

A Practical Guide to Assessment Literacy

Dr. Charles E. Notar (Emeritus)
Jacksonville State University

Dr. Patricia E. Allanson
Liberty University

ABSTRACT

In the age of accountability, it is important for educators to understand key principles of assessment practices. Too often terms are used interchangeably. This article reviews core knowledge of assessment literacy as used within educational context. History and definition of assessment terms are provided. Tables provide examples of assessment tools and characteristics of formative and summative assessments.

Key Terms: Assessment, Measurement, Test, Evaluation, Assessment Literacy, Assessment systems, Authentic Assessment, Formative Assessment, Performance-based Assessment, Summative Assessments

INTRODUCTION

As student's advance through the K-12 education system they will have taken an average of 112.3 mandated standardized tests (Strauss, 2015). This number does not include optional tests such as diagnostic test for exceptional students or English learners, district mandated tests, or teacher developed tests. The basic premise for standardized testing, subject-specific knowledge achievement tests or scholastic readiness aptitudes tests, is to provide a method to evaluate student performance on individual state standards. These standardize test only provide a snapshot of a year's worth of instruction, therefore, it is important for educators to know the fundamental differences between measurement, assessment and evaluation as they are used within educational context throughout the year. This core knowledge is referred to as *assessment literacy* and defined as "the ability to create and use valid and reliable assessments as a classroom teacher to facilitate and communicate student learning" (Gareis & Grant, 2015, p. 178)

Historically speaking, standardized testing has been around for a long time used in determining governmental positions in Imperial China to placing new recruits in World War I. The SAT that we know today was developed in 1926, and in 1936 IBM revolutionized the way students take tests also known as the infamous "bubble sheets". In the past couple of decades, two major policies (No Child Left Behind and Common Core State Standards Initiative) were passed that take standardized testing to a new level of "accountability". According to proponents, standardized testing measures student achievement and holds teachers and schools accountable to taxpayers, however, at the same time, opponents argue that standardized testing promotes "teaching to the test" and a decline in students' critical thinking skills. Although standardized testing is attractive because it is a quick way to quantify and compare how students, schools and educators are performing, there are many other ways to capture students' abilities and readiness. Of the many aspects of teaching, "quality classroom assessment plays a critical role in affecting student learning and has a research-based set of best practices", (Frey, 2014, p.3). However, after several centuries of education, the definition of the terms assessment, measurement, test, and evaluation are still being debated and used

interchangeably. Mertler (2003) provides one of the clearest definition and purpose of these terms as they pertain to classroom practice. Simply stated, *assessments* are used to obtain information, collection of data is *measurement*, the instruments used for measurement are *tests*, and *evaluation* is an interpretation. These four terms, their definitions and as they relate within the context of the classroom and assessment literacy, will be the focus of the remaining content of this article.

DEFINITIONS OF KEY TERMS

Assessment

Assessment is thought by many educators as one in the same with evaluation, measurement, and testing because of their similarities, and is used in a variety of ways. Classroom assessments are part of a process of measurement, evaluation and testing tools, and referred to as an *assessment system*. Assessment systems are defined by Mertler (2003) as "all the systematic methods and procedures that are used to obtain information about behaviors and upon which educational decisions are based" (p. 7).

A synonym for assess is to measure, therefore when we assess something, we are measuring it. An example of such a practice is assessing for exceptionalities or verbal and quantitative skills. Assessments are generally used for planning, conducting, and evaluating instruction; diagnosing student difficulties; placing students (grouping based on ability, sequencing of coursework, etc.); providing feedback; and grading and evaluating learning. Assessments are generally tied to an objective, goal, or skill and used for many educational purposes such as student aptitude, motivation, self-concept, and achievement.

Tests

Tests are the catalyst in which assessments are performed and referred to as a collection instrument for accumulating information on student performance for a predetermine set of cognitive skills. They are in fact assessments, however, not all assessments are tests as will be discussed later. Tests are a special form of assessments administered under manufactured circumstances and connected to goals or objectives from which it was designed. A test can be a formal set of questions, or performance task generally administered to groups of students. Tests are given at the end of a lesson, unit, or some point in the year to assess progress of goal attainment.

Measurement

Measurement is defined as "a process involving a structured situation that includes samples of particular characteristics or behaviors that results in a numerical or narrative score" (Mertler, 2003, p. 7) and is generally quantitative in nature. Measurement involves observing an object or condition that includes or excludes criteria from specific measurement scale sets. The four basic measurement scales, discussed briefly below, include nominal measurement, ordinal measurement, interval measurement, and ratio measurement.

