganda National Household Survey questionnaire:

"Now I want to ask you about different kinds of learning or training experiences that NAME may have had outside of the regular school system"

"Has NAME ever participated in a literacy program that involves learning how to read or write? (not including primary or secondary school)"

If YES 'Is NAME currently participating in this program?'

"Has NAME ever participated in an apprenticeship or a similar program that involved learning a specialized skill or trade while working for someone who performs the same skill or trade? (not including primary or secondary school)?"

If YES 'Is NAME currently participating in this program or period of learning?'

"Has NAME ever participated in any other program or training that involved learning about how to do job or skill or how to improve at a particular job or skill? (not including primary or secondary school)?"

If YES 'Is NAME currently participating in this program or period of learning?'

We don't see a need for the full set of follow-up questions on training or apprenticeships unless survey designers have a compelling reason to ask them. If questionnaire developers do want to gather detailed information on the nature of each type of training, the Nigeria LSS questionnaire seems to set a good precedent for these questions.

Universe

Most household surveys ask these questions of all household members above a minimum age of 3, 4, or 5 years. In keeping with the minimum age standard recommended for the Core modules, we recommend asking these questions of all respondents above the age of 3. The questions should be asked of all household members without regard for their literacy status or whether they have or are attending formal school.

Auxiliary Attainment: Highest Diploma Earned

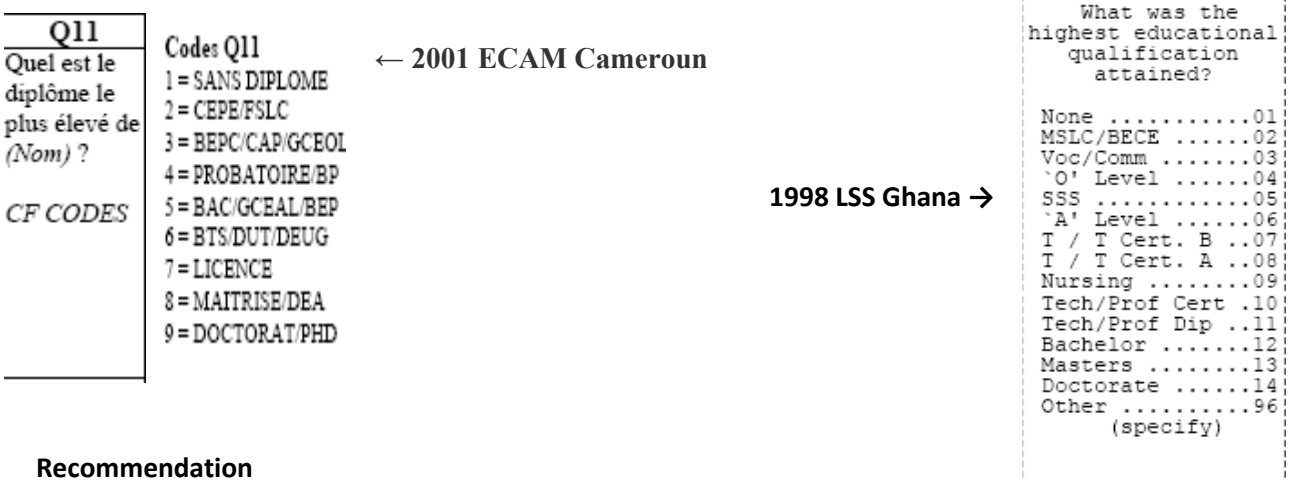
Eight questionnaires include some variation on a question asking about the highest diploma, certification, or qualification earned by a household member. For the current UN definition of Educational Attainment, this question does not replace the questions we recommend in the Educational Attainment but rather complement. Information gathered through this question can be used to contribute to a more nuanced understanding of educational attainment data by helping policymakers distinguish between household members who completed a level of schooling *and* completed the additional requirements (such as an exam) needed to earn a particular diploma or certification, and household members who may have completed the level of schooling but did not complete the additional requirements needed to earn the diploma. If the UN definition of Educational Attainment is revised in the near future, a question similar to

what is recommended here may be used to meet that new definition.

