Learning-Gain Measurement Protocol

Evidencing teaching effectiveness by how much students learn, not only what they score.

AcademicOS · Production Layer · Assessment Bench

Principle. Measure the same construct at entry and at exit, on the same scale, and report the movement. A learning-gain metric credits the distance a learner travels, which a raw exit mark cannot, because it adjusts for who walked in the door.

The seven steps

1. **Define the construct.** Fix the outcomes the course promises; the pre- and post-test assess these, nothing more.
2. **Build matched instruments.** Two parallel forms of equal difficulty, blueprinted by Bloom level, with a few **transfer items** on the post-test.
3. **Take the baseline.** Pre-test in week one, kept **low-stakes and ungraded**.
4. **Add a mid-point check.** A short formative measure turns two bookends into a trajectory.
5. **Take the exit measure.** Same conditions; record **who is missing**.
6. **Compute the gain** (use more than one) (see formulae opposite).
7. **Close the loop.** Review with teacher and mentor, agree one change, record it.

Decide before you start

State plainly whether the metric is **developmental** (seen by teacher and mentor, used to improve) or **evaluative** (used in appraisal). The moment it is high-stakes, the incentive to corrupt the baseline appears.

The formulae

Normalised gain

`g = (post − pre) / (max − pre)`

Bands: ≥0.7 high · 0.3 to 0.7 medium · <0.3 low

Effect size

`d = (mean_post − mean_pre) / SD_pooled`

Hedges correction for small classes. 0.2 small · 0.5 medium · 0.8 large

Value-added (programme level)

`residual = actual_exit − predicted_exit`

predicted exit modelled from prior attainment

Worked check (one minute)

Two sections of one course: A moves 42→71 (raw +29), B moves 55→78 (raw +23). Raw gain says A won. Normalised says otherwise: A g = 29/58 = **0.50**, B g = 23/45 = **0.51**, near-identical medium gains. The raw ranking was an artefact of the intake. This is why you normalise.

The six guards

- **Same construct, not the same paper.** Parallel forms or equating, so gain is not test familiarity.
- **Keep transfer items.** Improvement here, not just the rehearsed format, signals understanding.
- **Report attrition.** Quote the gain with who sat both tests, never survivors alone.
- **Check reliability.** Confirm the instrument is stable (e.g. Cronbach's alpha).
- **Aggregate small classes.** One course's gain is a signal, not a sentence.
- **Decouple from grades and reward.** So lowering standards can never raise the number.