POLICY BRIEF: Can Teacher-Administered Formative Reading Assessments Produce Relevant and Reliable Data?

Introduction. The availability and use of reading assessment data are critical to the goal of improving foundational literacy outcomes for students and realizing the potential of assessment-informed instruction. There is an overwhelming need for teachers to have timely data that they trust and understand, and for educational systems to have cost-efficient and relevant student reading assessments.

To advance the evidence on this topic, Room to Read conducted a study on the robustness of teacher-administered, classroom-based reading assessments within a broader study on assessment-informed instruction.[1] More specifically, the analyses discussed in this brief sought to determine whether data produced by teachers using Room to Read's **Student Tracking**[2] assessment are consistent and reliable indicators of student reading skills as measured by independently administered early grade reading assessments (EGRAs).

This brief highlights the key recommendations resulting from our study, in addition to providing an overview of its design, analytical methods, and key questions and findings.

KEY RECOMMENDATIONS

Governments, implementing partners, and researchers should analyze the reliability and predictive validity of current teacher-administered formative student reading assessments. If these teacher-administered reading assessments are found to provide strong and consistent signals of student performance**,** they can offer trusted, lower-cost, and more frequent measures of classroom-level outcomes than the typical larger, externally administered assessments.

These data can be used to:

➢ Determine school and classroom performance thresholds
➢ Categorize schools and classrooms into performance threshold levels
➢ Tailor support to schools, classrooms, and students

For formative assessments, care should be taken to:

➢ Use data to inform instruction and limit the use of data for accountability or other high-stakes purposes
➢ Provide robust training and coaching for teachers to administer assessments and utilize the data
➢ Streamline assessment scope and processes as much as possible
➢ Review the full distribution of student scores to optimize use of the data
➢ Incorporate ongoing use of data to refine teacher and student support systems

---

[1] C. Beggs, P. Joddar, *Supporting Assessment-Informed Instruction: New Evidence on the Implementation, Utilization, and Validity of Classroom-Based Assessments* (San Francisco: Room to Read, 2023),

[2] **Student Tracking** is a student reading skills assessment that is part of Room to Read's broader literacy program. The assessment is administered twice a year by teachers to all students in their classroom with support from Room to Read coaches. Teachers and coaches develop action plans in response to scores, and the model is contextualized across countries depending on the needs, capabilities, and preferences of education system actors.

Analytical methods. We used the following analyses in this study:

➢ **Descriptive analyses:** Descriptive analyses of grade 1 and grade 2 Student Tracking and grade 2 EGRA data sets.[3]

➢ **School performance category mapping:** Mapping of school performance categories, based on Student Tracking data, to EGRA scores.

➢ **Regression analyses:** Assessment of the associations between end-of-grade-2 EGRA scores and Student Tracking-based school performance levels.

➢ **Predicted probability analyses:** Estimation of the probability of children achieving benchmarks and zero scores on the oral reading fluency and reading comprehension tests in grade 2 EGRAs across school performance categories derived from Student Tracking data.

Data sampling. We utilized end-of-grade-2 EGRA samples collected during 2018 and 2019 as the sample frame for Student Tracking data across six countries: **Bangladesh, Cambodia, India, Laos, Nepal, and South Africa**. We used Student Tracking data from the same schools where EGRA data were collected[4] (outlined in Table 1).

Table 1. Final sample for EGRA and Student Tracking data

	Grade 1 Student Tracking		Grade 2 Student Tracking		Grade 2 EGRA	
	Oral reading	Reading comp.	Oral reading	Reading comp.	Oral reading	Reading comp.
School sample (all countries)	102	102	162	162	162	162
Student sample (all countries)	3,039	3,064	5,071	5,053	2,591	2,591

KEY QUESTIONS AND FINDINGS

QUESTION 1. Do teacher-administered assessments provide accurate information about student reading skills as validated by independent assessments?

FINDING. YES. Our analysis of data from six countries and several academic years shows **a consistently positive and statistically significant association** between school performance based on Student Tracking data from grade 1 and grade 2 and end-of-grade-2 EGRA scores. Additionally, based on predicted probability results, we found that students in well-performing schools for Student Tracking have **a greater chance of achieving higher reading fluency and comprehension scores** on EGRAs.