Nominal measurement scales have no numerical value and represent categorical descriptive variables. Examples of variables measured on a nominal scale include gender or religious affiliations. Ordinal scales are used to place the measurement variables in order and ranking for example academic grades or the results of a race (1st, 2nd, 3rd places). The third measurement scale, interval, involves the properties of equal intervals representing equal and meaningful differences such as in the case of measuring the distance of a football field. The distance between 10 yards and 40 yards is the same as the distance between 50 yards and 80 yards. A thermometer is an example of a ratio measurement scale. It is like ordinal

measurement with one important distinction – it includes the number zero which allows for a comparison between negative and positive values.

Evaluation

According to Mertler (2003), *evaluation* is “the process of making a *value judgment* about student skills or capabilities. Evaluation goes beyond measurement not only to quantify performance, but also to judge the merits of that performance” (p. 7). Although used interchangeably with measurement, evaluation is different from measurement in regards that evaluation uses the results of measurement and assessments in specific ways to make valued decisions or quantified judgements. Evaluation is what occurs after a measurement or assessment activity is implemented and requires extensive decision making on the part of the educator. Evaluations are generally used to make informed decisions about students, instructors, or academic program effectiveness.

A distinctive characteristic of evaluation is the inherent value placed on results to determine the “worthiness, appropriateness, goodness, validity, legality, etc.” (Kizlik, 2014, para. 6) of the assessment or measurement. To evaluate a situation, a criteria or purpose must be initialized in some valid and reliable way. You can collect measurements, for example the average temperature of the classroom, but this simple measurement does not provide information as to what constitutes an ideal classroom temperature for learning.

Another characteristic of evaluation is its comparative nature. As educators, we are continuously evaluating in comparison to the intended purposes (objective, standard, behavior, progress, etc.) and the obtained results. Student achievement is annually measured and compared to established national norms or state standards.

Kizlik (2014) summarizes the key terms as follows: “we measure distance, we assess learning, and we evaluate results in terms of some set of criteria”, (para. 7). Each method is part of an *assessment system* and should be used comprehensively with other samples of student work both formally and informally.

Context in the Classroom

Assessment systems are used to develop baseline data by educators to communicate performance and are an integral part of effective instruction. Over the past twenty years, there have been major shifts in assessment practices. A primary shift leans toward summative learning as in “assessments *of* learning” (student-centered) whereas “assessments *for* learning” is directed towards data driven instruction (teacher-centered). A third shift implies that “assessment *as* learning” should be student centered through self-reflection and self-regulation (Wang, 2017). Assessments have moved away from using results to compare students with other students to comparing individual performance based on some set criteria, giving students multiple opportunities to demonstrate knowledge, and opportunities to participate in the assessment process.

Figure 1 provides a basic list (not all inclusive) of assessment system tools available to educators. As indicated, there are many tools, and a variety of theoretical approaches to assess student learning depending on purpose.

- “Assessment can be used for learning, to actually increase learning,
- Assessment can be used to improve instruction and allow students to control their own learning,
- Assessments can be designed to mirror the real world with the real world's accompanying realistic expectations and evaluations, and

- Assessments can finally be designed to meaningfully include everyone in today's diverse classrooms". (Frey, 2014).

Assessment Approaches

Traditionally most assessments are paper-pencil tests administered after a unit or lesson to demonstrate mastery of skills and important concepts. According to the American Association for Higher Education (2018), "assessment is most effective when it reflects an understanding of learning as multidimensional, dynamic, contextual, and experiential" (para. b). Assessments should be designed in a fair and equitable manner to address diverse student needs, designed to assess learning outcomes, and provide information that is meaningful and relevant. When choosing an assessment approach, teachers should consider assessment *for*, *as*, and *of* to drive their choices:

- Assessment for learning – used to design and revise instruction and to know where students are "at".
- Assessment as learning- used to help students understand how they learn and have control in their own learning.
- Assessment of learning – used to reach conclusions about how much learning has occurred after instruction. This is the type of assessment most educators are familiar with and what has been most used in the past (Frey, 2014).

