

**4.1 Utilizes valid and reliable assessment tools that are aligned with standards and benchmarks and that assess meaningful learning in all content areas. (CO: 3.2)**

	<b>Basic (1.0 - 1.9)</b>	<b>Developing (2.0 - 2.9)</b>	<b>Proficient (3.0 - 3.9)</b>	<b>Advanced (4.0)</b>
<i>Criteria related to the assessment environment, test administration, and accuracy of interpreting, which affect reliability, and criteria related to use of multiple assessments, which affect validity, are included in 4.2;</i>				
<b>Alignment</b>	No evidence OR develops behavioral objectives for lessons and unit plans that do not specify all of the following: observable, measurable behaviors; conditions or circumstances important for performing the task, and criteria or level of proficiency for at least 2 lesson plans	Develops behavioral objectives for lessons and unit plans that specify all of the following: observable, measurable behaviors; conditions or circumstances important for performing the task, and criteria or level of proficiency; two accurate objectives are required for a rating of 2.0	Develops behavioral objectives with appropriate behavior, conditions, and criteria for at least two skill domains (cognitive/ academic, affective, psychomotor, social, learning) and different academic areas of responsibility; in unit plans (TWS) objectives are aligned with unit goals	Meets the criteria for "proficient" and shows advanced skills by developing effective objectives for complicated skills that required task analysis to accomplish (e.g., reasoning skills, performance skills, proficiency in creating products, learning dispositions)
	No evidence of assessments OR assessments do not have content validity -- content in the assessment or tool used in the assessment do not match the content in plans or in instruction	Provides evidence of content validity through alignment of standards (expectations, outcomes), lesson objectives, and content in the lessons with the content of the lesson assessment (including rubrics or other tools used in assessment); at least 2 lessons with alignment meet requirement for 2.0	Provides evidence of content validity through alignment of standards (expectations, outcomes), lesson objectives, and content in the lessons with the content of the lesson assessment (including rubrics or other tools used in assessment)	Provides evidence of content validity through alignment of standards (expectations, outcomes), lesson objectives, and content in the lessons with the content of the lesson assessment (including rubrics or other tools used in assessment)
	No evidence of assessments on some lesson plans	Includes assessments on all lesson plans, but some aspects may be undeveloped, incapable of being replicated	Lesson plans consistently include detailed information on assessment that would allow the reader to replicate the plan (consistently = >75%)	Meets criteria for "proficient" consistently, without support, across a variety of types of assessments
<b>Reliable</b>	No evidence of assessments OR assessments are consistently unreliable because of inappropriate length (too short), ambiguous directions and questions, and questions of inappropriate difficulty level	Assessments developed for lessons to be taught generally are appropriate length (some inconsistent errors may be present), unambiguous directions and questions, and questions of appropriate difficulty level	Assessments developed and implemented for lessons and for summative assessment (inconsistent errors may be present) are the appropriate length, directions and questions unambiguous, and questions are an appropriate difficulty level; some evidence exists that s/he analyzes reliability of assessment and makes changes or reflects on changes after using	Consistently develops a variety of different types of assessments that are the appropriate length, with unambiguous directions and questions, and questions appropriate difficulty level; some evidence exists that s/he analyzes reliability of assessment and makes changes independently
	No evidence of assessments OR plans consistently are insensitive to learners' interests, backgrounds (e.g., tasks, non-meaningful examples, unfamiliar vocabulary)	Generally no evidence that assessments have been adapted or planned to address the interests/backgrounds of learners OR some plans may be insensitive to learners' interests, backgrounds, language (e.g., tasks, non-meaningful examples, unfamiliar vocabulary)	Evidence that assessments have been adapted or planned to address the interests/ backgrounds of learners and are sensitive to learners' interests, backgrounds, language (e.g., tasks, non-meaningful examples, unfamiliar vocabulary); evidence in TWS links contextual information to nature of assessments	Meets criteria for "proficient" and provides evidence of adapting assessments to improve their meaningfulness based on student interests, backgrounds, language differences, etc.
	No evidence of interpretation of assessment data	Makes errors in interpreting test results because of bias built into assessment	Evidence that s/he recognizes possible bias in interpreting results of assessments	Consistently considers possibility of different forms of bias in interpreting test results and considers bias in reporting outcomes