---

[3] These data sets were aligned to follow the same student cohorts, and we estimate that there is a very high level of overlap between the students included in the EGRA sample and the students included in the Student Tracking sample given that the schools are the same and the Student Tracking assessments were administered on a census basis in all of the schools included in the sample.

[4] We included round 2 Student Tracking data for grade 2 for all schools in the EGRA sample across countries except for South Africa. For grade 1, our school sample included four countries. In South Africa, only one round of grade 2 Student Tracking data was available that corresponds to the end-of-grade 2 EGRA school sample. For grade 1 analyses between Student Tracking and EGRA, we have excluded India and South Africa due to the unavailability of Student Tracking data.

Significant and positive associations between Student Tracking data and EGRA scores

The study looked at the associations between end-of-grade-2 EGRA scores and Student Tracking school performance for grade 1 and grade 2 separately through regression analyses. We found a consistent and clear pattern of statistically (highly) significant and positive associations between each of the four types of grade 1 and grade 2 Student Tracking school performance indicators and students' oral reading fluency (ORF) scores on the end-of-grade-2 EGRA in those schools.

For example, referencing grade 1 Student Tracking scores, keeping all other predictors constant, the grade 2 EGRA mean ORF in "doing well" schools is 15.61 cwpm higher (with a CI of 13.06 to 18.16 cwpm) than the grade 2 EGRA mean ORF in "struggling" schools. Also, for grade 2 Student Tracking scores, again holding all other predictors constant, the grade 2 EGRA mean ORF in "doing well" schools is 13.19 cwpm higher (with a CI of 10.63 to 15.76 cwpm) than the grade 2 EGRA mean ORF in "struggling" schools. As expected, we see a negative and statistically significant association between school performance based on Student Tracking data (grade 1 and grade 2) and the proportion of students scoring zero on the end-of-grade-2 EGRA.

Predicting benchmark performance

We found that if a school is performing well in Student Tracking, there is a higher chance that students in that school will meet benchmarks for oral reading fluency (=>45 cwpm) and reading comprehension (=>80% correct questions) on the grade 2 EGRA. The opposite is true for students in schools where Student Tracking scores are lower. The results using grade 1 Student Tracking data are outlined below. Results for using grade 2 Student Tracking data displayed similar patterns. See Table 2 for full results.

> **Grade 1, oral reading fluency.** For a student in a school categorized as "doing well" by grade 1 Student Tracking data, there is a 42% chance that the student will achieve the oral reading benchmark in the grade 2 EGRA. A student in a school categorized as "struggling" by grade 1 Student Tracking data is 12 percentage points less likely than students in "doing well" schools to achieve the same benchmark in the grade 2 EGRA.

> **Grade 1, reading comprehension**. For a student in a school categorized as "doing well" by grade 1 Student Tracking data, there is a 50% chance that the student will cross the =>80% correct questions benchmark in the grade 2 EGRA. A student in a school categorized as "struggling" by grade 1 Student Tracking data is 15 percentage points less likely to achieve that benchmark.

Table 2. Summary of predicted probability results

Predicted probability of children achieving oral reading fluency and reading comprehension benchmarks in grade 2 EGRA across categories of schools as determined by Student Tracking				
School categories	**Oral reading (source: grade 2 EGRA)**		**Reading comp. (source: grade 2 EGRA)**	
	ORF >= 45+ cwpm	**Zero score**	**Comp. >=80%**	**Zero score**
Categories based on grade 1 Student Tracking				
Doing well	42%	2%	50%	6%
Needs improvement	38%	7%	41%	12%
Struggling	30%	15%	35%	21%
Categories based on grade 2 Student Tracking				
Doing well	48%	8%	44%	11%
Needs improvement	35%	11%	39%	21%
Struggling	26%	16%	28%	31%

QUESTION 2. Do the school performance categories generated by teacher-administered Student Tracking assessments provide meaningful differentiation between school performance levels to enable tailored support to schools and teachers?

