



# School Sampling Preparation Guide

**Assessments for Minimum Proficiency Levels (AMPL) at  
17 November 2022, Version 1**

The Global Education Monitoring (GEM) Centre drives improvements in learning by supporting the monitoring of educational outcomes worldwide. The GEM Centre is a long-term partnership between the Australian Council for Educational Research (ACER) and the Australian Government's Department of Foreign Affairs and Trade (DFAT).



# Contents

AMPLab Sampling Preparation Guide .....	1
Introduction .....	1
Sampling Preparation Process .....	1
AMPLab Sampling Framework .....	3
The Target Population .....	3
Sampling Form 1: National Target Population .....	3
Reduced Coverage .....	7
Exclusions .....	7
School-level exclusions .....	7
Student-level exclusions .....	7
Implications of Reduced Coverage and Exclusions .....	8
Sampling Form 2: Coverage and Exclusions .....	8
Stratification .....	10
Overview .....	10
Explicit and Implicit Stratification .....	10
Stratification Guidelines .....	10
Sampling Form 3: Stratification .....	11
Sample Form Submission .....	12
Preparing the Sampling Frame .....	12
Sampling Frame Template .....	13
National School ID .....	13
School contact information .....	13
Stratification variables .....	14
Measure of school size .....	14
School exclusions .....	16
Sample frame submission .....	16
Final school sampling steps .....	16



# AMPLab Sampling Preparation Guide

## Introduction

The Assessments for Minimum Proficiency Levels ab (AMPLab) project aims to measure learning outcomes in reading and mathematics in order to understand and describe how many children at the end of primary school

- meet the Minimum Proficiency Levels (MPL) for SDG4.1.1b (end of primary school)
- do not meet the MPL for SDG4.1.1b but meet the MPL for SDG 4.1.1a (end of lower primary school)
- do not meet the MPL SDG 4.1.1a.

Participating countries can also choose to:

- Administer only the AMPLb at the end of primary school, which will describe how many students meet the Minimum Proficiency Levels (MPL) for SDG4.1.1b.
- Administer only the AMPLa at the end of lower primary school, which will describe how many students meet the Minimum Proficiency Levels (MPL) for SDG4.1.1a.

The *AMPLab School Sampling Preparation Guide* is designed to support National Project Managers (NPMs) through the process of aligning their national target population definition to the sampling framework, considering the sample design best suited to their country, and preparing the national sampling frames.

A series of sampling forms have been created to support NPMs as they work through the sampling preparation process. These sampling forms are described throughout this manual and provide illustrative examples of how to perform the various steps.

This guide is intended for the members of the NC team responsible for the sampling activities within their respective countries. For further details on roles and responsibilities please refer the *AMPLab Sampling Framework*.

## Sampling Preparation Process

The main steps required to prepare for national sampling are summarised in the table below:

<b>Step</b>	<b>Description</b>	<b>Supporting Resources</b>
1	Describe the national target population	Sample Form 1: National Target Population
2	Describe national coverage and exclusions	Sample Form 2: Coverage and Exclusions
3	Determine stratification variables	Sample Form 3: Stratification
4	Prepare a sampling frame	Sampling Frame Example

If you plan to participate with two target grades (using AMPLa and AMPLb or AMPLa+b) you will need to repeat this process for each target grade (that is, make a copy of the forms and complete two). It is important that as you work through the sampling preparation process, key stakeholders within the country are consulted about sample design decisions with respect to this assessment.

Please note that the space provided on the forms should not limit responses. It is useful to include more information rather than less. The space will expand with more content.

# AMPLab Sampling Framework

Sampling for this project is guided by the *AMPLab Sampling Framework*, which outlines the regional target population definition and a set of 12 sampling standards<sup>1</sup>. The standards aim to achieve precise estimates of student learning outcomes, enabling high quality, consistent data across the participating countries, and over time.

The sampling framework represents a commitment by countries to aim for high standards of participation and therefore careful planning will be required to ensure that these standards are achieved. During the sampling preparation phase, NPMs should regularly refer to these standards. Where standards are not achieved, the results will be annotated within international reporting to help readers evaluate the quality of outcomes.

