CORE GUIDE / 11

Baseline, Midline & Endline

Measure change over time with comparable tools, clean data and careful interpretation.

START WITH A DECISION, NOT A FORM. Ask what decision the evidence should inform before deciding what to measure.

The essential idea

A **baseline** records starting conditions before, or at the very start of, implementation. A **midline** checks progress and implementation part-way through, while there is still time to adjust. An **endline** measures where things stand at the end, so the team can see **how much changed** and for whom.

WHY IT MATTERS

Without a baseline you can only show where people ended up, not how far they moved. Without a midline you learn about problems when it is too late to fix them.

Plan the three rounds as one design

Change can only be measured if every round is collected the same way. Decide these before the baseline, then hold them steady.

- Use the **same indicators, tools and scoring rules** in every round. Pilot first, then freeze the tool.
- Assess the **same people, or a comparable sample**, and give each person a unique ID so records can be matched.
- Time the rounds to the programme calendar, for example the start, middle and end of the school year.
- Train assessors the same way each time and record who collected what.
- Plan for drop-outs: note who is missing at each round and why.

What each round is for

Round	Main purpose	Typical questions
Baseline	Set starting values, realistic targets and priorities for support.	Where are participants starting from? Who needs the most help?
Midline	Check fidelity and early outcomes; correct course.	Is delivery happening as planned? Are early changes visible? What should we adjust?
Endline	Measure change against the baseline and the targets.	How much changed, for whom, and how confident can we be about why?

Change is not the same as impact

A difference between baseline and endline shows change over time. It does not, by itself, prove the programme caused it.

- Use a **comparison group** where you can, or build a clear contribution argument where you cannot.
- Account for **natural growth**: children learn over a school year with or without extra support.
- Watch for **seasonality** and outside events between rounds.
- Report **attrition**: if the people who left differ from those who stayed, averages can mislead.

• Compare **matched individuals**, not only class or group averages.

Worked field example

A school runs remedial classes for foundational reading and arithmetic across one academic year.

Programme question	Evidence to use	Decision it can support
Which students are below grade level in reading and arithmetic?	Baseline ASER-style levels by grade and by student, before remedial classes begin.	Group students by actual level and plan targeted sessions.
Is the remedial support working mid-year?	Midline levels for the same students, plus attendance and session records.	Regroup students, add sessions or support facilitators where progress is slow.
How much did each student grow by year end?	Endline levels matched to baseline using student IDs; attrition noted.	Report growth honestly and decide what to scale, change or stop.

Try it in your work

• List three indicators and confirm each will be measured in exactly the same way in all three rounds.

• Put the three rounds on your programme calendar now, with an owner for each.

• Decide how you will match the same participants across rounds.

Common mistakes

• Changing the tool or scoring between rounds.

• Collecting the baseline after activities have started.

• Comparing different groups of people and calling it change.

• Ignoring who dropped out.

A stronger habit

• Pilot, then freeze the tool for all rounds.

• Collect the baseline before the first session.

• Use unique IDs and analyse matched records.

• Record attrition and the reasons for it.

REFLECT

If your endline showed no change, what would you need from the baseline and midline to understand why?