Policymakers may be interested in gathering information on the highest diploma earned in settings where a diploma requires more than simply completing the sequence of grades associated with a school level, and where the achievement of the diploma conveys meaningful information about an individual's human capital accumulation, job prospects, etc.

Precedence:

These questions are all structured similarly, with slight variations in the wording of the question. While most questions ask about the highest "diploma" earned, some ask about the highest "certification" or "qualification." The examples provided here are illustrative:



Q11	Codes Q11
Quel est le diplôme le plus élevé de *(Nom)* ? *CF CODES*	1 = SANS DIPLOME 2 = CEPE/FSLC 3 = BEPC/CAP/GCEO1 4 = PROBATOIRE/BP 5 = BAC/GCEAL/BEP 6 = BTS/DUT/DEUG 7 = LICENCE 8 = MAITRISE/DEA 9 = DOCTORAT/PHD

← **2001 ECAM Cameroun**

1998 LSS Ghana →

3
What was the highest educational qualification attained?
None 01
MSLC/BECE 02
Voc/Comm 03
`O' Level 04
SSS 05
`A' Level 06
T / T Cert. B ..07
T / T Cert. A ..08
Nursing 09
Tech/Prof Cert .10
Tech/Prof Dip ..11
Bachelor 12
Masters 13
Doctorate 14
Other 96
(specify)

Recommendation

The basic structure of this question should follow the precedent followed above, though questionnaire designers may want to adjust the specific wording of the question to keep it clear to their anticipated audience:

If Ever Attended = 'Yes'

What is the highest diploma or certification NAME has earned?

Questionnaire designers may choose to replace 'diploma or certification' with 'qualification' or some other term, and likewise may choose to replace 'earned' with 'achieved,' 'attained,' 'obtained, or some other similar term if they expect one of these other terms to be more meaningful to their survey respondents.

Response Categories

The list of response options for this question should be developed by the questionnaire developer since diploma or certification names vary from country to country. The list of response categories could include both academic degrees (ie: 'O-levels,' 'Ph.D.') and professional certifications (ie: 'teaching certificate,'

'nursing certificate') if desired, though questionnaire developers should follow the our guidelines for developing response categories.

Parents and Guardians

The EdData surveys include series of questions designed to gather information about parents' and guardians' attitudes towards education or involvement in education. While these questions appear to be soundly worded and the issues they explore are of potential value to some researchers and policymakers, we are hesitant to recommend them for inclusion in the IHSN question bank because of the questions specialized nature. This hesitation is reinforced by the fact that these questions do not appear in any of the surveys analyzed for our 2008 report, or in the standards DHS and MICs questionnaires. These questions are not recommended for the IHSN question bank, but have been included in the report as examples for questionnaire designers who may be interested in pursuing in-depth information on a specific education policy related topic. Questionnaire designers interested in pursuing these, or other similarly specialized lines of research would want to design and test questions that fit their research needs.

Parents attitudes towards education

EdData surveys include a series of questions directed at the parent or guardian of each child attending school. These questions are designed to gather information on parents attitudes towards education (such as the advantages and disadvantages of sending a child to school), and education related issues (such as whether corporeal punishment/motivation by teachers is justified). None of the household surveys analyzed in EPDC's 2008 report included questions like these.

Data collected through these questions are not as much of priority to the international community as the other indicators discussed in this report and are not maintained in international databases. The data collected through these questions may be of value to local/national policymakers who are interested in addressing cultural or values-based misalignments between school services and peoples' needs or values.

The EdData surveys filter these questions so that they are asked only of parents/guardians of children who are attending primary school (or primary and secondary in some cases). We believe that, from a policy perspective, it would be equally or more interesting to learn about the attitudes/values of parents/guardians of school-aged children who are *not* attending school, and suggest that if this family of indicators is included in a questionnaire, designers consider revising the universe of respondents to include parents/guardians of all school-aged children, regardless of whether those children are attending school.