## The Target Population

The target populations for AMPLab are as follows:

All students enrolled in the grade that corresponds to the final year of primary school, where the language of instruction corresponds with the language of the assessment.

and/or

All students enrolled in the grade that corresponds to the end of lower primary school, where the language of instruction corresponds with the language of the assessment.

All students enrolled in the target grade(s) in each participating country are included in the target population. This includes students from schools across all educational sub-systems and types within a country.

### Sampling Form 1: National Target Population

The purpose of **Sampling Form 1** is to define the national target population and detail contextual information related to the participating country's education system and any previous national sample-based studies that have occurred.

---

<sup>1</sup> Refer the *AMPLab Sampling Framework* to view the sampling standards.

The national target population must align with and comprehensively address the requirements of the project target population definition. Every effort should be made to ensure complete coverage when defining the national target population.

The following provides an example of Sampling Form 1, using a model country (Country ABC) as a reference, to illustrate the type and level of information needed to complete the form. Further detail about the information required for completing Sampling Form 1 is below.

Question 1 asks the name of the grade level that is intended to be used as the target population for the survey which scenario the country will use for the AMPLab study using AMPLa, AMPLa+b or AMPLb. If a combination of these options is to be used, please complete a form for each option.

Questions 2-3 ask about the language of the assessment and the target grade for the study. Questions 4 and 5 ask some additional questions about the target grade and how it corresponds with SDG1.1.4a and b reporting.

When responding to Question 6, if available, please provide an estimate of:

- The total enrolment for the target grade across the nation
- The number of schools
- Average class sizes
- The number of regions or states where eligible students are enrolled
- A brief description of the breakdown of the education system in the country relevant to the target grade that could be relevant for the survey, for example, the proportion of public/private schools, languages spoken

Please reference the source(s) used to obtain the above data in Question 7.

Question 8 asks if the AMPLab study will occur at the same time as another national educational assessment or survey. We ask for details of this survey and whether you wish for the school sample to overlap with the AMPLab sample. Included in the response should be any details available of the sample design such as:

- The school and student sample sizes;
- How the sample is clustered (for example, whole classes, or randomly chosen students within a grade irrespective of class);
- Stratification variables (for example, if students are sampled from all regions of the country, or from private and government funded schools);

Question 9 asks about the maximum target cluster size you wish to use for this project.

Question 10 relates to the start and end date of the school year in which the testing period will occur (i.e. 2023). Then detail the actual testing period in Question 11.



Figure 1: Sampling Form 1 Example

AMPL	
Sampling Form 1	National Target Population
AMPL Participant:	Country ABC
National Project Manager:	Jane Doe
Date form completed:	30 November 2022
<p><b>1 Which scenario option do you intend to use for this survey?*</b></p> <p>*If you intend to participate in more than one scenario, please complete a second set of sampling forms</p>	<p>Scenario 1: Administration of AMPLb (end of primary) <input checked="" type="checkbox"/></p> <p>Scenario 2: Administration of AMPLa+b (end of primary) <input type="checkbox"/></p> <p>Scenario 3: Administration fo AMPLa (end of lower primary) <input type="checkbox"/></p>
2 State the language in which the survey will be administered.	English
3 State the name of the grade you intend to use for the target population for the AMPL survey.	Grade 6
4a If you are participating in Scenario 1 or 2, state the name of the grade in your country that corresponds to the final year of primary school.	Grade 6
4b If you are participating in Scenario 3, state the name of the grade in your country that corresponds to the end of lower primary.	N/A
5a State the name of the grades which have previously been used to report against SDG 4.1.1a.	N/A
5b State the name of the grades which have previously been used to report against SDG 4.1.1b.	Grade 6
<p><b>6 Briefly describe the size and distribution of the target population, that is eligible for AMPL.</b></p> <p>There are approximately 1200 primary schools in Country ABC. Schools are either public (75%) or private (25%). Schools are spread across 6 provinces in the following proportions: Province A (33%), Province B (25%), Province C (18%), Province D (8%), Province E (10%) and Province F (6%). There are approximately 32,000 students in grade 6 across these provinces. The average class has 25-30 students in it.</p>	

.....

