

Types of Evaluation Metrics in Postgraduate Training

Why This Matters

Different metrics answer **different questions**:

- *Where is the learner developmentally?*
- *Did they meet expectations?*
- *Are they progressing faster than expected?*
- *What level of independence is safe?*

No single scale does all this well.

Evaluation Type	Example Scale	Description	What it answers	Best use	Strengths	Limitations
Developmental Metrics	<i>Novice → Expert Models</i>	A developmental framework that places learners along a progression of growth from early beginner to advanced practitioner.	Where is this learner in their professional development right now?	Longitudinal tracking, milestones, coaching conversations, and CCC reviews.	Growth-focused, normalizes early struggle, and aligns with adult learning.	Less precise for pass/fail decisions and requires a shared understanding of levels.
Achievement Metrics	<i>Met / Not Met / Partially Met</i>	A goal-based metric that evaluates whether stated objectives or	Were the stated objectives achieved?	Didactics, projects, QI initiatives, and program or site evaluations.	Simple, clear, objective-driven, and accreditation-friendly.	Does not show degree of excellence and has limited

		requirements were achieved.				developmental nuance.
Performance Quality Metrics	<i>Exceeds / Meets / Below Expectations</i>	A comparative quality scale that judges how well performance matches expectations for the learner's stage.	How well was this performed compared to expectations for stage?	Later-stage learners, staff or preceptor evaluations, and summative reviews.	Familiar to stakeholders and allows recognition of excellence.	Risk of rating inflation and can discourage learners early in training.
Excellence-Oriented Metrics	<i>Excellent / Very Good / Satisfactory</i>	A quality-oriented scale used to judge overall strength of performance relative to standards or peers.	How strong is performance relative to peers or standards?	Faculty or staff evaluations, program quality reviews, and organizational assessments.	Useful for quality benchmarking and less threatening for non-learner evaluation.	Not developmentally grounded and a poor fit for early learners.
Entrustment / Autonomy Metrics	<i>Supervision-Based Scales</i>	A supervision-based scale that indicates how much oversight a learner needs to perform safely in clinical practice.	How much supervision does this learner require to practice safely?	Clinical performance, patient safety decisions, and readiness for independent practice.	Clinically meaningful, directly tied to patient safety, and ideal for postgraduate training.	Requires faculty calibration and is less useful for didactics.

Milestone-Anchored Metrics	<i>Stage-Expected Performance</i>	A stage-based metric that evaluates whether a learner is below, at, or above the expected level for their point in training.	Is this learner progressing as expected for their point in training?	Quarterly reviews, CCC decisions, and progression determinations.	Balances development with accountability and normalizes variability.	Requires clearly defined milestones.
Rubric-Based Metrics	<i>Criteria-Referenced Scoring</i>	A structured scoring approach that evaluates performance against specific predefined criteria.	How well were specific criteria met?	Projects, QI initiatives, scholarly work, and leadership activities.	Transparent, objective, and highly defensible.	Time-intensive to build and needs regular review.
Narrative & Qualitative Metrics	<i>Descriptive, Not Numeric</i>	A descriptive feedback approach that captures context, nuance, and specific observations rather than reducing performance to a numeric score.	What specifically is happening, and why?	Formative feedback, professional identity development, and explaining numeric ratings.	Provides rich context and is often the most meaningful to learners.	Harder to aggregate and requires faculty skill.

How to Choose the RIGHT Metric

Ask These 3 Questions:

- 1. Is this developmental or summative?**
- 2. Is patient safety involved?**
- 3. Do I need to show growth, achievement, or readiness?**

Critical Rule

Never rely on a single metric type for high-stakes decisions.

Use:

- Developmental + Entrustment + Narrative or**
- Milestones + Rubrics + Faculty consensus**