



**USAID**  
FROM THE AMERICAN PEOPLE

# MEASURING AND COMMUNICATING PROGRESS TOWARDS GOAL ONE

## METHODOLOGY FOR THE USAID EDUCATION STRATEGY

OCTOBER 2015

## Executive Summary

This document describes the methodology USAID will use to measure progress towards the Goal One target of the USAID Education Strategy: improved reading for 100 million children in primary grades by 2015. This methodology replaces the methodology that was outlined in the [Education Strategy Technical Notes](#). The new methodology is compatible with existing implementation and data collection guidance. Partners and implementers whose work is aligned with previous guidance and who follow rigorous practices in the measurement of student learning do not need to – and should not – change their practices in response to this methodology.

As always, the ultimate objective of USAID Goal One programming is to ensure that all children are able to read with fluency and comprehension by the end of their second year of primary school and read at grade level by the end of the primary cycle. This ambitious objective can be achieved over an extended timeframe. The target of “100 million children with improved reading” provides insight into the pace of progress towards our ultimate objective.

The new Goal One methodology is premised on the comparison of the reading ability of successive cohorts of grade 2 students. If a USAID-supported program targets a particular set of schools, the reading ability of a sample of children attending grade 2 at the end of the program is compared with the reading ability of a sample of children attending grade 2 before the program began. The percentage of endline scores that exceed baseline scores by a defined amount is calculated, and this figure is interpreted as representing the percentage of students with improved reading ability.

The majority of USAID programs use the oral reading fluency (ORF) measure of the number of correct words read per minute (CWPM) to assess student reading ability. This is based on evidence that there is a close relationship between ORF and reading comprehension. For programs that use ORF, USAID has defined two increments of change that will be counted as indicative of improved reading ability. The increment varies according to the reading ability of baseline cohort members:

- For members of the baseline cohort who cannot read a single word (0 CWPM), an endline improvement of at least one correct word per minute (1+ CWPM) over the baseline score is indicative of improved reading ability.
- For members of the baseline cohort who can read at least one word (1+ CWPM), an endline improvement of at least ten correct words per minute (10+ CWPM) over the baseline score is indicative of improved reading ability.

For the purposes of the Education Strategy, USAID will apply these increments of change uniformly across the diverse contexts and languages reached through our Goal One programming. Because greater nuance is possible at the national, subnational, or program levels, these increments must not take the place of contextually appropriate norms and targets developed in consultation with local stakeholders.

The methodology described will be applied to all Goal One programs initiated over the life of the 2011-2015 USAID Education Strategy and to programs initiated beyond the 2011-2015 Strategy. In future years, should large scale national assessments be implemented as a means for measuring early grade reading ability, USAID will examine how these efforts can be integrated into existing accountability and reporting structures.

## Contents

Executive Summary .....	2
Background and Introduction.....	3
Methodology .....	4
Data Required for the Methodology .....	11
Communicating Goal One Outcomes .....	15

## Background and Introduction

In February 2011, the USAID Office of Education announced a bold five-year Education Strategy designed to ensure that the Agency’s global education investments would be grounded in evidence and aimed at maximizing the impact and sustainability of development results. The [2011-2015 Education Strategy](#) identified early grade reading as a critical outcome for developing and sustaining learning for children, and USAID committed to improving reading skills for 100 million children in early grades by 2015. USAID also committed to measure, analyze, and share progress towards this target, recognizing that doing so would increase accountability through the disclosure of information about education program effectiveness, relevance, and efficiency.

In the [Education Strategy Technical Notes](#) that accompanied the 2011-2015 Education Strategy, the USAID Office of Education detailed the data needed under Goal One of the Strategy and recommended a methodology for measuring progress towards the target of improved reading for 100 million children. As Missions began to generate data in keeping with this guidance, however, it became clear that the methodology identified at the inception of the strategy did not satisfy the Agency’s needs. Recognizing that a new methodology was required, the Office of Education launched a consultative process to identify a new methodology for measuring progress towards Goal One of the USAID Education Strategy.

This document outlines the revised methodology that USAID will use as the basis for estimating the number of children with improved reading skills in connection with USAID programming and partnerships.

**USAID’s ultimate objective is all children reading with comprehension and fluency. The purpose of this methodology is to measure progress towards the target of “improved reading skills for 100 million children in primary grades.”**

It is critical to distinguish between the number of children reading with fluency and comprehension, which consists of learners who have reached a particular ability level, and the number with improved reading skills, which consists of learners demonstrating meaningful positive change in their reading ability level. Learners can show meaningful positive change regardless of whether this change brings them across the threshold of reading with fluency and comprehension. Some learners who make meaningful gains in reading ability may begin below the threshold and remain below the threshold despite their gains; others may begin above the threshold and progress further beyond it; still others may begin below the threshold and make meaningful gains that carry them beyond the threshold. The USAID Education strategy counts learners in all three of these scenarios as contributing to the Goal One target of “learners with improved reading ability.”

