

Evaluation Matrix Template Worked Example

October 2026

A made-up program. All names, places, organizations and figures are invented.

Evaluation Questions

- Question 1 (Relevance): To what extent does the design of the Sunnowen Early-Grade Reading Program match the reading needs of students, teachers and caregivers in its three districts (Pike Hollow, Verrin Ford and Tessaly), and does it stay relevant as needs change?
- Question 2 (Coherence): How well does the program fit with the national early-grade curriculum, the district teacher-support system and other reading work in the schools?
- Question 3 (Effectiveness): Do students in program schools improve their reading, do the gains differ for girls and boys, students with disabilities, students with a different home language and students in remote schools, and do teachers use the structured lesson method as designed?
- Question 4 (Efficiency): Are resources used well and on time, and what does the program cost for each student reached?
- Question 5 (Impact): What longer-term changes can be linked to the program, including attendance, reading at home and use of the method outside the program grades, and are there unintended effects?
- Question 6 (Sustainability): Will the teaching practices, reader supply and coaching continue as program funding comes to an end, and what would keep them going?

Coverage Check

Criterion	Questions	Question Numbers	Check
Relevance	1	1	Covered
Coherence	1	2	Covered
Effectiveness	1	3	Covered
Efficiency	1	4	Covered
Impact	1	5	Covered
Sustainability	1	6	Covered
Other	0	None	Not used

Evaluation Matrix

Question and Criterion	Sub-Questions	Judgment Criteria	Indicators	Data Sources and Methods	Analysis
Question 1, Relevance	Did the lessons and readers address the gaps found at baseline? Do teachers and caregivers find them useful and in the right language? Did the changes after midline answer what midline found?	Design addresses the gaps found at baseline; teachers and caregivers rate lessons and readers as useful and in the right language; changes after midline answer what midline found.	Share of teachers rating lessons and readers useful or very useful; share of caregivers who say books match their child's level; number of design changes after midline, with reasons.	Program documents; baseline and midline data; teacher survey; caregiver survey; key informant interviews. Methods: document review, surveys, interviews	Compare survey ratings with the gaps found at baseline; list design changes against midline findings; code interview themes.
Question 2, Coherence	Do program lesson plans follow the skills and sequence of the curriculum? Do program and district coaches overlap or support each other? Does any other reading effort work against the program?	Lesson plans cover the same skills and sequence as the curriculum; program and district coaches do not duplicate visits; no other reading effort works against the program.	Share of curriculum learning outcomes covered by program lesson plans; number of schools visited by both program and district coaches in the same term; overlapping partner programs found.	Curriculum and district plans; program documents; coaching and training records; key informant interviews. Methods: curriculum mapping, document review, interviews	Map lesson plans against curriculum outcomes; count schools with overlapping visits; code interview themes.
Question 3, Effectiveness	What share of grade 2 students meet the reading target? Do program schools gain more than comparison schools, and which groups gain least? Do teachers use the lesson method as designed, and do more coaching visits lead to fuller use?	Met: 40% or more of grade 2 students read 45 or more correct words a minute with 80% comprehension (baseline 11%); partly met: 30 to 39%. Program schools gain more than comparison schools. Gaps between groups do not widen. At least 70% of observed lessons show all five core steps.	Share of grade 2 students meeting the reading target; share reading no words (baseline 38%); difference in gain between program and comparison schools; gain by sex, disability, home language and travel time to town; share of observed lessons with all five steps; teachers trained and coached as planned (410 trained, 6 visits a year).	Endline reading test; baseline and midline data; classroom observation; teacher survey; coaching and training records; focus groups; key informant interviews. Methods: timed oral reading test with comparison schools, secondary data analysis, lesson observation, survey, records review, focus groups, interviews	Compare the share meeting the target and the gain in program and comparison schools; split gains by group and name any group gaining less than half the average; compare lesson observation results by number of coaching visits.