When designing assessments, teachers should always take into consideration those approaches that encourage long-term learning and are both engaging and intrinsically meaningful. The focus, according to Earl (2013), is to use assessments "as the vehicle for helping students develop and practice the necessary skills to become critical thinkers who are comfortable with reflection and the critical analysis of their learning" (p. 28). Teachers also must consider if the assessment type will be administered in a group or individual setting, or on a formal or informal basis. Formal group assessments, usually pre-determined by administration and lack spontaneity, may include text and curriculum-embedded tests, commercial criterion and normed tests, rating scales, performance tests, and questionnaires. Performance assessments, standardized norm and criterion-referenced tests, and curriculum-embedded tests are individual formal assessment options. There are several options to assess students and groups informally, which are more spontaneous and less obvious. Examples of this type of assessment includes oral questions, writing samples, portfolios, self-assessments, error analysis, etc. (Table 1)

Anecdotal records	Projects
Achievement test	Portfolios (documentation of progress, showcase)
Aptitude test	Progress chart/reports
Attitude test	Profiles
Checklist	Questioning (Oral)
Class groups	Quizzes
Conference/interviews	Records (behavior, class, conduct, daily, group, student, test)
Daily record	Report cards
Diagnostic testing	Research projects
Demonstrations	Rubric
Documents (homework, seat work, etc.)	Self-reporting
Exams	Simulations
Grades	Speaking
Grade book	Standardized tests (Achievement, Aptitude, and Attitude test)
Hands on testing (Performance)	Teacher made tests
Observation	Writing
Oral presentation	
Process assessment	
Product assessment	

Table 1: Assessment, Evaluation, Measurement tools

Most often, teachers will elect to use one of the four major assessment approaches: summative assessments, formative assessments, performance-based assessments, and authentic assessments. Dixson and Worrell (2016) provide an overview of summative and formative assessment use in the classroom, the purpose for each, and how the four approaches are related (see Table 2). It is important to note that many assessments can be used interchangeably meaning a formative assessment can be designed for use as a summative assessment and summative assessments used as a formative assessment.

Summative Assessments

Summative assessments are used for formally assessing student knowledge and are usually administered at the end of a unit, lesson, or year of study. The data collected from summative assessments indicates progress and achievement in skills, content knowledge, and understanding. Summative assessments are associated most often as a way for state or districts to gauge student learning of content standards and accountability measures (AYP). Since summative assessments are used to evaluate progress and programs after instruction, they do not yield information that is useful for the classroom teachers during instruction to provide for modifications or interventions (Garrison & Ehringhaus, 2013).

Characteristics of Formative and Summative Assessments

<i>Characteristic</i>	<i>Formative Assessment</i>	<i>Summative Assessment</i>
Purpose	To improve teaching and learning To diagnose student difficulties	Evaluation of learning outcomes Placement, promotion decisions
Formality	Usually informal	Usually formal
Timing of administration	Ongoing, before and during instruction	Cumulative, after instruction
Developers	Classroom teachers to test publishers	Classroom teachers to test publishers
Level of stakes	Low-stakes	High-stakes
Psychometric rigor	Low to high	Moderate to High
Types of questions asked	What is working What needs to be improved How can it be improved	Does student understand the material Is the student prepared for next level of activity
Examples	Observations Homework Question and answer sessions Self-evaluations Reflections on performance Curriculum-based measures	Projects Performance assessments Portfolios Papers In-class examinations State and national tests

Table 2: Characteristics of Formative and Summative Assessments

Formative Assessments

Black and Wiliam (2010) define formative assessments as “activities undertaken by teachers—and by their students in assessing themselves— that provide information to be used as feedback to modify teaching and learning activities” (p. 82). Teachers use the feedback from formative assessments to adjust instructional strategies and learning behaviors. The purpose of a formative assessment is to increase student learning and works best when used as an assessment *for* learning rather than an assessment *of* learning. Formative assessments should be used to heavily involve students in their own learning process and thought of in terms of “practice. Formative assessments are traditionally not assigned a grade but used to provide feedback of where improvements or adjustment can be made in the learning process leading up to a final summative assessment. Formative assessment: A basic foundation for teaching and learning (Akpan, Notar, & Padgett, 2012).

If students received a grade every time they initially learned a new concept, the final grade would not accurately reflect learning and lessen confidence and motivation to learn (Garrison & Ehringhaus, 2013).

