

Appendix A

Classroom Assessment Literacy Inventory (CALI) (Mertler & Campbell, 2005)

**Description of the inventory:**

This inventory consists of two parts. Part I consists of 35 items related to the seven “Standards for Teacher Competence in the Educational Assessment of Students.” Some of the items are intended to measure general concepts related to testing and assessment, including the use of assessment activities for assigning student grades and communicating the results of assessment to students and parents; other items are related to knowledge of standardized testing and the remaining items are related to classroom assessment. Part II consists of items related to your background as a classroom teacher.

**Directions:**

Please read each item carefully and select the response you think is the best one by shading the corresponding circle. Even if you are not sure of your choice, but you think you know which is best, mark that response.

**PART I**

1. What is the **most** important consideration in choosing a method for assessing student achievement?  
☐ The ease of scoring the assessment.  
☐ The ease of preparing the assessment.  
☐ The accuracy of assessing whether or not instructional objectives were attained.  
☐ The acceptance by the school administration.
2. When scores from a standardized test are said to be “reliable,” what does it imply?  
☐ Student scores from the test can be used for a large number of educational decisions.  
☐ If a student retook the same test, he or she would get a similar score on each retake.  
☐ The test score is a more valid measure than teacher judgments.  
☐ The test score accurately reflects the content of what was taught.
3. Mrs. Bruce wished to assess her students' understanding of the method of problem solving she had been teaching. Which assessment strategy below would be **most** valid?  
☐ Select a textbook that has a "teacher's guide" with a test developed by the authors.  
☐ Develop an assessment consistent with an outline of what she has actually taught in the class.  
☐ Select a standardized test that provides a score on problem solving skills.  
☐ Select an instrument that measures students' attitudes about problem solving strategies.
4. What is the **most** effective use a teacher can make of an assessment that requires students to show their work (e.g., the way they arrived at a solution to a problem or the logic used to arrive at a conclusion)?

- ☐ Assigning grades for a unit of instruction on problem solving.
  - ☐ Providing instructional feedback to individual students.
  - ☐ Motivating students to attempt innovative ways to solve problems.
  - ☐ None of the above.
5. Ms. Green, the principal, was evaluating the teaching performance of Mr. Williams, the fourth grade teacher. One of the things Ms. Green wanted to learn was if the students were being encouraged to use higher order thinking skills in the class. What documentation would be the **most** valid to help Ms. Green to make this decision?
- ☐ Mr. Williams' lesson plans.
  - ☐ The state curriculum guides for fourth grade.
  - ☐ Copies of Mr. Williams' unit tests or assessment strategies used to assign grades.
  - ☐ Worksheets completed by Mr. Williams' students, but not used for grading.
6. A teacher wants to document the validity of the scores from a classroom assessment strategy she plans to use for assigning grades on a class unit. What kind of information would provide the **best** evidence for this purpose?
- ☐ Have other teachers judge whether the assessment strategy covers what was taught.
  - ☐ Match an outline of the instructional content to the content of the actual assessment.
  - ☐ Let students in the class indicate if they thought the assessment was valid.
  - ☐ Ask parents if the assessment reflects important learning outcomes.
7. Which of the following would most likely **increase** the reliability of Mrs. Lockwood's multiple choice end-of-unit examination in physical science?
- ☐ Use a blueprint to develop the test questions.
  - ☐ Change the test format to true-false questions.
  - ☐ Add more items like those already on the test.
  - ☐ Add an essay component.
8. Ms. Gregory wants to assess her students' skills in organizing ideas rather than just repeating facts. Which words should she use in formulating essay exercises to achieve this goal?
- ☐ compare, contrast, criticize
  - ☐ identify, specify, list
  - ☐ order, match, select
  - ☐ define, recall, restate
9. Mr. Woodruff wanted his students to appreciate the literary works of Edgar Allen Poe. Which of his test items shown below will **best** measure his instructional goal?
- ☐ "Spoke the raven, nevermore." comes from which of Poe's works?
  - ☐ True or False: Poe was an orphan and never knew his biological parents.



Edgar Allen Poe wrote:

1. Novels
2. Short stories
3. Poems
4. All of the above.



Discuss briefly your view of Poe's contribution to American literature.

10. Several students in Ms. Atwell's class received low scores on her end-of-unit test covering multi-step story problems in mathematics. She wanted to know which students were having similar problems so she could group them for instruction. Which assessment strategy would be **best** for her to use for grouping students?



Use the test provided in the "teacher's guide."



Have the students take a test that has separate items for each step of the process.



Look at the student's records and standardized test scores to see which topics the students had not performed well on previously.



Give students story problems to complete and have them show their work.

11. Many teachers score classroom tests using a 100-point percent correct scale. In general, what does a student's score of 90 on such a scale mean?