The primary purpose of this methodology is to count progress towards the target of improved reading skills for 100 million children in primary grades, but it is important to keep this count within the context of the

ultimate objective of all children reading with fluency and comprehension. To maintain this perspective, USAID will also track: 1) the number of children who gain the ability to read with fluency and comprehension and 2) the number of children who gain the ability to read their first word. These are separate from the top-line count required for the Education Strategy and do not replace or change the need to perform this count.

At all levels, USAID Goal One programming should be designed to achieve the ultimate goal of children reading with fluency and comprehension. Programming must not be designed around the methodology for the Goal One count.

### **What changes and what does not change**

The **2011-2015 Education Strategy** and **Education Strategy Technical Notes** establish the terms of USAID's focus on primary grade reading and define the data that missions must collect in support of Goal One of the Education Strategy. This methodology document does not change USAID guidance on the implementation of the strategy or on the data that must be collected under the strategy.

The methodology outlined in this document replaces the methodology that was outlined in the Education Strategy Technical Notes. The new methodology is designed to be compatible with existing implementation and data collection guidance. Partners and implementers whose work has been aligned with 2011 guidance, and who have been reporting in keeping with the [2014 Update to Education Strategy Reporting Guidance](#), will not need to do anything differently in order to support the new methodology.

Missions are no longer responsible for calculating program contributions to the Goal One target through the process described in Education Strategy Technical Notes. The Office of Education will manage the process of calculating progress towards the Goal One target. Missions are responsible for ensuring the appropriate data are produced to perform this task.

Missions are still responsible for reporting on Standard Foreign Assistance Indicators 3.2.1-27 (Proportion of students who, by the end of two grades of primary schooling, demonstrate that they can read and understand the meaning of grade level text), and 3.2.1-28 (Proportion of students who, by the end of the primary cycle, are able to read and demonstrate understanding as defined by a country curriculum, standards, or national experts). These indicators are calculated independently of the methodology described here, and should be included in performance monitoring plans as relevant.

