

Handout: Assessment of Writing

Assessment Method	Benefits	Constraints
Portfolios	Provide opportunities for assessment conversations to consider readers' interpretations	Judgments of writing performance too often based on highest quality pieces rather than entire body of work
	Focus on authentic writing tasks rather than writing in response to assessment-driven prompts	Samples written and selected under diverse conditions and levels of support, rendering judgments difficult to interpret
	Accommodate evaluation of writing knowledge, skills, and processes as well as writing outcomes	Wide array of portfolio models can obfuscate alignment between portfolio design, data collection and analysis, and assessment purposes
	Can represent a balanced view of communicative competence because oral and visual performances can be included	Tensions between assessment and learning functions of portfolios can negatively impact reliability and validity of evidence
Indirect	Measurement reliability is enhanced through ease of scoring and high degree of consistency across test items	Lack content and face validity because tasks do not evaluate application of writing knowledge, skills, or processes
	Permit large-scale assessment in a single session	Do not reflect real-world writing demands
On-Demand	Assessment conditions and writing measures can be standardized	Measurement error is introduced through poorly written prompts; presentation effects (e.g., legibility, spelling mistakes) and writer identity lead to scoring bias
	Holistic, analytic trait, and primary trait rubrics are robust for evaluating multiple aspects of written composition in multiple forms for multiple purposes	Rubrics tend to not be sensitive to small increments of change and are difficult to use reliably
Curriculum-based Measurement	Brief, repeated sampling of timed on-demand writing (for young writers, copying of text is acceptable) performance is standardized, as are scoring procedures	Brevity of writing sample influences validity of inferences regarding writing performance—older students need longer sampling periods
	Several writing measures can be applied to a given sample, including total words written, words spelled correctly, correct letter sequences, and correct word sequences	No single measure is sufficiently reliable and valid for all students at all grades; more complex measures appear to be better predictors of writing competence
	Progress monitoring of individual students for evaluating responsiveness to intervention is permitted	Though long-term gains are found, research has yet to identify how frequently samples should be collected, the rates of growth that can be expected, or how these measures are impacted by instruction