Because there is very low precedence for the use of these questions, because we anticipate that the questions will be of interest to policymakers only under narrowly specific circumstances, and because questions of this nature should be tailored to address the specific context of each household survey, we will not recommend specific questions here. Instead, questions from two EdData questionnaires, Zambia (2002) and Egypt (2006), are provided for illustrative purposes. Because the information gathered through these questions would not used to calculate specific indicators but rather for exploratory purposes, we will not recommend specific code for data extraction.

714	Do you think girls in primary school should be taught by a male teacher, a female teacher, or does it not matter?	MALE . . . . . 1 FEMALE . . . . . 2 EITHER MALE OR FEMALE 3	
715	Do you think girls in preparatory and secondary school should be taught by a male teacher, a female teacher, or does it not matter?	MALE . . . . . 1 FEMALE . . . . . 2 EITHER MALE OR FEMALE 3	

716	Do you think boys in primary school should be taught by a male teacher, a female teacher, or does it not matter?	MALE . . . . . 1 FEMALE . . . . . 2 EITHER MALE OR FEMALE 3	
717	Do you think boys in preparatory and secondary school should be taught by a male teacher, a female teacher, or does it not matter?	MALE . . . . . 1 FEMALE . . . . . 2 EITHER MALE OR FEMALE 3	
718	Sometimes students cheat on exams. Do you think this is ever justified?	YES . . . . . 1 NO 2	
719	I am interested in knowing your opinions about what makes schools good and about the importance of schooling. Do you agree or disagree with the following statements? Whenever necessary, parents should keep their children home from school to work or help in the household. It is more important to send a boy to school than to send a girl to school. Primary schools should teach more practical skills, like carpentry or sewing.	AGREE DISAGREE DK 1 2 8 1 2 8 1 2 8	
720	I am interested in knowing what kinds of things you think affect the quality of schools. Does each of the following things make a school better, make a school worse, or have no effect on the quality of the school? Students being required to wear uniforms. Teachers hitting pupils to maintain discipline. Teachers hitting pupils to improve their performance. Parents being actively involved in the school. Crowded classes. Well-maintained buildings. Teachers making learning interesting.	BETTER / NO EFFECT / WORSE / DK 1 2 3 8 1 2 3 8 1 2 3 8 1 2 3 8 1 2 3 8 1 2 3 8 1 2 3 8	
721	Now I would like you to think about the benefits of primary school. Think of a 12-year-old boy who received his primary school certificate, and then dropped out of school. What advantages does this boy have compared to a boy of the same age who never attended primary school? PROBE: Anything else? RECORD ALL MENTIONED.	FIND (BETTER) JOB . . . . . A PROVIDE SUPPORT TO HOUSEHOLD/PARENTS . . . . . B CHANCE TO GO TO PREPARATORY . C LEARN TO READ AND WRITE . . . . . D LEARN OTHER LANGUAGES . . . . . E LEARN MATHEMATICS . . . . . F LEARN VOCATIONAL SKILLS . . . . . G DEVELOP MORALS/DISCIPLINE . . . . . H CRITICAL THINKING SKILLS . . . . . I MAKE A BETTER MARRIAGE . . . . . J LEARN TO BE A GOOD PARENT . . . . . K BETTER HYGIENE . . . . . L SOCIAL INTERACTION SKILLS . . . . . M OTHER _____ X (SPECIFY) NO BENEFITS . . . . . Y	