The student answered 90% of the items on this test correctly.



The student knows 90% of the instructional content of the unit covered by this test.



The student scored higher than 90% of all the students who took the test.



The student scored 90% higher than the average student in the class.

12. Students in Mr. Jakman's science class are required to develop a model of the solar system as part of their end-of-unit grade. Which scoring procedure below will **maximize** the objectivity of assessing these student projects?



When the models are turned in, Mr. Jakman identifies the most attractive models and gives them the highest grades, the next most attractive get a lower grade and so on.



Mr. Jakman asks other teachers in the building to rate each project on a 5-point scale based on their quality.



Before the projects are turned in, Mr. Jakman constructs a scoring key based on the critical features of the projects as identified by the highest performing students in the class.



Before the projects are turned in, Mr. Jakman prepares a model or blueprint of the critical features of the product and assigns scoring weights to these features. The models with the highest scores receive the highest grade.

13. At the close of the first month of school, Mrs. Friend gives her fifth grade students a test she developed in social studies. Her test is modeled after a standardized social studies test. It presents passages and then asks questions related to understanding and problem definition. When the test was scored, she noticed that two of her students—who had been performing well in their class assignments—scored much lower than other students. Which of the following types of additional information which would be most helpful in interpreting the results of this test?

- ☐ The gender of the students.
- ☐ The age of the students.
- ☐ Reliability data for the standardized social studies test she used as the model.
- ☐ Reading comprehension scores for the students.
14. Frank, a beginning fifth grader, received a G. E. (grade equivalent score) of 8.0 on the Reading Comprehension subtest of a standardized test. This score should be interpreted to mean that Frank
- ☐ can read and understand 8th grade reading level material.
- ☐ scored as well as a typical beginning 8th grader scored on this test.
- ☐ is performing in Reading Comprehension at the 8th grade level.
- ☐ will probably reach maximum performance in Reading Comprehension at the beginning of the 8th grade.
15. When the directions indicate each section of a standardized test is timed separately, which of the following is **acceptable** test-taking behavior?
- ☐ John finishes the vocabulary section early; he then rechecks many of his answers in that section.
- ☐ Mary finishes the vocabulary section early; she checks her answers on the previous test section.
- ☐ Jane finishes the vocabulary section early; she looks ahead at the next test section but does not mark her answer sheet for any of those items.
- ☐ Bob did not finish the vocabulary section; he continues to work on that section when the testing time is up.
16. Ms. Camp is starting a new semester with a factoring unit in her Algebra I class. Before beginning the unit, she gives her students a test on the commutative, associative, and distributive properties of addition and multiplication. Which of the following is the **most** likely reason she gives this test to her students?
- ☐ The principal needs to report the results of this assessment to the state testing director.
- ☐ Ms. Camp wants to give the students practice in taking tests early in the semester.
- ☐ Ms. Camp wants to check for prerequisite knowledge in her students before she begins the unit on factoring.
- ☐ Ms. Camp wants to measure growth in student achievement of these concepts, and scores on this test will serve as the students' knowledge baseline.
17. To evaluate the effectiveness of the mathematics program for her gifted first graders, Ms. Allen gave them a standardized mathematics test normed for third graders. To decide how well her students performed, Ms. Allen compared her students' scores to those of the third-grade norm group. Why is this an **incorrect** application of standardized test norms?
- ☐ The norms are not reliable for first graders.

- ☐ The norms are not valid for first graders.
- ☐ Third grade mathematics items are too difficult for first graders.
- ☐ The time limits are too short for first graders.
18. When planning classroom instruction for a unit on arithmetic operations with fractions, which of these types of information have **more** potential to be helpful?  
norm-referenced information: describes each student's performance relative to other students in a group (e.g., percentile ranks, stanines), or  
criterion-referenced information: describes each student's performance in terms of status on specific learning outcomes (e.g., number of items correctly answered for each specific objective)
- ☐ Norm-referenced information.
- ☐ Criterion-referenced information.
- ☐ Both types of information are equally useful in helping to plan for instruction.
- ☐ Neither, test information is not useful in helping to plan instruction.
19. Students' scores on standardized tests are sometimes inconsistent with their performances on classroom assessments (e.g., teacher tests or other in-class activities). Which of the following is **not** a reasonable explanation for such discrepancies?
- ☐ Some students freeze up on standardized tests, but they do fine on classroom assessments.
- ☐ Students often take standardized tests less seriously than they take classroom assessments.
- ☐ Standardized tests measure only recall of information while classroom assessments measure more complex thinking.
- ☐ Standardized tests may have less curriculum validity than classroom assessment.
20. Elementary school teachers in the Baker School system collectively designed and developed new curricula in Reading, Mathematics, and Science that is based on locally developed objectives and objectives in state curriculum guides. The new curricula were not matched directly to the content of the fourth grade standardized test. A newspaper reports the fourth grade students in Baker Public Schools are among the lowest scoring districts in the State Assessment Program. Which of the following would **invalidate** the comparison between Baker Public Schools and other schools in the state?
- ☐ The curriculum objectives of the other districts may more closely match those of the State Assessment.
- ☐ Other school systems did not design their curriculum to be consistent with the State Assessment test.
- ☐ Instruction in Baker schools is poor.
- ☐ Other school systems have different promotion policies than Baker.
21. Which of the following choices typically provides the **most** reliable student-performance information that a teacher might consider when assigning a unit grade?
- ☐ Scores from a teacher-made test containing two or three essay questions related directly to instructional objectives of the unit.
- ☐ Scores from a teacher-made 20 item multiple-choice test designed to measure the specific