<b>Reliable</b>	No evidence of assessments OR plans consistently are insensitive to learners' developmental levels (e.g., nature of tasks, childish examples, inappropriate attentional requirements, too complicated instructions, non-meaningful examples)	Some aspects of plans for assessment may be insensitive to learners' developmental levels, but errors are inconsistent (e.g., nature of tasks, childish examples, inappropriate attentional requirements, too complicated instructions, non-meaningful examples)	Evidence that assessments have been adapted or planned to address the learners' developmental level (e.g., reading level, length of assessment, complexity of directions, examples); evidence in TWS links contextual information to nature of assessments	Meets criteria for "proficient" and provides evidence of adapting assessments to improve their appropriateness for students at different developmental levels
<b>Meaningful Learning</b>	No evidence of assessments OR constructs assessments with frequent (more than 2) consistent errors in: linguistics (syntactic, semantic) and/or conventional errors (spelling, capitalization, punctuation); logic and/or organization; errors which result in failure to communicate meaning; errors indicate failure to edit	Constructs assessments with infrequent errors in any of the areas of writing: linguistics (syntactic, semantic) and/or conventional errors (spelling, capitalization, punctuation); logic and/or organization; errors which result in failure to communicate meaning; sometimes errors indicate failure to edit	Rarely constructs assessments with errors in any of the areas of writing listed in "developing" (<90% have any errors)	Meets criteria for "proficient" -- "advanced"+E21 determined by performance in other areas of the dimension
	No evidence OR makes consistent errors in planning or implementing assessments in any of the following areas:  1.) written questions/items are confusing/poorly worded 2.) assessment/tools fail to assess most important knowledge and/or concepts 3.) assessment lacks sufficient comprehensive scope to lead to clear conclusions 4.) assessment evaluates more than one concept/trait and would lead to unclear conclusion 5.) plans some assessments which do not match level of difficulty of objectives/benchmarks (either too complex or much easier) 6.) assessment is not at appropriate reading level	Plans some assessments which would produce invalid results because of any of the following:  1.) written questions/items are confusing/poorly worded 2.) assessment/tools fail to assess most important knowledge and/or concepts 3.) assessment lacks sufficient comprehensive scope to lead to clear conclusions 4.) assessment evaluates more than one concept/trait and would lead to unclear conclusion 5.) plans some assessments which do not match level of difficulty of objectives/benchmarks (either too complex or much easier) 6.) assessment is not at appropriate reading level	Plans and implements a variety of different assessments across areas of responsibility that produce meaningful results because of adherence in all of the following (with only minor, inconsistent errors): 1.) written questions/items are clearly worded 2.) assessment/tools assess most important knowledge and/or concepts 3.) assessment includes sufficient comprehensive scope to lead to clear conclusions 4.) assessment does not evaluate more than one concept/trait that would lead to unclear conclusion 5.) assessments match level of difficulty of objectives/ benchmarks (either too complex or much easier) 6.) assessment is at appropriate reading level	Meets criteria for "proficient" and demonstrates "advanced" skills by providing evidence of improving the validity of a commercial/previously used instrument by modifying the instrument for any of the criteria listed in the preceding column OR by providing evidence of frequent evaluation of the utility of strategies/tools used in assessment
	No evidence of assessments OR evidence of assessments that evaluate quality of learning	Majority of assessments fail to evaluate quality of learning (e.g., only evaluates number of items incorrect and not nature of responses)	Majority of assessments evaluate quality of learning as demonstrated in assessments themselves and/or in analyses and summaries of student learning	Meets criteria for "proficient" and demonstrates "advanced" skills by routinely monitoring quality of learning for key
	No evidence of assessments OR assessments consistently fail to evaluate higher levels of thinking	Most plans for assessment evaluate rote learning, comprehension, and recall and not higher levels of thinking AND during instruction often assesses recall and comprehension; does not adapt commercial assessments to ensure assessment of higher order skills	Consistently plans and evaluates most important knowledge and critical skills, including all of the following different ability and skill levels (knowledge, comprehension, application, analysis, synthesis, evaluation)	Meets criteria for "proficient" and demonstrates advanced skills at consistently modifying and adapting assessments for diverse learners to assess critical and creative thinking (e.g., tiered assessment)