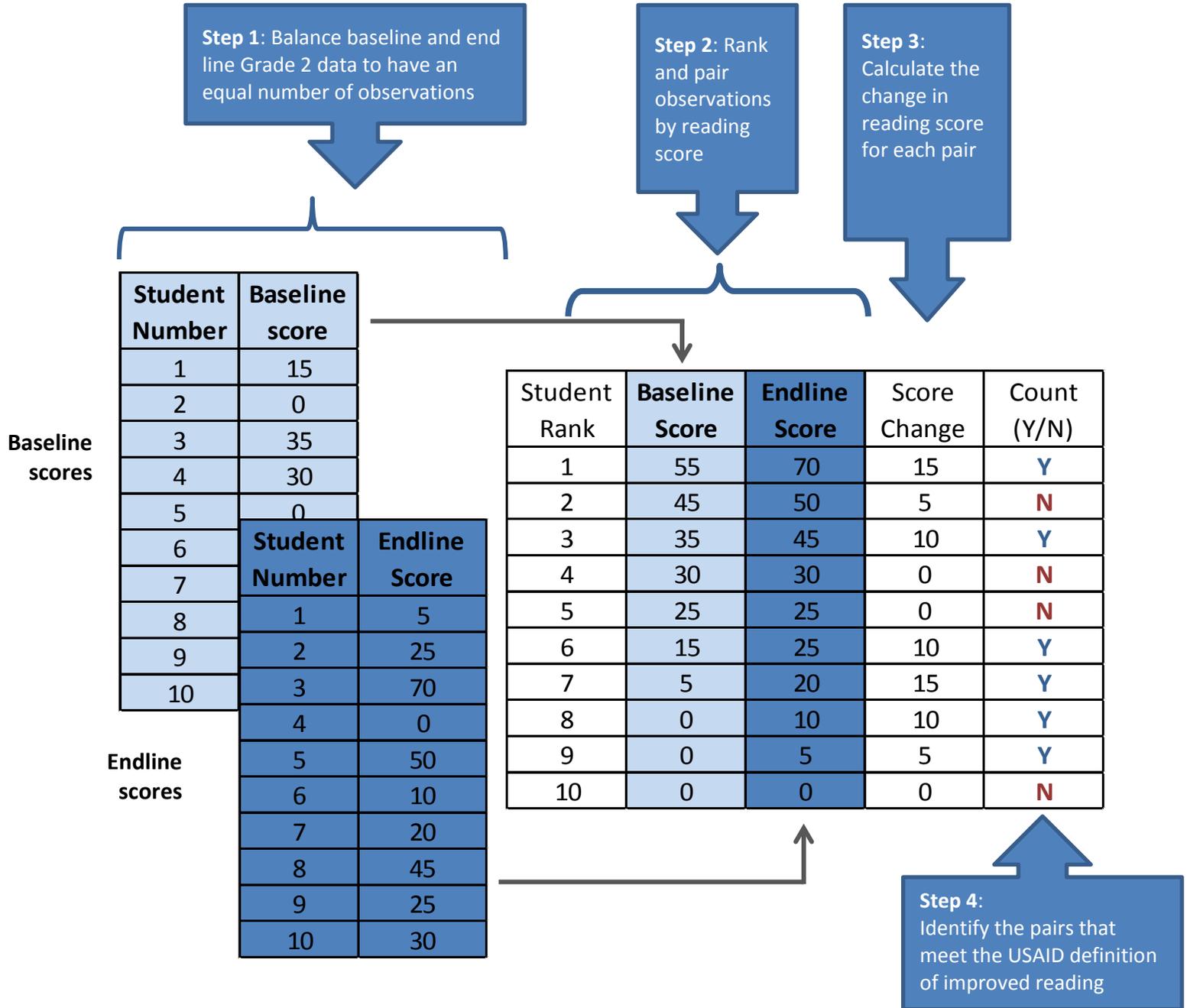
## **Methodology**

The methodology is described in this section for application using data that have been collected in a manner consistent with previous USAID guidance. The methodology is also based on the use of Oral Reading Fluency (ORF) as the measure of reading ability. This will be the most common approach to calculating progress towards the Goal One target, but there will be deviations to it. Some of these deviations are in keeping with previous USAID guidance – some programs have opted to use alternatives to ORF as the measure of reading ability. Other deviations are not in keeping with previous USAID guidance – some programs have measured learning outcomes at the beginning and end of a single school year rather than at the ends of two consecutive school years.

Whenever E3/ED encounters program data for which deviations to the methodology are needed, E3/ED will document any deviations necessary to estimate the count contribution of the program. Given that this will be assessed in a case-by-case basis, E3/ED will not discuss deviations from the methodology in this document.

## Methodology Overview

The new Goal One methodology can be described as a sequence of six steps. The steps are illustrated here and described in greater detail in the pages that follow.



This process is repeated separately for male and female students so that progress towards the target can be disaggregated by sex. For programs that are segmented by treatment (e.g., different groups of students receive partial or full treatment) and/or language (i.e., different groups of students receive the program intervention in different languages), the process is also repeated by treatment and/or language groups.

If there is only a baseline and endline, the count contribution of a program will be calculated once in its lifecycle. If there is a midline, the count contribution will be calculated twice. Following the midline data collection, an interim count will be generated based on improvements between baseline and midline. Following endline data collection, a final count will be produced based on improvements in endline reading ability as compared to baseline reading ability. For the program contribution to the count, the endline count replaces the midline count. This document describes the process for estimating the endline contribution of a program to the Goal One target. The mechanics of estimating the midline contribution of a program are the same with the exception of using midline data rather than endline data.

If a program reaches indirect beneficiaries and assessment data have been collected for them, the methodology can be applied to indirect beneficiaries. If learning assessment data have not been collected for indirect beneficiaries, the methodology cannot be applied to this group.

## Methodology in Detail

### **STEP 1: Balance baseline and endline grade 2 data to have an equal number of observations**

The methodology is based on a comparison of the reading scores of two separate cohorts of students in the same grade but in subsequent school years. Because observations from the baseline and endline data must be paired in order to allow for a one-to-one comparison of reading scores, the first step is to balance the baseline and endline datasets so that they have the same number of observations.

It is crucial that the balanced data sets resemble the unbalanced data sets, i.e., the original baseline and endline data sets. In order to balance the data sets, the phase with fewer students is kept intact, and the same number of students in this phase is randomly selected from the phase with more students. Given that this process involves randomly selecting a sub-sample of students, the process is repeated through multiple iterations and the results from all iterations are averaged so that differences between the unbalanced and balanced data sets of the phase with fewer students are minimized.

### **STEP 2: Sort baseline and endline observations by descending order of reading scores and pair observations based on their rank order**

Once a balanced data set has been produced, the next step is to pair students in the baseline with those in the endline. This pairing is based on their ability level; e.g. the highest performing student from the baseline is paired with the highest performing student from the endline, the second-highest performing student from the baseline is paired with the second-highest performing student from the endline, and so forth. The use of the student score for the pairing is based on evidence showing that baseline scores are generally the best predictor of endline scores (e.g. high-performing students in the baseline tend to be high-performing students in the endline); that is, student scores tend to track.<sup>1</sup> The fact that multiple iterations of the

---

<sup>1</sup> Tracking is a term to describe a pattern in which individuals maintain their relative positions with respect to some attribute among a group of peers over time. The concept of tracking is closely tied to prediction. It indicates that a person's past contains useful information about the individual's present and future status. Tracking has been used primarily in the health field, e.g., to predict height and weight. It has also been used to describe changes over time in intelligence and other psychological attributes, as

balanced data set are produced implies that in different iterations students may be paired with different students, which provides an approximation of the performance of the pair to the expected performance of an actual student with a similar ability level.