**7 Provide any references or links to documentation used to estimate the eligible population.**

Numbers were obtained from the national schools database in the department of education.

.....

.....

**8 Will the AMPLab study occur at the same time as another national education survey? Please describe any this assessment or survey and provide a brief description of any sample design for these assessments, including if you wish for the sample to overlap with AMPLab.**

We are planning a sample for a national survey at Grade 5 this year. We would be interested in having this sample overlap with the AMPLab sample. Let's discuss the requirements further. This sample is of 200 schools across the country but is still being designed.

.....

**9 What is the maximum number of students per class (maximum target cluster size (TCS)) you wish to sample per school? When the size of a sampled class exceeds this size, a sub-sample of the students from the sampled class will be selected.**

35

**10 Please specify the start and end dates of the school year.**

Start of School Year (DD-MM-YYYY)	End of School Year (DD-MM-YYYY)
25-Jan-23	24-Dec-23

**11 Please specify the start and end date of the expected testing period.**

Start of Testing Period (DD-MM-YYYY)	End of Testing Period (DD-MM-YYYY)
01-Jun-23	30-Jun-23

## Coverage and Exclusions

The aim of AMPLab is to provide complete coverage of the national target population. However, a small number of exclusions at school or student level are sometimes necessary, and in rare cases there may be reductions in coverage (for example a region recently affected by a major flood). For each participant, exclusions and reductions in coverage will be documented and quantified and will form part of the regional reporting.

A 5% threshold has been adopted as the upper limit for the exclusion of members of the survey population, either as a result of excluding schools or excluding individual students within schools. The objective is to minimise exclusions as much as possible.

The following outlines conditions under which reduced coverage and exclusions might occur.

## Reduced Coverage

In rare cases it might be necessary to reduce coverage to less than the full national target population due to political, operational or administrative reasons. For example, particular schools or geographic regions might be removed from the target population in the event of a natural disaster or political unrest.

## Exclusions

Exclusions can occur at a school or student level. School-level exclusions refer to whole schools being excluded from the sample while student-level exclusions, refer to specific students within sampled schools being excluded.

### School-level exclusions

Schools are generally excluded from the survey for practical reasons, such as increased survey costs or difficult survey conditions. Examples of school-level exclusions include:

- very remote locations (inaccessibility)
- very small schools (for example, less than 5 students in the target grade)
- international schools (offering a curriculum other than the prescribed national curriculum)
- schools catering exclusively to students who are considered student-level exclusions

### Student-level exclusions

The aim of AMPLab is to be as inclusive of as many students as possible, however, some students from sampled schools may be unable to access the assessment and will therefore need to be excluded. The international student-level exclusion categories are defined below:

- Students with functional disabilities—these are students who have permanent physical disabilities in such a way that they cannot take the AMPL test. Students with functional disabilities who can perform the assessment should be accommodated in the test situation, within reason, rather than excluded.
- Students with intellectual disabilities—these are students with a cognitive, behavioural, or emotional disability confirmed by qualified staff, such that they cannot take the AMPLab assessment. These are students who are cognitively, behaviourally, or emotionally unable to follow even the general instructions of the assessment. Students should not be excluded solely because of poor academic performance or normal disciplinary problems. It should be noted that students with dyslexia, or other such learning disabilities, should be accommodated, within reason, rather than excluded
- Students with insufficient assessment language experience—these are students who cannot read or speak the language(s) of the assessment and would be unable

to overcome the language barrier. Such students meet **ALL** the following three criteria:

- they are not native speakers of the assessment language(s),
- they have limited proficiency in the assessment language(s), and
- they have received less than one year of instruction in the assessment language(s).

## Implications of Reduced Coverage and Exclusions

Any exclusions and reductions in coverage need to be well-documented and quantified during the sampling preparation process.