Performance-based Assessments (PBA)

This type of assessment is considered the oldest form of testing commonly used throughout the 1800s prior to being replaced by multiple choice test items around 1914. Mertler (2003) provides a brief definition of performance-based assessments: assessments that address real life situations. Where multiple choice type assessments identify how well a student knows information (recall, identify, list, match), performance-based assessments show how students can apply their knowledge (classify, compare, analyze, evaluate). Performance-based assessments go beyond measuring lower level thinking skills through application and evaluation. Research findings suggest that performance-based assessments increase students’ vocabulary and writing skills while developing motivation and self-confidence in addition to improving teaching practices (Espinosa, 2015). The typical use for performance-based assessments is to assess skills or ability through the “3 P’s: Performance, portfolios and products” (Madaus, & O’Dwyer, 1999, p. 688).

Authentic Assessment

Performance-based assessments are often thought of as one in the same with authentic assessment, however, not all authentic assessments are considered performance-based assessments. The key difference between performance-based assessments and authentic assessments is that the latter requires “students to perform in ways that are valued outside of the classroom” (Frey, 2014, p. 7)- in simpler terms authentic assessments align with real-world application and higher-order thinking skills. More specifically “Assessment is authentic when the tasks, content, expectations, and evaluation methods of the assessment are similar to the meaningful tasks, content, expectations, and evaluation methods outside the classroom in the real world” (p. 203). Authentic assessments include context that is realistic, performance-based, and cognitively complex. Scoring criteria for authentic assessments are understood and sometimes developed by students either through multiple indicators or portfolios.

Involving Students in the Development of Assessment Criteria

Performance-based assessments and authentic assessments often leads to the involvement of students in developing assessment criteria. One advantage of student involvement is that it can help motivate students to learn (Butler & McMunn, 2011). At the center of any assessment must be the student, and one way to harness the power of assessment is to involve them in the process. Utilizing student-centered assessment practices align with Vygotsky’s (1978) theoretical framework (zone of proximal development) where students develop deeper cognitive understanding by engaging in problem solving and tasks guided by their teacher. Using this approach enables students to draw on their current understanding and construct new knowledge.

Assessment should be thought of in terms of what is *done* with students and not *to* the students. Assessment used as a collaborative endeavor, either teacher-student or student-student, maximizes student engagement, motivation to learn, and creates ownership. According to Chan et al. (2014), teachers who promote student ownership of learning enable them in the following ways:

- Being informed about their learning goals, in terms they can understand, from the very beginning of the teaching and learning process;
- Using accurate assessment information to become confident in themselves as learners;
- Receiving frequent feedback that provides them with specific insights as to how to improve;
- Engaging in regular self-assessment, with standards held constant, so that they can watch themselves grow over time; and
- Actively communicating with their teacher and their families about their achievement status and improvement” (p. 106).

An underlying goal of assessment practices is to promote student ownership of their own learning. This of course will demand a shift in practice moving away from a teacher-centered to a more student-centered classroom environment. In doing so, students will know where they are going, where they currently are, and how to close any gaps in their own learning.

Assessment as Reflection (Self-Assessment)

One approach to developing a student-centered classroom environment is to provide authentic and formative assessments that require students to reflect on their own learning process. Reflective practice is accomplished when students “monitor and evaluate the quality of their thinking and behavior when learning and identify strategies that improve their understanding and skills (Ratminingsih, Artini, & Padmadewi, 2017). Through this metacognitive practice, students are involved in their own learning, they can determine what they know and do not

know, and they are able to identify next steps in their progress (Garrison & Ehringhaus, 2013). The students begin to set higher goals and improve their critical thinking skills. Other benefits of reflective practice are promoting intrinsic motivation and empowering students in their own personal learning.

CONCLUSION

Better understanding of the choices teachers make when testing students can be of great benefit to those concerned with improving the quality of classroom assessment. Today's classroom assessments are multifaceted meaning that they overlap and can be derived from different theoretical frameworks focusing on purpose and consistency with intended goals. A formative assessment can be used to provide feedback to both teachers and students and can also be performance based and authentic. It is important for teachers to know the different types of assessments and their appropriate purposes. Assessments must not be thought of as something done to students but as an important part of instructional practices, student learning tools, and a key to learning in its own right.