### **STEP 3: Calculate the difference between baseline and endline reading scores for each ranked pair**

After the data from each phase have been balanced, sorted in descending order of scores, and paired by rank-order, calculating improvement in reading ability between the two cohorts is a simple arithmetic process. For each paired observation, the baseline reading score is subtracted from the endline reading score. For each pair, the difference between these values indicates the direction and magnitude of change in reading ability associated with the program.

### **STEP 4: Identify the proportion of ranked pairs representing meaningful improvement in reading**

USAID has set what it considers as a meaningful improvement in ORF for the purposes of the count. If improvement in reading ability for a given pair is at least as great as the level of change identified by USAID, E3/ED will consider the pair as representing an improved reader. The percentage of pairs achieving the meaningful improvement set by USAID is the number of pairs achieving the expected increment divided by the total number of pairs. For the purpose of the count, USAID established two increments that will count as meaningful, depending on the baseline ability level of the pair.

- For pairs whose baseline score demonstrates an inability to read a single word, a gain of the ability to read at least one correct word per minute is counted as improved reading.
- For pairs whose baseline score demonstrates an ability to read at least one word, a gain of the ability to read at least ten correct words per minute is counted as improved reading.

### **STEP 5: Estimate the number of unique student beneficiaries reached by the program**

Missions provide estimates of the number of unique student beneficiaries reached through each program to E3/ED on an annual basis. Figure 1 below is illustrative of the format of submitted beneficiary count data. Because the objective is to estimate the number of individual students reached by the program, students are counted only once – the first year they are reached by the program. In the case of a program that reaches grades 1-3 over the course of four years, the count of the unique beneficiaries reached by the program would include students in all three grades in the first year of the program, and then only the grade one students in subsequent years (shaded blue in Figure 1).

For the endline count estimation, the total number of unique beneficiaries reached over the life of the program should be used. For the midline count estimation, data should be used for only the unique beneficiaries reached over the span of years represented by the baseline and midline assessments.

In the example below, the number of unique student beneficiaries reached at endline is 15,400 (6,400 + 3,000 + 3,000 + 3,000). The number of unique student beneficiaries used for a midline conducted at the end of the 2012-2013 school year would be 9,400 (6,400 + 3,000).

---

well as in educational attainment, such as reading ability. There is also research around those individuals who do not track “properly” and their treatment as exceptions to the general tracking trend. For more information on tracking see Foster, T., Mohr, D., and Elston, R. (1989). A statistical model of tracking. *Communications in Statistics*, 18(8), 2861-2881; Foulkes, M., and Davis, C. (1981). An index of tracking for longitudinal data. *Biometrics*, 37, 439-446; Goldstein, H. (1981). Measuring the stability of individual growth patterns. *Annals of Human Biology*, 8, 549-557; McMahan, C. (1981). An index of tracking. *Biometrics*, 37, 447-455; Pekow, P. (1991). *A mixed model of tracking for inconsistently timed longitudinal data*. Chapel Hill, NC: Unpublished doctoral dissertation.

**Figure 1: Number of Beneficiaries Reached over the Life of the Program**

		Year 1	Year 2	Year 3	Year 4
		2011-2012	2012-2013	2013-2014	2014-2015
		Actual	Actual	Actual	Actual
Total Beneficiaries (Annual)	Fem.	3,700	3,700	3,700	3,700
	Male	3,700	3,700	3,700	3,700
	<b>Total</b>	<b>6,400</b>	<b>6,400</b>	<b>6,400</b>	<b>6,400</b>
Total Unique Beneficiaries (Cumulative)	Fem.	3,700	1,500	1,500	1,500
	Male	3,700	1,500	1,500	1,500
	<b>Total</b>	<b>6,400</b>	<b>3,000</b>	<b>3,000</b>	<b>3,000</b>
Grade 1	Fem.	1,500	1,500	1,500	1,500
	Male	1,500	1,500	1,500	1,500
	<b>Total</b>	<b>3,000</b>	<b>3,000</b>	<b>3,000</b>	<b>3,000</b>
Grade 2	Fem.	1,200	1,200	1,200	1,200
	Male	1,200	1,200	1,200	1,200
	<b>Total</b>	<b>2,400</b>	<b>2,400</b>	<b>2,400</b>	<b>2,400</b>
Grade 3	Fem.	1,000	1,000	1,000	1,000
	Male	1,000	1,000	1,000	1,000
	<b>Total</b>	<b>1,000</b>	<b>1,000</b>	<b>1,000</b>	<b>1,000</b>

**STEP 6: Estimate the number of unique student beneficiaries with improvements in reading skills**

To obtain the estimated number of beneficiaries with improved reading, the estimated unique number of student beneficiaries reached by the program (calculated in Step 5) is multiplied by the proportion of ranked pairs that reached the meaningful improvements in reading set by USAID (calculated in Step 4).