Significant reduced coverage and exclusions from the national target population mean that the survey results will no longer be nationally representative. Therefore, the objective is to minimise the total number of exclusions. A 5% threshold has been adopted as the upper limit for combined exclusions from the survey population.

If the total of the school and student exclusions exceeds 5% of the national target population, this will be annotated in the final reports.

## Sampling Form 2: Coverage and Exclusions

Sampling Form 2 lists the school- and student-level exclusions from the national target population.

Student-level exclusions may not be known in advance. But to avoid having outcomes annotated due to excessive exclusions, it is necessary to at least estimate these in advance of sampling, for example, by applying a rate calculated from a previous survey. If you expect extremely few or no such exclusions, note this on the form.

Below is an example of a completed Sampling Form 2.

Figure 2: Sampling From 2 Example

AMPL															
Sampling Form 2	Coverage and Exclusions														
AMPL Participant:	Country ABC														
National Project Manager:	Jane Doe														
Date form completed:	30 November 2022														
Total enrolment in the target grade	[a] 32,512														
1 School-level exclusions															
<table border="1"> <thead> <tr> <th>Description of exclusions</th> <th># of students</th> </tr> </thead> <tbody> <tr> <td>Schools affected by natural disaster</td> <td>206</td> </tr> <tr> <td>Very remote schools in province D</td> <td>105</td> </tr> <tr> <td></td> <td></td> </tr> <tr> <td></td> <td></td> </tr> <tr> <td></td> <td></td> </tr> <tr> <td><b>TOTAL</b></td> <td><b>311</b></td> </tr> </tbody> </table>	Description of exclusions	# of students	Schools affected by natural disaster	206	Very remote schools in province D	105							<b>TOTAL</b>	<b>311</b>	[b]
Description of exclusions	# of students														
Schools affected by natural disaster	206														
Very remote schools in province D	105														
<b>TOTAL</b>	<b>311</b>														
2 Percentage of school-level exclusions	1.0%														
3 Estimated student-level exclusions (if applicable)															
<table border="1"> <thead> <tr> <th>Description of exclusions</th> <th># of students</th> </tr> </thead> <tbody> <tr> <td>Students with special needs</td> <td>57</td> </tr> <tr> <td></td> <td></td> </tr> <tr> <td></td> <td></td> </tr> <tr> <td></td> <td></td> </tr> <tr> <td><b>TOTAL</b></td> <td><b>57</b></td> </tr> </tbody> </table>	Description of exclusions	# of students	Students with special needs	57							<b>TOTAL</b>	<b>57</b>	[c]		
Description of exclusions	# of students														
Students with special needs	57														
<b>TOTAL</b>	<b>57</b>														
4 Expected percentage of student-level exclusions	0.2%														
5 Expected percentage of reduced coverage and exclusions:	[d] 1.1%														

# Stratification

## Overview

Stratification is a process of organising the sampling frame to improve the efficiency of the sample design, resulting in more precise and reliable survey estimates, and ensuring proportional representation of specific groups within the national target population. Stratification can lead to improved reliability of estimates, provided that the variables are related to the major survey outcomes and national objectives.

## Explicit and Implicit Stratification

Stratification variables can be based on a range of contextual variables. There are two types of stratification, explicit stratification and implicit stratification.

Explicit stratification involves partitioning the sampling frame into mutually exclusive parts according to certain variables, from which separate, independent samples are drawn. Explicit stratification allows different sample designs to be applied to individual explicit strata.

Implicit stratification involves sorting the sample frame by the implicit stratification variables and then systematically sampling from the sorted list. The purpose of implicit stratification is to ensure a proportional sample allocation across all implicit strata and to improve the reliability of survey estimates.

## Stratification Guidelines

Stratification variables are identified by the NC and determined following consultation with ACER. The following guidelines should be applied when identifying stratification variables:

- selection of two to four stratification variables is usually adequate
- data must exist for each stratification variable with a valid entry for each school within the national target population
- avoid defining very small strata, as this is unlikely to improve the overall level of sampling precision
- variables should be related to key characteristics of the AMPLab survey outcomes.