722	Now think of a 12-year-old girl who received her primary school certificate, and then dropped out of school. What advantages does this girl have compared to a girl of the same age who never attended primary school? PROBE: Anything else? RECORD ALL MENTIONED.	FIND (BETTER) JOB ..... A PROVIDE SUPPORT TO HOUSEHOLD/PARENTS ..... B CHANCE TO GO TO PREPARATORY . C LEARN TO READ AND WRITE ..... D LEARN OTHER LANGUAGES ..... E LEARN MATHEMATICS ..... F LEARN VOCATIONAL SKILLS ..... G DEVELOP MORALS/DISCIPLINE ... H CRITICAL THINKING SKILLS ..... I MAKE A BETTER MARRIAGE ..... J LEARN TO BE A GOOD PARENT ... K BETTER HYGIENE ..... L SOCIAL INTERACTION SKILLS ..... M OTHER _____ X (SPECIFY) NO BENEFITS ..... Y	

728	Now I would like you to think about the disadvantages of schooling. What are the disadvantages of sending a boy to primary school? RECORD ALL MENTIONED.	EXPENSIVE ..... A LOSE CHILD'S LABOUR ..... B BAD MANNERS ..... C NOT WILLING TO WORK ..... D MIGRATES FROM VILLAGE ..... E NO DISADVANTAGES ..... G OTHER _____ X (SPECIFY)	

729	What are the disadvantages of sending a girl to primary school? RECORD ALL MENTIONED.	EXPENSIVE ..... A LOSE CHILD'S LABOUR ..... B BAD MANNERS ..... C NOT WILLING TO WORK ..... D MIGRATES FROM VILLAGE ..... E LATER MARRIAGE/HARDER TO FIND HUSBAND ..... F NO DISADVANTAGES ..... G OTHER _____ X (SPECIFY)	

725	Now I would like to learn about how decisions are made in your household. More than one person may be involved in this decision, but who has the final say in your household about whether or not children attend school?	MOTHER ..... 01 FATHER ..... 02 BOTH PARENTS ..... 03 GUARDIAN(S) ..... 04 CHILD HIMSELF/HERSELF ..... 05 PARENT(S)/GUARDIAN W/ CHILD 06 SOMEONE ELSE _____ 96 (SPECIFY) DON'T KNOW 98	

Parents' involvement in education, schooling, the school

Many EdData surveys include a set of questions designed to gather information about the relative degree to which parents or guardians are actively involved in their childrens' education, school, or school activities. These questions appear to be soundly designed, but gather information on a narrow, specific policy area. Comparable questions do not appear in any other questionnaires investigated for this report. Because of the specialized nature of these questions, we do not recommend them for inclusion in the IHSN question bank. The questions are displayed here as an example to questionnaire designers interested

in developing similarly specialized questions for their surveys.

515	Do you or anyone else in the household frequently, sometimes or never help (NAME) with his/her homework?	FREQUENTLY ..... 1 SOMETIMES ..... 2 NEVER ..... 3 DON'T KNOW 8

515	Do you or anyone else in the household frequently, sometimes or never help (NAME) with his/her homework?	FREQUENTLY ..... 1 SOMETIMES ..... 2 NEVER ..... 3 DON'T KNOW 8

722	Have you or has any adult in your household attended a meeting of the PTA in the last 12 months?	YES ..... 1 NO 2

723	In the last 12 months, have you or has any adult in your household gone to a primary school for any of these reasons?		YES	NO
	For a school celebration, performance, or sports event.	EVENT	1	2
	For a meeting or conference with a head teacher or teacher.	MEETING	1	2
	To collect report forms.	REPORTS	1	2

709	Have you, one of your children, or anyone else in your household provided any of the following kinds of support to any school in the last 12 months?		YES	NO	DK
	Money to support school buildings, grounds or teacher housing.	MONEY	1	2	8
	Materials to support school buildings, grounds or teacher housing.	MATERIALS	1	2	8
	Labour to support school buildings, grounds or teacher housing.	LABOUR	1	2	8

710	In the last 12 months, have you, one of your children, or anyone else in your household provided any of these kinds of support to a teacher for the teacher's own use?		YES	NO	DK
	Money, other than for private tuition.	MONEY	1	2	8
	Food.	FOOD	1	2	8
	Labour, other than for maintenance of teacher housing.	LABOUR	1	2	8