Note the steps allow us to make an estimate, not an exact count. For instance, the estimated unique number of beneficiaries reached by the program integrates information on learners who may never have passed through grade two at the time of an assessment (those in grade 3 in the first year of the program, and those in grade 1 in the final year of the program), whereas the calculated proportion of ranked pairs representing meaningful reading improvement is limited to grade 2 data. As discussed earlier, the results of this counting exercise are not meant to be a precise accounting of USAID Goal One programming but rather indicative of the number of students reaching a minimum expected improvement.

## Methodology Rationale

The methodology and approach to measurement outlined in this document are designed for the specific purpose of measuring progress towards the target of “improved reading skills for 100 million children in primary grades.” Because the methodology is designed to operate at the global level and to incorporate data collected from programs in dozens of countries, it contains design elements that, while appropriate at the global level, may not be immediately evident to practitioners and policymakers more accustomed to methodologies designed to work at the level of an individual program or country. This section outlines major characteristics of the design of the Goal One methodology and their effects on how results of the methodology should be used and interpreted.

### **The Education Strategy methodology is intended to be *Indicative* of the scale and effectiveness of USAID programming.**

USAID-supported Goal One programming reaches tens of millions of children in dozens of countries across the globe. Given the huge scale of this endeavor and great diversity of students and contexts reached through it, a methodology designed to generate a precise count of the number of students with improved reading ability simply is not warranted. The methodology is instead designed to generate a count that is indicative of the scale and effectiveness of Goal One programming. This is by no means an excuse for reduced methodological rigor. It is instead a practical approach to generating data sufficiently precise for their intended use, while maintaining a reasonable balance of investment in measurement relative to implementation.

The indicative nature of the methodology is evident in a number of areas:

- Strategy guidance directs Missions to estimate the number of individual students reached by counting them only the first time they participate in the intervention, rather than attempt to track individual students over the multi-year course of their educational careers;
- USAID’s strategic focus on reading as an outcome is reinforced through guidance indicating that the methodology will look exclusively at Oral Reading Fluency (ORF) because of its close relationship with reading comprehension;
- USAID’s strategic focus on developing reading ability early in a student’s educational career is reinforced through guidance directing Missions to measure learning outcomes at the end of grade 2, rather than attempt measurement of learning outcomes at each grade in every year of the program;
- Although Missions may choose to rigorously evaluate individual Goal One programs, USAID will not use the strategy as a way to evaluate Goal One implementation practices. The practices have already been proven in many contexts. At the global level, the challenge and cost of maintaining a rigorous treatment-control evaluation is not justified. USAID will not incorporate counterfactual information when calculating count contributions.

The selective measurement focus outlined here was purposefully chosen as a means for reinforcing the central priority of Goal One programming – teaching children to read with fluency and comprehension by the end of grade 2.

### **The methodology requires a consistent increment of improvement across all programs. This does not translate to a uniform measure of change across languages and contexts.**

This methodology applies a single increment of change as the measure of meaningful improvement in reading across all of the languages and contexts where USAID supports Goal One programming. USAID recognizes, however, that every program is characterized by unique circumstances that can affect reading gains by students. For example, differences in the orthographic complexity of the languages mean that

gains of the same increment may be much more difficult to achieve in some languages than in others. At present, there are not sufficient data for the languages USAID works in to support adjusting incremental gain expectations to factors such as orthographic complexity.

Despite the many factors that confound comparing reading gains across programs, USAID is confident that the gains targeted in this methodology are large enough to represent meaningful change in every context and language where we work. The target increment of 10 CWPM corresponds roughly to a 0.30 SD effect size for most of the programs for which data were available at the time this document was written. To put this in perspective, a 0.30 SD effect size equates to a gain of about half of a grade level in the U.S.<sup>2</sup> For cohort members who could not read a single word correctly at baseline (non-readers), an incremental gain of 1 CWPM represents the culmination of the pre-literacy skills required before a child can read their first word.

**The methodology is designed to be appropriate at the global- rather than country- or program-level.**

The increments of improvement outlined above represent the minimum improvement USAID is willing to count as improved reading for the purposes of the strategy. They are appropriate for the purposes of the Education Strategy, but should not be replicated as country or program level standards for several reasons.

First, the increments of change outlined here are the smallest increments USAID is willing to count as progress towards the Goal One target. They are in no way equivalent to the minimum expectation for program outcomes. Available data indicate that much larger gains – in excess of 30 CWPM -- are possible when programs are well-designed and implemented. Programs are expected to reach these much larger gains and the ultimate goal of children reading with fluency and comprehension.