Examples of stratification variables include:

- Geographic region (state, province, district);
- Sector (public, private);
- Urbanisation (urban, rural);

- Gender (male, female, coeducation);
- Socio-economic status (low, medium or high income areas);
- School type (primary, middle school);
- School-level performance measures (e.g. school average score on a national assessment)

## Sampling Form 3: Stratification

The purpose of **Sampling Form 3** is to outline the stratification variables that will be used for AMPLab.

The first task when completing Sampling Form 3 is to list, in order of importance, each of the proposed stratification variables. Determination of final explicit and implicit stratification variables will be finalised during discussions between the NC and ACER following submission of the form.

Provide any further details about the proposed stratification variables that you believe may be of relevance to the sample design or that you wish to discuss with ACER in the section below the table.

This section provides an example of Sampling Form 3.



Figure 3: Sampling Form 3 Example

**AMPL**

**Sampling Form 3** **Stratification**

AMPL Participant: Country ABC

National Project Manager: Jane Doe

Date form completed: 30 November 2022

**1 List and describe proposed variables to be used for stratification in their order of importance. For each variable list the number of levels.**

Stratification Variables				
	Variable name	Description	Value labels	# of levels
	Province	Political division	A/B/C/D/E/F	6
	Sector	School governance type	Public/Private	2
	Location	Urbanisation	Urban/Rural	2
Additional Stratification Information				
We may have a national examination for grade 5 at the same time and it would be good to explore overlap between these samples				

## Sample Form Submission

The first draft of the sampling forms are submitted to ACER securely via your National Team private channel, within the project ‘Microsoft Team’ titled ‘AMPL Partner Collaboration’. Once sampling forms have been submitted, ACER will review the information provided and prepare a summary report containing any queries or clarifications. Before finalising the sampling forms, an iterative consultation process will be conducted between the NCs and ACER to address any queries and clarifications.

## Preparing the Sampling Frame

The next step following the finalisation of the sampling forms is to prepare the sampling frame. The sampling frame is a list of all schools in the country serving the target population, along with the number of students enrolled in the target grade and other information about those schools. The contents of the frame are described in detail in the following sections of this guide. The sampling frame will be prepared at the local level and will be used by ACER to select the sample.



A list of all eligible schools in the national target population will be required for sampling. Identifying and accessing the most reliable and up-to-date database can take a considerable amount of time, so it is recommended that NCs plan these tasks well ahead of schedule. If you perceive any difficulties in being able to access a suitable database of schools, please discuss this with ACER early in the planning process.

The sampling frame should provide complete coverage of the defined target population, without containing incorrect entries, duplicate entries or entries that do not belong to the defined target population.

## Sampling Frame Template

The sampling frame will be prepared using Microsoft Excel. A template has been provided to assist NCs to construct their national sampling frame. The template includes a school frame template and a school exclusions template, as well as variable formatting instructions.

The sampling frame template contains the following fields for each school entry:

- a unique National School Identifier (ID)
- school contact information
- stratification variables
- measure of school size

## National School ID

The unique National School ID is the identification code used for each school within the participating country. Most countries have a National School ID and will use this to complete the sampling frame. A unique National School ID is essential to uniquely identifying schools during sampling.

## School contact information

School contact information will be used after the sample is drawn as part of the school liaison process to facilitate identification and location of the sampled schools as well as to communication with the sampled schools. School contact information may include the following:

- School name
- Street address
- Phone number
- Email address

School contact information is not required for sampling, therefore for data protection purposes the NC may choose to provide ACER with a reduced version of the sampling

frame in which this information is omitted. Each NC should be familiar and comply with their own country's data protection requirements.

### **Stratification variables**

Stratification variables must be provided each with a value for each school consistent with those outlined in Sampling Form 3. Each school in the sampling frame should only belong to one category of each stratification variable. For example, a school can be classified as urban or rural, not both.

