

Measuring whether equity work is working, without scoring people

Tell the difference between a platform that works and a program that improves equity. Use a logic model, baselines, mixed methods, disaggregation with small-group protection, and honest limits. Never use individual learning, participation, or assessment results as performance data.

Why it matters

Invented targets, composite scores without validation, and individualized learning surveillance destroy trust and prove nothing. Good measurement is possible and protects people.

Content

Two different questions. Does the platform work: can people find, use, and apply material? Does the program improve equity practice or outcomes: are decisions, services, and workplace patterns changing? Measure both, separately.

Logic model. Inputs and activities, outputs, outcomes, and impact. Write what you expect to change and by when before you measure.

Baselines first. No causal claim without a baseline. If there is no baseline, say so and start one.

Mixed methods. Numbers show patterns; people's accounts show why. Use both, and give qualitative evidence equal standing.

Disaggregate with protection. By community, language, disability, and place where the data allows. Set the smallest group you will report, and suppress below it. Ask the data owner about denominators.

Protect private learning. Learning, reflection, Ask, and practice paths are private. Report only combined totals that cannot identify anyone, and never use them for supervision or evaluation.

Never use assessment results as performance data. Intercultural assessments and similar instruments describe development; they are not scores of competence and must not enter personnel decisions.

State the limits. Every report says what the evidence does not show. That sentence is what makes the rest trustworthy.

Next steps

- **Measurement and accountability area** — program page /areas/measurement
- **Six commitments for working in partnership** — program page /library/pn-partnership-spine