High Quality Assessment Content Validity Review Tool

How to use the Assessment Review Tool

Content Area: Mathematics		
Name of Assessment: Mathematics Assessment Program, Middle School Mathematics, MS-2 Reviewer: Content Collaborative		
Assessment Profile		
Item Types - check all that apply (note: there is often overlap among certain item types): Selected Response (multiple choice, true-false, matching, etc.) Short Answer (short constructed response, fill in a graphic organizer or diagram, explain	Check All That Apply	
your thinking or solution, make and complete a table, etc.) Extended Response (essay, multi-step response with explanation and rationale required for tasks)	x x	
Product (research paper, editorial, log, journal, play, poem, model, multimedia, art products, script, musical score, portfolio pieces, etc.)		
Performance (demonstration, presentation, science lab, dance or music performance, athletic performance, debate, etc.) Process (creation, development, design, exploration, imagining, visualization, experimentation, invention, revision)		
The assessment includes:	Check All That Apply	
Teacher directions (may include prerequisites/description of instruction before giving the assessment e.g., this assessment should be given after students have learned)		
Scoring Guide/Rubric Sample evidence to show what student performance might look like Materials (if needed to complete the assessment) Estimated time for administration	x	
Student Directions & Assessment Task/Prompt – what does the student see/use? Other:	x	
Other:		

A high quality assessment should be...Aligned

To understand the review process and the use of the review tool, go to:

Alignment	Rating Column	Comments
1a.		
Grade Level(s): 6-8		
Indicate the Colorado Academic Standards and Grade Level Expectations evaluated by		
the Assessment: MA10-GR.6-S.1-GLE.1-EO.cvi; MA10-GR.6-S.3-GLE.1; MA10-GR.7-S.3-		
GLE.1; MA10-GR.8-S.2-GLE.2		
Indicate the intended DOK range of the Grade Level Expectations:		
Indicate the intended DOK of the assessment (list DOK levels): 1-3		
1b. Describe the content knowledge/concepts assessed by the set of items or the		
performance task: systems of linear equations, Pythagorean Theorem, probability,		
proportional reasoning, percent, nets, measures of central tendency		
1c. List the skills/performance assessed: Mathematical Practices 1, 2, and 4		
1d. To what extent do you see a strong content match between the set of items		
reviewed or the task and the corresponding Colorado Academic Standard/s? Use the		
definitions below to select your rating.		
Full match – all tasks or items fully address or exceed the relevant skills and		
knowledge described in the corresponding state standard/s.		
X Close match		
Partial match – many tasks or items partially address the skills and knowledge described		
in the corresponding state standard/s.		

Rigor Level Rating	2	
	Rigor=1, Less Rigor=1	
	Similar Rigor=2, More	
problems from simple to complex.		
make connections among concepts and use mathematics to model situations and solve		
while others have multiple solution paths. In some problems, students are asked to		
support your response: Skill-based problems are included where procedures are clear		
Please provide evidence from both the grade level expectations and assessment to		
for the grade level expectations.		
 Less rigor – most items or the task reviewed are lower than the DOK range indicated 		
for the grade level expectations.		
X Similar rigor – most items or the task reviewed are similar to the DOK range indicated		
range indicated for the grade level expectations.		
More rigorous – most items or the tasks reviewed are at a higher DOK level than the		
expectations? Use the definitions below to select your rating.		
1e . Are the set of items or tasks reviewed as cognitively challenging as the grade level		
	Rating Column	Comments
Aligned to Colorado Academic Standards Rating		
	Match=2; No Match= 1	
	Match=3; Minimal	
	Match=4; Partial	
	Full Match=5; Close	
aligned to standards for grades 6-8.		
Evidence Outcome within grades 6-8. Short tasks numbers one and three are not		
response: In most cases, each individual task aligns closely to a specific GLE and		
Please provide evidence from both the standards and assessment to support your		
in the corresponding state standard/s.		
□ No match – task or most items are not related to the skills and knowledge described		
described in the corresponding state standard/s.		
Minimal match – some tasks or items match some relevant skills and knowledge		