Second, the standards developed for the Education Strategy were developed to be broadly applicable across the numerous contexts and languages where USAID supports Goal One programming. As a result, they are bereft of local context or expectations. When setting targets or expectations at the program or country level, Missions have the opportunity to engage in a more rich and inclusive process of developing contextually appropriate measures.

**Oral Reading Fluency will be used to measure progress towards the Goal One Target.**

Based on global best practices and research, USAID will use ORF to measure progress towards the Goal One target. For almost 40 years, ORF measures have consistently been shown as direct and reliable correlates with students' ability to recognize words and to comprehend text.<sup>3</sup> Put simply, this is because it is literally impossible for a child to understand a text that they cannot read fluently. Many correlation studies have demonstrated that ORF is a powerful predictor of reading comprehension ability when children are reading a language they speak and understand. Conclusive studies of the ability of ORF scores to predict comprehension outcomes have found concurrent validity coefficients of between .92 and .96 percent (a validity coefficient of 1.0, which almost never occurs, indicates a perfect correlation).<sup>4</sup>

While no child can understand text without being able to read it fluently, small minorities of children may demonstrate an ability to pronounce words without a corresponding ability to understand their meaning.<sup>5</sup>

---

<sup>2</sup> The U.S. Department of Education's Institute of Educational Sciences (IES) used results from several large standardized tests to estimate the expected spring-to-spring improvements in student scores from grades K-12 in different subjects. According to the U.S. data, students are expected to improve their reading scores on average by 0.60 standard deviation units from the end of grade 2 to the end of grade 3; thus, an effect size of 0.30 SD in reading speed would represent a half of a grade level improvement over the annual expected gain for grade 2.<sup>2</sup>

<sup>3</sup> See Deno, 1985; Fuchs, Fuchs, & Maxwell, 1988; Hosp & Fuchs, 2005; Jenkins, Fuchs, van den Broek, Espin, & Deno, 2003; Shinn, Good, Knutson, Tilly, & Collins, 1992; Wanzek et. al., 2010.

<sup>4</sup> See Children's Educational Services, 1987; Good, Kaminski, Smith, & Bratten, 2001.

<sup>5</sup> [http://www.pbs.org/launchingreaders/fluentreading/meettheexperts\\_2.html](http://www.pbs.org/launchingreaders/fluentreading/meettheexperts_2.html)

This could occur if students are forced to learn to read in languages that they do not speak and understand (a practice that contravenes best practice), or if students have insufficient language comprehension in a language they do understand. In USAID-supported EGRAs however, the number of students demonstrating satisfactory fluency performance and unsatisfactory comprehension performance has never been sufficiently large to greatly alter the validity coefficients between the ORF scores and comprehension scores for the target learner populations.

The use of ORF measures to calculate students' reading skill is based on the best available research, is reliable, enables governments to track progress towards benchmarks, is possible in every context where the Agency works, and enables aggregate calculations of progress towards the Goal One target. For these reasons, USAID measures progress towards the Goal One target by collecting ORF data measured in terms of CWPM.

USAID will use ORF to measure change in reading ability for the purposes of the USAID Education Strategy. This does not mean that reading comprehension or other valid measures related to the development of reading ability are invalid or inappropriate for use in other ways. Missions, implementing partners, and evaluating partners are encouraged to collect and make use of a range of appropriate measures to support the more detailed and nuanced analyses that are appropriate at the national and program levels.

**The Education Strategy is a tool for monitoring Agency-level outcomes, not a tool for program management or evaluation.**

The targets outlined in the 2011-2015 USAID Education Strategy were developed to add technical rigor to USAID's accountability for its programming through the addition of an objective outcome measure. As a management and accountability tool, the Education Strategy represents an addition to tools for management and accountability at the Mission level. The Education Strategy does not replace existing Mission-level and program-level management tools, but rather complements these tools by creating a lens for understanding program outcomes in the aggregate. The Education Strategy methodology should not be used to evaluate the effectiveness of an individual program.

## Data Required for the Methodology

Two types of data are needed in order to estimate a program's contribution to the Goal One target:

- The estimated total number of unique student beneficiaries reached over the life of the program. Missions provide E3/ED with annual updates to this figure.
- Baseline, midline, and endline learning assessment datasets and documentation. Missions ensure that implementers provide these to E3/ED within 90 days of data collection.

## Estimated Number of Unique Beneficiaries Reached Over the Life of a Program

As stated earlier, in order to estimate the number of students with improved reading ability it is necessary first to have an estimate of the number of individual student beneficiaries reached by the program. Because progress towards the Goal One target is reported cumulatively over a multi-year period, it is important to count individual student beneficiaries only once, even if they benefit from the program over multiple years. In other words, the goal is to count a student benefitting from a program only once, even if she participates as a first grader in the first year of the program, a second grader in the second year of the program, and so on.