instructional objectives of the unit.

☐

Oral responses to questions asked in class of each student over the course of the unit.

☐

Daily grades designed to indicate the quality of in-class participation during regular instruction.

22. A teacher gave three tests during a grading period and she wants to weight them all equally when assigning grades. The goal of the grading program is to rank order students on achievement. In order to achieve this goal, which of the following should be **closest** to equal?

☐

Number of items.

☐

Number of students taking each test.

☐

Average scores.

☐

Variation (range) of scores.

23. When a parent asks a teacher to explain the basis for his or her child's grade, the teacher should

☐

explain that the grades are assigned fairly, based on the student's performance and other related factors.

☐

ask the parents what they think should be the basis for the child's grade.

☐

explain exactly how the grade was determined and show the parent samples of the student's work.

☐

indicate that the grading scale is imposed by the school board and the teachers have no control over grades.

24. Which of the following grading practices results in a grade that **least** reflects students' achievement?

☐

Mr. Jones requires students to turn in homework; however, he only grades the odd numbered items.

☐

Mrs. Brown uses weekly quizzes and three major examinations to assign final grades in her class.

☐

Ms. Smith permits students to redo their assignments several times if they need more opportunities to meet her standards for grades.

☐

Miss Engle deducts 5 points from a student's test grade for disruptive behavior.

25. During the most recent grading period, Ms. Johnson graded no homework and gave only one end-of-unit test. Grades were assigned only on the basis of the test. Which of the following is the **major** criticism regarding how she assigned the grades?

☐

The grades probably reflect a bias against minority students that exists in most tests.

☐

Decisions like grade assignment should be based on more than one piece of information.

☐

The test was too narrow in curriculum focus.

☐

There is no significant criticism of this method providing the test covered the unit's content.

26. In a routine conference with Mary's parents, Mrs. Estes observed that Mary's scores on the state

assessment program's quantitative reasoning tests indicate Mary is performing better in mathematics concepts than in mathematics computation. This probably means that

- ☐ Mary's score on the computation test was below average.
- ☐ Mary is an excellent student in mathematics concepts.
- ☐ the percentile bands for the mathematics concepts and computation tests do not overlap.
- ☐ the mathematics concepts test is a more valid measure of Mary's quantitative reasoning ability.

27. Many states are revising their school accountability programs to help explain differences in test scores across school systems. Which of the following is **not** something that needs to be considered in such a program?

- ☐ The number of students in each school system.
- ☐ The average socio-economic status of the school systems.
- ☐ The race/ethnic distribution of students in each school system.
- ☐ The drop-out rate in each school systems.

28. The following standardized test data are reported for John.

**Subject -- Stanine Score**

*Vocabulary -- 7*

*Mathematics Computation -- 7*

*Social Studies -- 7*

Which of the following is a **valid** interpretation of this score report?

- ☐ John answered correctly the same number of items on each of the three tests.
- ☐ John's test scores are equivalent to a typical seventh grader's test performance.
- ☐ John had the same percentile rank on the three tests.
- ☐ John scored above average on each of the three tests.

29. Mr. Klein bases his students' grades mostly on graded homework and tests. Mr. Kaplan bases his students' grades mostly on his observation of the students during class. A major difference in these two assessment strategies for assigning grades can **best** be summarized as a difference in

- ☐ formal and informal assessment.
- ☐ performance and applied assessment.
- ☐ customized and tailored assessment.
- ☐ formative and summative assessment.

30. John scored at the 60th percentile on a mathematics concepts test and scored at the 57th percentile on a test of reading comprehension. If the percentile bands for each test are five percentile ranks wide, what should John's teacher do in light of these test results?

- ☐ Ignore this difference.
- ☐ Provide John with individual help in reading.
- ☐ Motivate John to read more extensively outside of school.