Question and Criterion	Sub-Questions	Judgment Criteria	Indicators	Data Sources and Methods	Analysis
Question 4, Efficiency	Did spending follow the plan for each component? Did readers and coaching arrive before the terms they were meant for? What did the program cost for each student reached?	Spending follows plan within 10% by component; readers and coaching arrive before the terms they are meant for; cost per student is within 20% of the cost per student of two comparable early-grade reading programs.	Spending against budget by component; days of delay for reader delivery and coaching visits; cost per student reached and per student meeting the target.	Financial records; coaching and training records; program documents; cost benchmarks; key informant interviews. Methods: financial analysis, delivery-note review, benchmark comparison, interviews	Compare spending with budget by component; compute cost per student reached and per student meeting the target; compare with the benchmark.
Question 5, Impact	Have attendance and reading at home changed? Does the link to the program hold after other explanations are checked? What unplanned effects, good or bad, appeared?	Attendance and reading at home move as the theory of change expects, and the link to the program holds after other explanations are checked; any harm or unplanned effect is named.	Attendance rate in grades 1 to 3, 2023 to 2026; share of caregivers reading with a child 3 or more days a week; share of grade 1 and 3 teachers using the method without program training; unplanned effects named.	School registers; caregiver survey; teacher survey; focus groups; key informant interviews; endline reading test. Methods: register review, reading test, survey, focus groups, interviews	Contribution analysis: test each step of the theory of change against the evidence and rule out other explanations; compare attendance trends.
Question 6, Sustainability	Will districts fund coaching after the program ends? Is there a funded plan to reprint and deliver readers? Who takes ownership at school and district level?	Lesson steps still in use; districts fund or run coaching; a funded plan exists to reprint and deliver readers; principals and district officers take ownership.	District budget lines for coaching and readers in 2027; share of schools visited by a coach since January 2027; share of schools with working reader stocks; observed use of the method.	Curriculum and district plans; classroom observation; key informant interviews; coaching and training records; focus groups. Methods: budget review, lesson observation, interviews, records review, focus groups	Compare district budget lines with the cost of coaching and readers; code interview themes on ownership; rate each sustainability factor as in place, partly in place or missing.

Triangulation Check

Question	Sources (Count)	Methods (Count)	Triangulation Check
Question 1	5	3	Met
Question 2	4	3	Met
Question 3	7	7	Met
Question 4	5	4	Met
Question 5	6	5	Met
Question 6	5	5	Met

Data Sources

Data Source	Method and Tool	Sample or Coverage	Lead and Dates	Questions
Program Documents	Document review; review checklist	Results framework, annual reports, design papers and program theory	Evaluation lead, February 2027	1, 2, 4
Baseline and Midline Data	Secondary analysis of reading test datasets	March 2023 baseline and March 2025 midline, program and comparison schools	Data analyst, February to March 2027	1, 3
Endline Reading Test	Timed oral reading test; student assessment tool with child functioning and home language questions	1,440 grade 2 students, 24 per school in 30 program schools (drawn at random from the 64) and 30 comparison schools	Field coordinator, 8 March to 16 April 2027	3, 5
Teacher Survey	Tablet survey; teacher questionnaire	120 teachers, 4 per school in the 30 program schools: 3 trained grade 2 teachers and 1 grade 1 or 3 teacher without program training	Field coordinator, 8 March to 16 April 2027	1, 3, 5
Classroom Observation	Structured lesson observation; observation checklist	90 lessons, 3 per school in the 30 program schools	Education specialist, 8 March to 16 April 2027	3, 6
Coaching and Training Records	Records review; extraction sheet	Training attendance sheets and visit logs for all 64 program schools, 2023 to February 2027	Data analyst, February to March 2027	2, 3, 4, 6
Financial Records	Financial analysis; budget variance sheet	Budgets, spending reports and supplier delivery notes, 2023 to 2026	Evaluation lead, February to March 2027	4
Cost Benchmarks	Document review; cost comparison sheet	Published cost reports of two comparable early-grade reading programs	Evaluation lead, February 2027	4
School Registers	Register review; attendance extraction sheet	Enrollment and attendance, grades 1 to 3, the 60 tested schools, 2023 to 2026	Data analyst, March to April 2027	5
Key Informant Interviews	Semi-structured interviews; interview guides	24 people: 6 district officers, 3 provincial officers, 9 principals, 2 curriculum staff, 2 suppliers, 2 program staff	Evaluation lead, 8 March to 16 April 2027	1, 2, 3, 4, 5, 6
Focus Groups	Focus group discussions; discussion guides	12 groups: 8 with parents and caregivers, 4 with teachers; 6 to 8 people each	Education specialist, 15 March to 16 April 2027	3, 5, 6
Curriculum and District Plans	Curriculum mapping; document review	National early-grade curriculum; 3 district education plans (Pike Hollow, Verrin Ford, Tessaly) and 2027 budgets	Education specialist, February to March 2027	2, 6
Caregiver Survey	Phone and household survey; caregiver questionnaire	480 caregivers of tested students, 16 per program school	Field coordinator, 15 March to 16 April 2027	1, 5