The recommended approach to counting individual student beneficiaries of a program, which is outlined in the 2014 Update to Education Strategy Reporting Guidance and illustrated in Figure 1 of this document, is to count student beneficiaries only the first time they are reached by the program. This is generally achieved by counting all benefitting students in the first year of the program, and then only new entrants to the program in subsequent years. As discussed in the *Rationale* section of this document, numbers generated through this approach are acknowledged to be estimates because they do not attempt to incorporate the logistically daunting process of tracking individual student beneficiaries over the course of their educational careers.

The estimated count of student beneficiaries obtained through this process is referred to as a count of “unique beneficiaries” to emphasize the fact that it avoids double-counting student beneficiaries over time.

Missions provide E3/ED with annual updates to the estimated number of unique beneficiaries reached through each program. E3/ED typically leads a data call for this information every October.

### Data Needed for the Count:

#### 1. Estimated Number of Unique Beneficiaries Reached over the Life of the Program

- a. Disaggregated by sex, treatment group (in case more than one treatment group), and language (in case different population of students are tested in different languages)
- b. Data reported to E3/ED in October annually

#### 2. Baseline, Midline and End line Learning Assessments

- a. Representative of Grade 2 student beneficiaries
- b. Representative by sex
- c. Collected in the last two months of the school year
- d. Collected at the same time of year
- e. Baseline conducted the year before student beneficiaries are reached
- f. ORF strongly preferred
- g. Student-level observations for each subtask assessed
- h. Data transferred to E3/ED within 90 days of data collection
- i. Detailed guidance available on the EGRA Toolkit

## Learning Assessment Datasets and Documentation

USAID and the World Bank funded the development of the Early Grade Reading Assessment (EGRA) starting in 2005-2006 and ever since it has been the main source of reading assessment data used in USAID-funded evaluations. The methodology developed for the Goal One count is based on EGRA and, more specifically, the ORF subtask. Other reading assessments that capture ORF, such as the Annual Status of Education Report (ASER) assessment with the addition of timing to the component involving reading connected text, may also be used as sources for the Goal 1 count.

Assessments that measure ORF are the preferred instrument for measuring change in reading ability. However, the 2011-2015 USAID Education Strategy allows tests that do not measure ORF but have been shown to be reliable and appropriate measures of grade 2 literacy and pre-literacy skills.

This document outlines the general learning assessment parameters required to support the methodology. It is not intended to serve as a replacement for the detailed guidance on USAID-recommended practices that can be obtained in the updated EGRA Toolkit (expected 10/2015). Implementing and evaluating partners are required to use the guidance in the EGRA Toolkit when implementing a USAID-funded EGRA and contract/grant agreements should include language to that effect. Missions must monitor implementing and evaluating partners to ensure that their EGRA design and implementation include all parameters of the toolkit guidance, including sampling weights when applicable, inter-rater reliability (IRR), and statistically equated/comparable instruments.

At a minimum, grade 2 reading ability data from two points in time are required to perform the count. The baseline should be performed during the last two months of the school year preceding the first year the program reaches students in classrooms. The endline should be performed during the last two months of the final school year that program effects reach students in the classroom. A midline assessment is strongly recommended at the end of the second year the program reaches students in classrooms. Data must be representative of the grade 2 student population benefitting from program effects and representative by sex. If program treatment is segmented by language (i.e., different students receive treatment in different languages), data must also be representative of each language group.

In rare circumstances where, despite the best efforts of Mission and implementing or evaluating partner, it is not possible to collect a grade 2 baseline in the last two months of the school year preceding the first year that program effects reach students in classrooms, Missions may discuss with E3/ED the possibility of collecting a baseline from grade 3 children at the beginning of the first year of program implementation. This practice is strongly discouraged, but is seen as the least bad alternative when a true baseline cannot be collected.

The methodology anticipates making use of data that are cross-sectional rather than longitudinal. The purpose of the data is to allow a comparison of the reading ability of subsequent cohorts of grade 2 students against the reading ability of the baseline cohort of grade 2 students.

Since conducting one-on-one assessments such as EGRA has cost implications, the assessments are sample-based and often include multi-stage cluster sampling. Detailed guidance can be obtained in the EGRA Toolkit.

These are the minimum requirements of learning assessment data required to allow USAID to measure progress towards the Goal One target. Many Missions will want to add on to these minimum requirements in order to facilitate use of the data for other purposes beyond the Education Strategy. This may include establishing a control group or other counterfactual for use in an impact evaluation, collecting data for multiple treatment arms or data representative of additional grades or subpopulations of interest, or including a longitudinal element of data. All of these additions are fine as long as the minimum requirements for the Education Strategy are met.