A high quality assessment should be...Scored using Clear Guidelines and Criteria

Scoring Guide Present	Check all that apply:	Comments
X Answer key, scoring template, computerized/machine scored	X	
Generalized Rubric (e.g., for persuasive writing, for all science labs)		
Task-Specific Rubric (only used for the particular task)		
Checklist (e.g., with score points for each part) Track of Checklist (Checklist Checklist)	X	
Teacher Observation Sheet/ Observation Checklist	Rating Column	
2a.Does the rubric/scoring criteria align to Colorado Academic Standards in this assessment. Provide an explanation of your response: Checklist provides points for accuracy for problems but does not specifically reference standards.	Yes=3, Somewhat=2, No=1	
Rubric Aligned to Standards Ratin	g <u>1</u>	
2b. Are the score categories clearly defined and coherent across performance levels? Provide an explanation of your response: Performance levels not given.	Yes=3, Somewhat=2, No=1	
Rubric/Scoring Coherent Rating	g 1	
2c. To what degree does the rubric/scoring criteria address all of the demands within the task or item? Provide and explanation of your response. Specific points awarded for each element of the response. All parts of the assessment were addressed.	High=3, Moderate=2, Low or None=1	
Rubric/Scoring Aligned with Task Rating	g <mark>3</mark>	
2d. Based on your review of the rubric/scoring criteria, do you think the scoring rubric would most likely lead different raters to arrive at the same score for a given response. Rubric very specific about awarding points, thus would most likely lead to consistency among raters.	Yes=3, Somewhat=2, No=1	
Rubric/Scoring Different Raters Same Ratin	g <u>3</u>	

2e. Is there student work (e.g., anchor papers, video, portfolio) which illustrates student mastery? If so, describe. If not, what student work would be needed? No work provided. Authentic student work showing a variety of solution strategies would be beneficial to accompany the scoring check list.	Yes=3, Somewhat=2, No=1	
Student Work Samples Rating	1	

A high quality assessment should be...FAIR and UNBIASED

FAIR and UNBIASED (the areas below should be discussed relative to the needs of ELLs, gifted and talented students, and students with disabilities)	Rating Column	Comments
3a. To what extent are most of the items or the tasks designed and formatted to be visually clear and uncluttered (e.g., use of white space, graphics, and illustrations)? Provide an explanation of your response: Formatting is clear and easy to read with space provided for work. Graphics provided support student understanding of the problem.	High=3, Moderate=2, Low=1	
Clear & Uncluttered Rating	3	
3b. To what extent are most of the items or the task presented in as straightforward a way as possible for a range of learners? Provide an explanation of your response: Problems presented in a clear and straight-forward manner.	High=3, Moderate=2, Low=1	
Straight Forward Rating	3	
3c. To what degree is the vocabulary and context(s) presented by most of the items or task free from cultural or other unintended bias? Provide an explanation of your response: Contexts are of common situations with graphics provided to support understanding.	High=3, Moderate=2, Low=1	
Free of Cultural or Unintended Bias Rating	3	
3d. Does the assessment use appropriate levels of academic language for the grade and content area? Provide an explanation of your response. Grade-level appropriate language is used, e.g. equation, hypotenuse, median.	Yes=3, Somewhat=2, No=1	
Academic Language Rating	3	
3e. Does the assessment limit the usage of words that can be confused with one another (homonyms)? (Examples: ate/eight; sell/cell; allowed/aloud; beet/beat; by/buy). Use of homonyms not a problem.	Yes=3, Somewhat=2, No=1	
Confusing Language Rating	3	
*Please reference "Defining Features of Academic Language in WIDA's Standards" (http://wida.us/searchResults.aspx?cx=0001878867407992537742:bjkids4qwcy&cof=FORID:10&q= Defining%20Features%20of%20Academic%20Language)		
3f. If applicable, what type of accommodations are provided to ensure that English Learners and/or Students with Disabilities can fully access the content represented by the task or set of items reviewed? None addressed. Accommodations are commonly categorized in five ways: presentation, response,		
setting, and timing and scheduling: o Presentation Accommodations —Allow students to access information in ways that do not require them to visually read standard print. These alternate modes of access are		
auditory, multi-sensory, tactile, and visual. o Response Accommodations —Allow students to complete activities, assignments, and assessments in different ways or to solve or organize problems using some type of assistive device or organizer.		
 Setting Accommodations — Change the location in which a test or assignment is given or the conditions of the assessment setting. Timing and Scheduling Accommodations — Increase the allowable length of time to complete an assessment or assignment and perhaps change the way the time is organized. 		
o Linguistic Accommodations—Allow English language learners (ELLs) to access academic construct measured by reducing the linguistic load of an assessment. The accommodation is based on an ELL's limited English language proficiency, which is different than an accommodation based on a student's disability or a cognitive need.		
3g: Are there adequate accommodations permitted for this assessment? Provide an explanation of your response: Not addressed	Yes, Some identified=2; None identified =1	
Adequate Accommodations Allowed Rating 1		
A high quality assessmentIncreases Opportunities to L	earn	

Opportunities to Learn	Rating Column	Comments
(the areas below should also be discussed relative to the needs of ELLs, gifted and		
talented students, and students with disabilities)		

4a. Does this assessment engage a student in thinking that connects to a real world, new context, situation, problem or challenge? Provide an explanation of your response: Mix of problems provide various realistic contexts for using mathematics. Simple real- world problems are included. Engagement Rating	High=3; Moderate=2; Low or None=1 2	
	-	
4b. To what extent do you think the knowledge and skills tested by the assessment can provide good information about what students have learned in the classroom? Provide an explanation of your response: Good information but not connected to specific GLEs or Evidence Outcomes or specific grade level. Most individual items can provide good information about what should have been taught in the classroom.	High=3; Moderate=2; Low or None=1	
Classroom Learning Rating	2	
4c. To what degree do the results from this assessment (scores and student work analysis) foster meaningful dialogue about learning expectations and outcomes with students and parents? Provide an explanation of your response: Scoring checklist does not give a connection to proficiency expectations for a grade level; scoring only recognizes correct answers.	High=3; Moderate=2; Low or None=1	
Learning Expectations/Outcomes Rating	2	
4d. To what extent do you believe the assessment can clearly communicate expectations for academic excellence (e.g., creativity, transference to other content areas or 21st Century skills) to students? Provide an explanation of your response: Provides some opportunity for student to engage in Standards for Mathematical Practice, but does not address more global 21st Century Skills. Even more, extended-response tasks are scaffolded so as to give direction to student solution method.	High=3; Moderate=2; Low or None=1	
Communicate Academic Excellence Rating	2	
4e. Based on the content evaluated by the task or the set of items reviewed, to what extent do you think teachers can use the results (scores and student work analysis) to understand what competency on standard/s look like? Provide an explanation of your response: Worthwhile tasks are included and scoring guide shows grade-level appropriate solution method, but it does not account for reasoning, critical thinking, or alternative solution methods. It also does not explicitly connect to proficiency expectations for a given standards or grade level.	High=3; Moderate=2; Low or None=1	
Competency on Standards Rating	2	
4f: Based on the content evaluated by the task or the set of items reviewed, to what extent do you think teachers can locate where the assessment evidence is represented within the curriculum, student learning objectives, or lesson? Provide an explanation of your response: Challenge of this assessment is that it covers various standards across three grade levels. This makes it more difficult to clarify purpose and alignment.	High=3; Moderate=2; Low or None=1	
Locate Evidence Rating	2	
Summary	Earned	Possible
Standards Rating	4	5
Rigor Rating	2	2
Subtotal	6	7
		85.7%
Rubric Aligned w/Standards Rating	1	3
Rubric/Scoring Coherent Rating	1	3
Rubric/Scoring Aligned with Task Rating	3	3
Inter-rater Reliability Rating	3	3
Student Work Samples Rating	1	3
Subtotal	9	15
Clear & Uncluttered Rating	3	60.0% 3
Straight Forward Rating	3	3
Free of Cultural or Unintended Bias Rating	3	3
Academic Language Rating	3	3
Academic Language Rating	3	3

Confusing Language Rating	3	3
Adequate Accommodations Allowed Rating	1	2
Subtotal	16	17
		94.1%
Engagement Rating	2	3
Reflects Classroom Learning Rating	2	3
Reflects Learning Expectations/Outcomes Rating	2	3
Communicates Academic Excellence Rating	2	3
Competency on Standards Rating	2	3
Locate Evidence Rating	2	3
Subtotal	12	18
		66.7%
Grand Total	43	57
		75.4%

This assessment is: Place an 'X' in the appropriate box

Fully Recommended	
Partially Recommended	Х
Not Recommended	

This assessment is not for a particular grade level, as it spans standards from grades 6-8. It also is not a comprehensive assessment, only highlighting a couple of GLEs and specific evidence outcomes from grades 6-8