References

- Akpan, J., Notar, C., & Padgett, S. (2012). Formative assessment: A basic foundation for teaching and learning. *National Teacher Education Journal*, 5(1), 83-98.
- American Association for Higher Education. (2018). Principles of good practice for assessing student learning. 9 principles of good practice for assessing student learning. Retrieved from https://www.researchgate.net/publication/268399482_American_Association_for_Higher_Education_AAHE_Principles_of_Good_Practice_for_Assessing_Student_Learning_9_Principles_of_Good_Practice_for_Assessing_Student_Learning.
- Black, P., & Wiliam, D. (2010). Inside the black box: Raising standards through classroom assessment. *Phi Delta Kappan*, 92(1), 81 – 90. doi:10.1177/ 003172171009200119.
- Butler, S. M., & McMunn, N. D. (2011). A teacher's guide to classroom assessment: Understanding and using assessment to improve student learning. John Wiley & Sons, Incorporated. Retrieved from ProQuest Ebook Central, <https://ebookcentral-proquest-com.ezproxy.liberty.edu/lib/liberty/detail.action?docID=708115>.
- Chan, P. E., Graham-Day, K. J., Ressa, V. A., Peters, M. T., & Konrad, M. (2014). Beyond involvement: Promoting student ownership of learning in classrooms. *Intervention in School and Clinic*, 50(2), 105-113. doi-org.ezproxy.liberty.edu/10.1177/1053451214536039
- Dixon, D. D., & Worrell, F. C. (2016) Formative and summative assessment in the classroom. *Theory Into Practice*, 55(2), 153-159, doi: 10.1080/00405841.2016.1148989
- Earl, L. (2013). *Assessment as learning: Using classroom assessment to maximize student learning*. (2nd ed.). Thousand Oaks, CA: Corwin Press.
- Espinosa, L. F. (2015). Effective use of performance-based assessments to identify English knowledge and skills of EFL students in Ecuador. *Theory and Practice in Language Studies*, 5(2), 2441+.
- Frey, B.B. (2014). Modern classroom assessment. In *Modern classroom assessment* (pp. 1-14). 55 City Road, London: SAGE Publications, Ltd doi: 10.4135/9781506374536.n1
- Frey, B. B. & Schmidt, V. L. (2007). Coming to terms with classroom assessment. *Journal of Advanced Academic*, 18, 402-423.
- Gareis, C. R., & Grant, L.W. (2015) *Teacher-made assessments: How to connect curriculum, instruction, and student learning*. Routledge, ProQuest Ebook Central. Retrieved from: <http://ebookcentral.proquest.com/lib/liberty/detail.action?docID=2001973>.
- Garrison, C., & Ehringhaus, M. (2013). Formative and summative assessments in the classroom. *Association for Middle Level Education*. Retrieved from <https://www.amle.org/BrowsebyTopic/WhatsNew/WNDet/TabId/270/ArtMID/888/ArticleID/286/Formative-and-Summative-Assessments-in-the->
- Kizlik, B. (2014). *Measurement, assessment and evaluation in education*. Retrieved from <http://www.pdfmode.com/ed/education-evaluation-pdf>

- Madaus, G., & O'Dwyer, L. (1999). A short history of performance assessment: Lessons learned. *The Phi Delta Kappan*, 80(9), 688-695. Retrieved from <http://www.jstor.org/stable/20439537>
- Mertler, C. A. (2003). *Classroom assessment: A practical guide for educators*. Los Angeles, CA: Pycrcal Publishing.
- Ratminingsih, N. M., Artini, L. P., & Padmadewi, N. N. (2017). Incorporating self and peer assessment in reflective teaching practices. *International Journal of Instruction*, 10(4), 165-184. <https://doi.org/10.12973/iji.2017.10410a>
- Regional Educational Laboratories (1998). *Improving classroom assessment: A toolkit for professional developers*. Portland, OR: Northwest Regional Educational Laboratory.
- Strauss, V. (2015, October 24). Confirmed: Standardized testing has taken over our schools. But who's to blame? *Washington Post*. Retrieved from https://www.washingtonpost.com/news/answer-sheet/wp/2015/10/24/confirmed-standardized-testing-has-taken-over-our-schools-but-whos-to-blame/?noredirect=on&utm_term=.fe728e65621a
- Taylor, C. S., & Nolen, S. B. (2005). *Classroom assessment: Supporting teaching and learning in real classrooms*. Columbus, OH: Pearson/Merrill Prentice Hall, p. 190.
- Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Cambridge, MA: Harvard University Press.
- Wang, C. S. (2017). Three paradigms of classroom assessment: Implications for written feedback research. *Language Assessment Quarterly*. DOI: 10.1080/15434303.2017.1405423