### **Measure of school size**

A suitable measure of size (MOS) is an essential component of the school sampling frame since school selection probabilities are based on these values. The most suitable MOS is the enrolment (ENR) of students in the target grade within each school. In cases where target grade ENR is not available, other measures of size can be used, for example, the total school enrolment. The same type of MOS should be used for each school. Where none of these MOS figures are available an alternative sampling plan will be determined in consultation with ACER.

Up-to-date and accurate enrolment data can greatly improve the precision of the sample, however it is understood that even the most up-to-date figures may vary from the actual number of eligible students at the school when this information is collected. The actual number of eligible students will be recorded as part of the within-school sampling process and where this number varies significantly from the estimate these schools will be flagged for follow up checking.

Figure 4: School Sampling Frame Example

National School ID	School Name	Address	Town	Province	Postcode	Phone	Email	Target Grade ENR	Target Grade Classes	Total School ENR	Sector	Location
ABC0001	Broad River School	Roseda-Tinamba Road	BROAD RIVER	A	3001			36	2	216	Public	Urban
ABC0002	Broad River College	Shaw Drive	BROAD RIVER	A	3001			13	1	78	Private	Urban
ABC0004	Eastview Institute	Sale-Heyfield Road	EASTVIEW	B	4251			48	2	288	Private	Rural
ABC0005	Eastview Secondary School	Railway Avenue	EASTVIEW	B	4251			11	1	66	Private	Rural
ABC0006	Elk Grove High School	Carlisle Street	ELK GROVE	C	5418			21	1	126	Public	Urban
ABC0007	Forest Lake School	Crofts Road	FOREST LAKE	D	4115			56	2	336	Public	Urban
ABC0008	Golden Sierra School	Bourke Crescent	TARRANTYURK	E	0868			31	2	186	Private	Urban
ABC0009	Greenville College	Armstrong Street	GREENVILLE	F	6173			15	1	90	Public	Urban
ABC0010	Greenville Grammar School	Acheron Road	GREENVILLE	F	6173			60	2	360	Private	Urban
ABC0011	Hawking Institute	Redesdale Rd	SEDGWICK	A	2222			71	3	426	Public	Urban
ABC0013	Mountain Oak Institute	Cherokee Road	MOUNTAIN OAK	A	2132			90	3	540	Private	Rural
ABC0014	Northride School for Girls	Romawi Road	NORTHRIDE	A	2603			76	3	456	Public	Urban
ABC0015	Oakland School	Corio Street	OAKLAND	D	6339			30	1	180	Private	Urban
ABC0016	Pine Hill Academy	Edmundsons Road	PINE HILL	A	2558			15	1	90	Public	Rural
ABC0017	Riverdale School for Boys	Yarra Street	RIVERDALE	B	7541			79	3	474	Public	Urban
ABC0018	Riverdale Secondary School	Swanston Street	RIVERDALE	B	7541			24	1	144	Public	Urban

## School exclusions

Schools excluded in accordance with Sampling Form 2, should be identified on the sampling frame and removed prior to sampling. These schools and their enrolment sizes should be recorded on the School Exclusions sheet attached to the Sampling Frame Template, along with their reason for exclusion. The sampling frame should not contain excluded schools.

Figure 5: School Exclusions Example

National School ID	School Name	Target Grade ENR	School ENR	Reason for Exclusion
ABC0003	Glenn Russel Special School	21	147	Special education school
ABC0012	Metcalfe High School	3	21	Very remote school

## Sample frame submission

Once the sampling frame has been prepared, the Excel file should be uploaded as instructed. Once the sampling frame has been submitted, ACER will review the data provided. Before finalising the sampling frame an iterative consultation process will again be conducted between the NC and ACER to address any queries and clarifications.

## Final school sampling steps

Following the preparation of the sample frame, the school sample will be selected. The list of sampled schools will be shared with the NPM to ensure that no errors or misunderstandings have occurred. Once agreed, the sample will be passed to the data management team for implementation through ACER Maple.