



KDE Numeracy Diagnostic Assessment Decision Rubric

Review Date: Click or tap to enter a date.

Test Name: Click or tap here to enter text.

Submission form contains all requested information: Yes: ☐ No: ☐

Overall Rating: Choose an item. /8

Approved: ☐ Rejected: ☐ Date: Click or tap to enter a date.

Decision Rationale:

Resubmission window (if applicable) for rejected assessments: Click or tap to enter a date.

**If not approved, submitters may provide additional evidence or clarification within two weeks of the rejection notification. This additional information will be reviewed by the team, and a final decision will be made at that time. No further resubmissions will be allowed.*

Domains/Construct Assessed: Check all that apply.

Counting and Cardinality (K)	Number and Operations Base 10	Operations and Algebraic Thinking	Number and Operations Fractions (3-5)	Measurement and Data	Geometry	Procedural Fluency	Conceptual Understanding

Technical Adequacy Scoring Rubric

Part 1: Classification Accuracy: AUC values closer to 1 indicate the screening measure reliably distinguishes among students with satisfactory and unsatisfactory numeracy performance, whereas values at 0.50 indicate the predictor is no better than chance.

Score	Definition
	Evidence provided for each testing window, domain, subdomain and grade-level targeted by the tool:

Score	Definition
2	AUC $\geq$.70 OR Sensitivity and Specificity $\geq$.70
1	Evidence is mixed with some testing windows, domains, subdomains or grade-levels targeted by the tool not meeting the criteria of $\geq$.70 or not reported. OR Data provided demonstrates weak evidence in some but not all testing windows, domains, subdomains or grade-levels .
0	Evidence was not provided for these criteria, or information does not demonstrate an acceptable level of evidence .

Reviewer notes to support rating for classification accuracy:

Part 2: Reliability: the consistency of a set of scores that are designed to measure the same thing.
Internal consistency (usually reported as Cronbach's alpha) or alternate form where applicable.

Score	Definition
2	Evidence is provided for each domain, subdomain and grade-level targeted by the tool : Reliability measure is $\geq$.70
1	Reliability evidence is mixed with some domains, subdomains or grade-levels targeted by the tool not meeting the criteria of $\geq$.70 or not reported for a domain or grade-level.
0	Evidence was not provided for these criteria, or information does not demonstrate an acceptable level of evidence .

Reviewer notes to support rating for reliability:

Part 3: Validity: how accurately the assessment predicts future academic performance and/or how well the domain/subdomain represents the construct it measures.

Score	Definition
2	Evidence is provided for each domain, subdomain and grade-level targeted by the tool: Predictive Validity and/or content validity measure is $\geq .60$.
1	Evidence was mixed, some domains, subdomains or grade-levels targeted by the tool not meeting the criteria of $\geq .60$ or not reported for a domain or grade-level.
0	Evidence was not provided for these criteria, or information does not demonstrate an acceptable level of evidence.

Review notes to support rating for validity:

Part 4: Representative Sample: The sample used for psychometric analysis closely matches the characteristics of its population as a whole.

Score	Definition
2	The research sample used closely matches the population of the United States.
0	Insufficient or no evidence was provided on the characteristics of the sample used.

Reviewer notes to support rating for representative sample: