High Quality Assessment Content Validity Review Tool

To understand the review process and the use of the review tool, go to: How to use the Assessment Review Tool

ABSTRACT: This set of assessments is a set of summative exams given in NY at the end of 8th grade. However many of the test items have merit if pulled out, especially the written response, to inform instruction and use as mastery evaluation for classroom use. Although as evaluator we tend to score open-ended items higher, there are some good MC items which allow students to apply their knowledge versus clarifying what they have memorized. As a group we would use many if not most of these items in our own classrooms.

Content Area: Science

Name of Assessment: University of the State of New York Intermediate Level Science - Assessment
http://www.nysedregents.org/Grade8/Science/home.html

Reviewer: Content Collaborative

14-Nov-12

Assessment Profile	
Item Types - check all that apply (note: there is often overlap among certain item types):	**Check All That Apply**
Selected Response (multiple choice, true-false, matching, etc.)	X
Short Answer (short constructed response, fill in a graphic organizer or diagram, explain your thinking or solution, make and complete a table, etc.)	X
Extended Response (essay, multi-step response with explanation and rationale required for tasks)	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 …)	X
Scoring Guide/Rubric	X
Sample evidence to show what student performance might look like	X
Materials (if needed to complete the assessment)	X
Estimated time for administration	
Student Directions & Assessment Task/Prompt – what does the student see/use?	X
Other: Rubric to categorize level of mastery 1-4, score conversion sheet	

A high quality assessment should be...Aligned

Alignment	Rating Column	Comments
1a.		This set of released test items
Grade Level(s): 8		from NY address many if not all

Indicate the Colorado Academic Standards and Grade Level Expectations evaluated by the Assessment: SC09-GR.8-S.1-GLE.1; SC09-GR.8-S.1-GLE.2; SC09-GR.8-S.1-GLE.3; SC09-GR.8-S.1-GLE.4; SC09-GR.8-S.2-GLE.1; SC09-GR.8-S.2-GLE.2; SC09-GR.8-S.3-GLE.1; SC09-GR.8-S.3-GLE.2; SC09-GR.8-S.3-GLE.3; SC09-GR.8-S.3-GLE.4; SC09-GR.7-S.1-GLE.1; SC09-GR.7-S.2-GLE.1; SC09-GR.7-S.2-GLE.2; SC09-GR.7-S.2-GLE.3; SC09-GR.7-S.2-GLE.4; SC09-GR.7-S.2-GLE.5; SC09-GR.7-S.3-GLE.1; SC09-GR.7-S.3-GLE.2; SC09-GR.6-S.1-GLE.1; SC09-GR.6-S.1-GLE.2; SC09-GR.6-S.1-GLE.3; SC09-GR.6-S.1-GLE.4; SC09-GR.6-S.2-GLE.1; SC09-GR.6-S.2-GLE.2; SC09-GR.6-S.3-GLE.1; SC09-GR.6-S.3-GLE.2; SC09-GR.6-S.3-GLE.3	2	GLE's for grades 6-8 in the areas of Earth, Physical, and Life Sciences. The reviewers suggest that teachers look at the released items and pull those which match the specific GLE they want to address in their classroom.
Indicate the intended DOK range of the Grade Level Expectations: 1-3		DOK: The MC section of the test ranges from 1-3 though leaning more to the 2's and 3's. The written/ short response section is rated 1-3 as well but has more lower level questions than level 3's.
Indicate the intended DOK of the assessment (list DOK levels) : 1-3 (see comment)		
1b. Describe the content knowledge/concepts assessed by the set of items or the performance task: This test assesses core concepts from earth, life, and physical sciences at the MS level. Items are clearly identifiable as to what science they pertain to.		
1c. List the skills/performance assessed (what are students expected to do?): evaluate data in variety of forms and respond very directly or make inferences, regurgitation of information, problem solving.		
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. □ **Close match –** most tasks or items address the relevant skills and knowledge described in the corresponding state standard/s. □ **Partial match –** many tasks or items partially address the skills and knowledge described in the corresponding state standard/s. □ **Minimal match –** some tasks or items match some relevant skills and knowledge described in the corresponding state standard/s. □ **No match –** task or most items are not related to the skills and knowledge described in the corresponding state standard/s. **Please provide evidence from both the standards and assessment to support your response: As noted in the comments most if not all of the CAS are addressed within the scope of the released exam items at the appropriate DOK's. Teachers do need to be selective for the content and difficulty level of their EO's.**		
	Full Match=5; Close Match=4; Partial Match=3; Minimal Match=2; No Match= 1	
Aligned to Colorado Academic Standards Rating	4	
	Rating Column	**Comments**
1e. Are the set of items or tasks reviewed as cognitively challenging as the grade level expectations? **Use the definitions below to select your rating.** □ **More rigorous –** most items or the tasks reviewed are at a higher DOK level than the range indicated for the grade level expectations. □ **Similar rigor –** most items or the task reviewed are similar to the DOK range indicated 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. **Please provide evidence from both the grade level expectations and assessment to support your response: Overall the skills and knowledge required to show mastery on these items match very well to the DOK's listed for grade level evidence outcomes.**		
	Similar Rigor=2, More Rigor=1, Less Rigor=1	
Rigor Level Rating	2	