- ☐ Provide enrichment experiences for John in mathematics, his better performance area.

31. In some states testing companies are required to release items from prior versions of a test to anyone who requests them. Such requirements are known as

- ☐ open-testing mandates.
- ☐ gag rules.
- ☐ freedom-of-information acts.
- ☐ truth-in-testing laws.

32. Mrs. Brown wants to let her students know how they did on their test as quickly as possible. She tells her students that their scored tests will be on a chair outside of her room immediately after school. The students may come by and pick out their graded test from among the other tests for their class. What is wrong with Mrs. Brown's action?

- ☐ The students can see the other students' graded tests, making it a violation of the students' right of privacy.
- ☐ The students have to wait until after school, so the action is unfair to students who have to leave immediately after school.
- ☐ Mrs. Brown will have to rush to get the tests graded by the end of the school day, hence, the action prevents her from using the test to identify students who need special help.
- ☐ The students who were absent will have an unfair advantage, because her action allows the possibility for these students to cheat.

33. A state uses its statewide testing program as a basis for distributing resources to school systems. To establish an equitable distribution plan, the criterion set by the State Board of Education provides additional resources to every school system with student achievement test scores above the state average. Which cliché **best** describes the likely outcome of this regulation?

- ☐ Every cloud has its silver lining.
- ☐ Into each life some rain must fall.
- ☐ The rich get richer and the poor get poorer.
- ☐ A bird in the hand is worth two in the bush.

34. In a school where teacher evaluations are based in part on their students' scores on a standardized test, several teachers noted that one of their students did not reach some vocabulary items on a standardized test. Which teacher's actions is considered **ethical**?

- ☐ Mr. Jackson darkened circles on the answer sheet at random. He assumed Fred, who was not a good student, would just guess at the answers, so this would be a fair way to obtain Fred's score on the test.
- ☐ Mr. Hoover filled in the answer sheet the way he thought Joan, who was not feeling well, would have answered based on Joan's typical in-class performance.
- ☐ Mr. Stover turned in the answer sheet as it was, even though he thought George, an average student, might have gotten a higher score had he finished the test.
- ☐ Mr. Lund read each question and darkened in the bubbles on the answer sheet that represented what he believed Felicia, a slightly below average student, would select as the

correct answers.

35. Mrs. Overton was concerned that her students would not do well on the State Assessment Program to be administered in the Spring. She got a copy of the standardized test form that was going to be used. She did each of the following activities to help increase scores. Which activity was **unethical**?
- ☐ Instructed students in strategies on taking multiple choice tests, including how to use answer sheets.
  - ☐ Gave students the items from an alternate form of the test.
  - ☐ Planned instruction to focus on the concepts covered in the test.
  - ☐ None of these actions are unethical.

## PART II

36. What is your gender?
- ☐ female
  - ☐ male
37. Which of the following is the **most appropriate** description of the level at which you teach?
- ☐ elementary – primary (K – grade 3)
  - ☐ elementary – intermediate (grades 4 – 6)
  - ☐ elementary (K – 6)
  - ☐ middle (grades 6 – 8)
  - ☐ high (grades 9 – 12)
  - ☐ secondary (grades 6 – 12)
  - ☐ K – 12
  - ☐ other
38. Which **best** describes the educational level you have attained?
- ☐ B.A. or B.S.
  - ☐ M.A. or M.S.
  - ☐ Specialist
  - ☐ Ed.D.
  - ☐ Ph.D.

39. **Including the current year**, how many years of experience do you have as a classroom teacher?
- ☐ 1 – 5 years
- ☐ 6 – 10 years
- ☐ 11 – 15 years
- ☐ 16 – 20 years
- ☐ 21 – 25 years
- ☐ 26 – 30 years
- ☐ more than 30 years
40. To the best of your knowledge, did you take a **standalone** course in classroom assessment as part of your undergraduate teacher preparation?
- ☐ yes
- ☐ no
41. Which of the following best describes your perception of the level of preparation **for the overall job of being a classroom teacher** that resulted from your undergraduate teacher preparation program?
- ☐ very unprepared
- ☐ somewhat unprepared
- ☐ somewhat prepared
- ☐ very prepared
42. Which of the following best describes your perception of the level of preparation **for assessing student performance** that resulted from your undergraduate teacher preparation program?
- ☐ very unprepared
- ☐ somewhat unprepared
- ☐ somewhat prepared
- ☐ very prepared

#### Appendix B: Interview Questions

1. To what extent are you familiar with alternative assessment?
2. How often do you use alternative assessment in your classes?
3. If you ever use alternative assessment, what forms of alternative assessment do you use?
4. If no, why don't you use alternative assessment?
5. If you do not use alternative assessment, what do you use instead?

**Thank You**