FINDING. **YES, especially for higher-performing schools.** Findings indicate that **school performance categories** based on Student Tracking data are **useful and generally well aligned with EGRA scores**, even with context-specific assessment adaptations and more informal classroom-based administration.

As a partner working with governments to improve early reading outcomes, Room to Read must take special care to ensure that our monitoring systems produce high-quality and actionable data that can inform the levels of support we need to provide to schools and teachers. To this end, a key part of our Student Tracking model is the categorization of schools into performance categories, based on Student Tracking ORF and reading comprehension scores (see categories in Table 3). Having meaningful performance categorization is necessary to tailor support to schools and classrooms.

Table 3. Categories of school performance based on Student Tracking scores

School performance category	Oral reading fluency score	Reading comprehension grade 1 score	Reading comprehension grade 2 score
Doing well	75-100% of students reading 75% or more words correctly	75–100% of students answered at least one (out of three) question correctly	75–100% of students answered at least two (out of four) questions correctly
Needs improvement	50–74% of students reading 75% or more words correctly	50–74% of students answered at least one (out of three) question correctly	50–74% of students answered at least two (out of four) questions correctly
Struggling	Less than 50% of students reading 75% or more words correctly	Less than 50% of students answered one (out of three) question correctly	Less than 50% of students answered two (out of four) questions correctly

We explored whether the school performance categorization based on Student Tracking scores provides meaningful differentiation (based on EGRA scores) between school performance levels. The data indicate this to be broadly true. This is especially the case for higher-performing schools, as across data sets, the "doing well" schools have the highest EGRA mean ORF scores (see "doing well" category for Bangladesh, India, Nepal, and South Africa examples in Table 4), with only two exceptions across the data sets we analyzed.

Table 4. EGRA mean ORF scores by Student Tracking performance category (excerpted results)

	School performance category	School sample	EGRA mean ORF score (end of grade 2)
Bangladesh (Dhaka)	Doing well	117	60.9
	Needs improvement	81	47.4
	Struggling	36	48.0
India (Madhya Pradesh)	Doing well	14	31.1
	Needs improvement	44	25.0
	Struggling	37	24.8
Nepal (Nuwakot)	Doing well	12	30.6
	Needs improvement	42	16.0
	Struggling	25	23.2
South Africa (Limpopo)	Doing well	140	34.7
	Needs improvement	163	28.3
	Struggling	92	19.3

However, the alignment between school performance categories and eventual EGRA scores gets less clear and more varied as we look at the "needs improvement" and "struggling" performance categories, with some mean scores higher in the "struggling" category than the "needs improvement" category (see gray shaded cells in Table 4).

There are also patterns specific to context that should be taken into consideration when setting thresholds for performance categories. For instance, in Natore district in Bangladesh, we found higher EGRA ORF cwpm scores at the bottom of the distribution as compared to other geographies. In Natore district for grade 2 students, the lowest EGRA ORF score was 8 cwpm in the "struggling" schools, 18.8 cwpm in the "needs improvement" schools, and 31.1 cwpm in the "doing well" schools. Every other geography included in this analysis has zero scores as the minimum EGRA ORF score in each school performance category. These variations in score distributions need to be taken into consideration to optimize school performance categorization.

In sum, this study found that Room to Read's Student Tracking model and process is providing reliable data on student reading outcomes. We also found that some of our categorizations, especially for lower-performing schools, need to be refined to ensure that we can tailor our support to teachers. There is a more generalizable finding that it is possible for teacher-administered assessments to provide accurate data on student reading outcomes and that those data can be used to identify high-performing and struggling students and classrooms. This opens up possibilities for governments to utilize these types of data to check on the health of their education system and, most importantly, tailor their support to teachers.

See this study's Policy Brief and the full study Final Report for a broader discussion on recommendations for implementing effective assessment-informed instruction. Policy Brief: https://scienceofteaching.site/research/ Final Report: https://scienceofteaching.site/wp-content/uploads/2023/07/Aii-Study-Room-to-Read-Final-June-29-2023.pdf

Suggested citation for full study: Beggs, C., Joddar, P. (2023) "Assessment-*Informed Instruction: New Evidence on the Implementation, Utilization and Validity of Classroom-Based Assessments.* San Francisco: Room to Read.