E3/ED acknowledges that not all past programs will have gathered data in the manner required for the Education Strategy. To accommodate this, deviations from the methodology may be made in order to estimate counts for Goal One programs that did not collect data in accordance with guidance. However, it should be emphasized that **it is not acceptable to continue** generating data outside of the minimum requirements described here and in the EGRA Toolkit. E3/ED strongly advises that future programs pay particular attention to the requirements outlined in this document and the 2014 Update to Education Strategy Reporting Guidance.

## Data Management Responsibilities and Process

The information needed to estimate the Goal One count feeds into the six steps described below, with the responsible parties in parentheses:

**STEP 1:** Ensure that the EGRA design includes all parameters discussed in the EGRA Toolkit **(Mission, Implementer/Evaluator);**

**STEP 2:** Conduct a baseline and an endline on a sample that is representative of students who benefited from a USAID-funded intervention **(Implementer/Evaluator);**

**STEP 2:** Transmit the learning assessment data, codebook, and documentation to USAID via the SART Ed portal within 90 days of data collection **(Implementer/Evaluator);**

**STEP 3:** Transmit updates of student beneficiary count data to E3/ED through E3/ED's annual data call **(Mission);**

**STEP 4:** Use transmitted data to estimate the number of students showing improved reading skills between baseline and endline **(E3/ED);**

**STEP 5:** Aggregate the count contribution of individual programs and report at the Agency level on a regular basis **(E3/ED);** and

**STEP 6:** Make micro-data and supporting documentation available for public use via the SART Ed portal **(E3/ED).**

### Responsibilities of Missions, implementers/evaluators, and E3/ED are as follows:

#### USAID Missions

- Design implementation/evaluation mechanisms to ensure that student beneficiary data and learning assessment data are collected as required by the EGRA Toolkit and as needed for the Education Strategy;
- Ensure that contracts and agreements reference Education Strategy guidance and the EGRA Toolkit to make it evident that this is required implementation/evaluation guidance;
- Monitor implementer/evaluator adherence to the parameters of the Education Strategy guidance and the EGRA Toolkit;
- Communicate with host country governments to ensure learning assessment data can be made public in accordance with USAID policy;
- Provide annual updates to E3/ED on the number of unique student beneficiaries reached through E3/ED's annual data call;
- Ensure that implementers/evaluators transfer learning assessment data and documentation to E3/ED within 90 days of data collection; and
- Validate Goal One count results generated by E3/ED by checking figures used to generate the count.

#### Goal One Program Implementers/Evaluators

- Collect, store, and transmit learning assessment data in adherence to the parameters of the Education Strategy guidance and the EGRA Toolkit;
- Remove personal identifiable information (PII) from learning assessment datasets; and
- Transfer learning assessment data and documentation to E3/ED within 90 days of data collection.

### **USAID Office of Education (E3/ED)**

- Coordinate the collection of student beneficiary data from Missions and learning assessment data from implementers/evaluators;
- Coordinate E3/ED's application of the methodology to measure progress towards the target; and
- Manage the communication and dissemination of results.

To reiterate, Missions and implementers/evaluators are not responsible for calculating program contributions to the Goal One count.

## **Communicating Goal One Outcomes**

USAID is committed to principles of openness and transparency in Agency work towards the targets of the USAID Education Strategy. This extends well beyond transparency in the mechanics of estimating progress towards the targets themselves to include openly sharing descriptive information on the programs we support, detailed quantitative data on program level outcomes, learning assessment datasets, and documentation that underpin all of the above. We believe this open approach is the surest path to accelerated development of the field of knowledge surrounding early grade reading.

USAID anticipates sharing results and data related to Goal One of the USAID Education Strategy in the following ways:

- Publishing an annual update on progress towards the Goal One target of the USAID Education Strategy: improved reading for 100 million children in early grades;
- Publishing data-rich project briefs illustrating program-level reading improvements and progress towards the ultimate USAID objective that all children are able to read with fluency and comprehension by the end of their second year of primary school; and
- Disseminating learning assessment micro-data and documentation via the SART Ed website and other data platforms.

Finally, USAID recognizes a real danger in the potential consequence of placing undue emphasis on the mechanics of the Goal One count methodology. The Goal One methodology is an appropriate tool for the narrow purpose of measuring progress towards the target of the Education Strategy. As discussed elsewhere in this document, however, E3/ED is keenly aware that the methodology is not appropriate as a measure to be applied at the program or country level. USAID will report progress towards the Goal One target at the global level on an annual basis. When reporting on the midline and endline achievements of individual programs, USAID will emphasize progress towards the ultimate outcome of reading with fluency and comprehension and other measures of progress that do not place undue emphasis on any particular increment of change. This does not alter our commitment to sharing program micro-data and documentation whenever we are able.