<b>Practical</b>	No evidence of assessments OR plans/implements assessments that take so much time and energy that they consistently compromise instruction or teacher planning time	Many plans for assessment would require so much time and energy that they would compromise instruction or teacher planning time OR consistently fails to use practical strategies for formative assessment (e.g., choral systems, partner journals, post - gallery tours, fuzzy points, collaborative, comprehension checks)	Plans and implements practical assessments that (with only minor errors) AND uses several different strategies for formative assessment during instruction (e.g., choral systems, partner journals, post - gallery tours, fuzzy points, collaborative, comprehension checks)	Meets criteria for "proficient" and demonstrates "advanced" skills by routinely using a variety of practical strategies and systems for formative assessment AND developing manageable systems for organizing assessment information
------------------	---	--	--	--

## Operationalization/Criteria:

### Guidelines for Admission to Education:

1. Benchmark for admission is that the student can "*develop assessments for lesson plans in which the assessment matches the planned objective.*"
2. Benchmark for admission is a rating of "developing" on evaluation form completed by the K-12 classroom teacher who has worked with the student and a rating of "developing" for the first dimension of the standard based on review of at least 2 lesson plans and self/peer evaluations.
3. A rating below "developing" should always include a brief narrative explaining deficiencies and may result in a support plan.

**Evidence to be Evaluated:** 2 lesson plans and accompanying assessment tools (e.g., rubric); field experience teacher's evaluations, self and peer evaluations

### Guidelines for Admission to Student Teaching:

1. Benchmark is that the student can "*demonstrate development and application of a variety of assessment tools and can reflect on their reliability and validity as used.*"
2. Benchmark at admission to student teaching is "developing" on dimensions 1 and 2.

### Evidence to be Evaluated:

Lessons for different academic areas of responsibility, assessment plans from a unit implemented with students (mini work samples), field experience teachers' evaluations/feedback

### Guidelines for Program Completion/Student Teaching:

1. Required for program completion is a rating in the "proficient" range in all dimensions of the standard.
2. Observe at least three administered lessons and sample lessons from the lesson plan notebook; may need to refer to basal or student materials to determine validity. Evaluate both daily lessons and summative evaluations (e.g., unit tests, portfolio assessments); evaluate the TWS
3. Observe both the quality of the assessment and whether assessment is implemented.
4. Observe range of assessments: written test, observation of learner responses during a lesson, performance assessments, writing assignments, portfolio assessments, etc.
5. Observe assessments in each domain/teaching area (e.g., writing, reading, social studies for elementary teachers).
6. Ask student teacher and cooperating teacher whether administration was typical performance by students/student teacher.
7. Consistency = requires fluency/repetition across different types of assessments and content areas.
8. The OVERALL rating for the standard should be an average of the rating on all dimensions.
9. The narrative for the Inventory should specify an example of a skill/observation that led to the rating, e.g.: *TWS assessments were aligned with content standards in 3 areas and she included a range of types of assessments that evaluated various levels of skill acquisition and quality of learning.*

### Examples of Evidence

Assessments and reflections included in lesson plans notebook, TWS, assessments in eportfolio, direct observation of administration of assessments, interviews with others who have observed assessments

**Rationale:**

- Borich, G.D. (2007). *Effective teaching methods: Research-based practice*, 6th ed. Upper Saddle Rivers, NJ: Merrill/Prentice Hall.
- Gronlund, N.E., & Waught, C.K. (2009). *Assessment of student achievement*, 9th ed. Upper Saddle Rivers, NJ: Allyn & Bacon.
- Herbert, E.A. (2001). *The power of portfolios: What children can teach us about learning and assessment*. San Francisco: Jossey-Bass.
- Kauchak, D.O., & Eggen, P.D. (1998). *Learning and teaching: Research-based methods*, 3rd ed., Needham Heights, MA: Allyn & Bacon.
- Marzano, R., Pickering, D., & McTighe, J. (1993). *Assessing outcomes: Performance assessment using dimensions of learning*. Alexandria, VA: ASCD.
- McLoughlin, J.A. (2008). *Assessing students with special needs*, 7th ed. Upper Saddle Rivers, NJ: Merrill.
- Popham, W.J. *Classroom assessment: What teachers need to know*, 5th ed. Upper Saddle Rivers, NJ: Allyn & Bacon.
- Shepard, L.A. (2000). The role of assessment in a learning culture. *Educational Researcher*, 29(7), 4-14.
- Stiggins, R.J. (2008). *Introduction to Student-Involved Assessment for Learning*, 5th ed. Upper Saddle Rivers, NJ: Allyn & Bacon.
- Taylor, C.T. (1994). Assessment for the measurement of standards: The peril and promise of large scale assessment reform. *American Educational Research Journal*, 31, 231-262.