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

Scoring Guide Present	Check all that apply:	Comments
☐ **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)	X	
☐ **Checklist** (e.g., with score points for each part)	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: Scoring criteria for MC is simply is search of a correctly matched answer per DOK 1-2. The written response rubric delves into deeper content but all is matched to CAS.**	**Yes=3, Somewhat=2, No=1**	
Rubric Aligned to Standards Rating	**3**	
2b. Are the score categories clearly defined and coherent across performance levels? **Provide an explanation of your response: Between the scoring rubric, the score conversion table, and assignment of mastery rubric it is very clear as to how a student has performed on this exam at all levels.**	**Yes=3, Somewhat=2, No=1**	
Rubric/Scoring Coherent Rating	3	
2c. To what degree does the rubric/scoring criteria address all of the demands within the task or item? **Provide an explanation of your response. The rubric either clearly gives the correct answer or a range of acceptable responses for written items.**	**High=3, Moderate=2, Low or None=1**	
Rubric/Scoring Aligned with Task Rating	**3**	
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. **Provide an explanation of your response. Because of the 3 tiers of rubrics and score conversion worksheets we believe all teachers would rate a student similarly (i.e. standardized)**	**Yes=3, Somewhat=2, No=1**	
Rubric/Scoring Different Raters Same Rating	**3**	
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? In the written item rubric, each question/task has at least 3 samples of acceptable answers.**	**Yes=3, Somewhat=2, No=1**	
Student Work Samples Rating	**3**	

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: Within the MC section of the test, questions are very close together in a 2 column format, where student is to answer left column first then the right column. This could be confusing to a student new to the format. The written item section is very clearly organized with ample answer space.**	High=3, Moderate=2, Low=1	
Clear & Uncluttered Rating	**2**	
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: Most items seemed to clearly indicate what knowledge or application thereof was required to correctly answer the question.**	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: Overall the test is written without bias. This assessment received a 2 here because it does cover human genetics and that can have unintended bias.**	High=3, Moderate=2, Low=1	
Free of Cultural or Unintended Bias Rating	**2**	
3d. Does the assessment use appropriate levels of academic language for the grade and content area? **Provide an explanation of your response. Language is grade level appropriate but does include content specific terms.**	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). **Provide an explanation of your response. We did not note any areas where the language would confuse the student.**	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? **Provide an explanation of your response.**		
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.* o *Setting Accommodations —Change the location in which a test or assignment is given or the conditions of the assessment setting.* o *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. There were no accommodations specifically identified within the test however a teacher could use response, setting, timing, and linguistic accommodations at their discretion.**	Yes, Some identified=2; None identified =1	
Adequate Accommodations Allowed Rating	**1**	

A high quality assessment…Increases Opportunities to Learn

Opportunities to Learn *(the areas below should also be discussed relative to the needs of ELLs, gifted and talented students, and students with disabilities)*	**Rating Column**	**Comments**
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: Those items which have a DOK of 2-3 do provide these opportunities although the DOK 1 and some 2's do not.**	High=3; Moderate=2; Low or None=1	
Engagement Rating	**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: These items would allow a student to clearly demonstrate basic learning/memorization. DOK 3 items would need to be isolated for students to show application of content to new situations.**	High=3; Moderate=2; Low or None=1	
Classroom Learning Rating	**3**	
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: The results would be very assistive for teachers, students, and parents to communicate about a students progress towards mastery.**	High=3; Moderate=2; Low or None=1	
Learning Expectations/Outcomes Rating	**3**	
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: Again some DOK 2 and all 3's provide an opportunity and encourage students to apply knowledge to new situations.**	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: As these tests are written for NY standards it does not necessary demonstrate what mastery on CO standards would look like. Teachers would need to make their own interpretations.**	High=3; Moderate=2; Low or None=1	
Competency on Standards Rating	**1**	
4f: Based on the content evaluated by the task or the set of items reviewed, to what extent do you think teachers can identify what purpose the assessment serves (e.g. diagnostic, report card grades, adjusting instruction, etc.)? **Provide an explanation of your response: This test was written as a summative evaluation. However, items could be pulled and used for a formative, instruction modifying evaluation.**	High=3; Moderate=2; Low or None=1	
Clarity of Purpose Rating	**2**	

Summary	**Earned**	**Possible**
Standards Rating	4	5
Rigor Rating	2	2
Subtotal	6	7
		85.7%
Rubric Aligned w/Standards Rating	3	3
Rubric/Scoring Coherent Rating	3	3
Rubric/Scoring Aligned with Task Rating	3	3
Inter-rater Reliability Rating	3	3

Student Work Samples Rating	3	3
Subtotal	15	15
		100.0%
Clear & Uncluttered Rating	2	3
Straight Forward Rating	3	3
Free of Cultural or Unintended Bias Rating	2	3
Academic Language Rating	3	3
Confusing Language Rating	3	3
Adequate Accommodations Allowed Rating	1	2
Subtotal	14	17
		82.4%
Engagement Rating	2	3
Reflects Classroom Learning Rating	3	3
Reflects Learning Expectations/Outcomes Rating	3	3
Communicates Academic Excellence Rating	2	3
Competency on Standards Rating	1	3
Locate Evidence Rating	2	3
Subtotal	13	18
		72.2%
Grand Total	48	57
		84.2%

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

Fully Recommended	
Partially Recommended	X
Not Recommended	