International Journal of Social Science and Business, Vol. 4 No 2; June 2019
ISSN: 2519-9161(Online), ISSN: 2519-9153(Print)
Published by Center for Academic Research (CAR)

Assessment for Learning: What Does it Entail?

Dr. Charles E. Notar (Emeritus)
Jacksonville State University
United States of America

Abstract

The article is focused on specific needs of new teachers (18). However, any teacher can use the article to refresh what they are doing mentally from their classroom experiences. In the age of accountability, it is important for educators to understand key principles of assessment and evaluation planning. This article reviews core knowledge of this planning as used within educational context. History, definition, and step to design an assessment/evaluation plan are provided. Charts and tables provide examples of assessment planning tools.

Key Words: Assessment Plan, Evaluation Plan, Developing a Test Plan, Writing Test Plans, Test Plan Format

Introduction

Education provides schooling to train by formal instruction and supervised practice especially in a skill, trade, or profession and to develop mentally, morally, or aesthetically especially by instruction. Formal education has two principal objectives with respect to the cognitive development of the individual: the long-term acquisition and retention of stable, organized, and extensive bodies of meaningful, generalizable knowledge and b) growth in the ability to use this knowledge in the solution of particular problems, including those problems which, when solved, augment the learner's original store of knowledge. These two goals require an assessment to be made to ensure accomplishment.

As a teacher the author is sure you know the feeling of anticipation when you are about to take a quiz or test. Did you take detailed class notes and study enough? Have you studied what will be tested? When the test comes are you shocked that what was in the book or in class is not being tested? (26).

Exams and essays along with speeches and projects are forms of assessment. Assessment is a critical step in the learning process. It determines whether or not the course's learning objectives have been met. A **learning objective** is what students should know or be able to do by the time a lesson is completed. Assessment affects many facets of education, including student grades, placement, and advancement as well as curriculum, instructional needs, and school funding (26).

A comprehensive teacher assessment and evaluation system should have two distinct components: 1) ongoing, consistent, formative assessments of performance for the sole purpose of fostering professional growth and improved practice; and 2) periodic summative evaluations of teacher performance for the purpose of approving continued employment (27).

Student learning outcome assessment begins with specifying the objectives ... Finally, plans are made for revisions that will improve student outcomes in the ... tests, oral and written reports, individual and group projects, portfolios, and other (27). Assessment plans detail how you will work through the steps of the assessment cycle for the learning outcomes you have identified for your course (19). In the context of classroom assessment, the term refers to the entire process of measurement, evaluation, and, finally, use of the information by teachers and students.

An assessment system will be defined as all the systematic methods and procedures that are used to obtain information about behaviors and upon which educational decisions are based to include:

The learning outcomes for departments or programs •

The assessment methods used to demonstrate the achievement of each outcome •

The timeframe for collecting and reviewing the data •

Assessment should be meaningful, manageable, measurable, and sustainable (15).

What is an assessment?

Assessment is used for a multitude of educational purposes: person's aptitude to learn something; motivation to achieve in school; self-concept; achievement level in a scholastic area; and environmental factors that affect how much a person learns.

Teachers, counselors and all others must have a baseline for assessment, measurement, and evaluation!

Assessment is a term that is used in a variety of ways. The shorter term, assess, is a synonym for measure. When researchers say they "assessed" something, they mean that they measured it.
Sometimes assessment means "evaluation," and sometimes it refers to the more specific process of diagnosing of individual difficulties, such as assessing for learning disabilities. Some measurement specialists use assessment to refer to procedures used to obtain information about student performance (23).

What's the definition of assessment in education? Assessment is the systematic process of documenting and using empirical data on the knowledge, skills, attitudes and beliefs. By taking the assessment, teachers try to improve student learning. This is a short definition of assessment (24).

Assessment is the process of documenting knowledge, skills, attitudes and beliefs, usually in measurable terms. The goal of assessment is to make improvements, as opposed to simply being judged. In an educational context, assessment is the process of describing, collecting, recording, scoring, and interpreting information about learning (45).

An assessment plan is a written document that describes. how you will monitor and evaluate your program, as well as how you intend to use evaluation results for program improvement and decision making. The evaluation plan clarifies how you will describe the "What," the "How," and the "Why It … (38).

T The primary purpose of assessment is to improve student achievement.
T Assessment continually guides the development, implementation and support of instruction.
T Students need to receive timely, specific and directive feedback in order to meet the grade/course expectations and assessment targets.
T Diagnostic assessment of content knowledge, skill level, use of literacy, numeracy and metacognitive strategies will serve as baseline data and inform instructional starting points.

Importance of Assessment

Hopefully by this point in your life you have discovered that learning can be fun! You have probably also realized that you are constantly learning, whether you are in a classroom, a car, or a kitchen (41).
I'm sure you know the feeling of anticipation when you are about to take a quiz or test. Did you take detailed class notes and study enough? And you surely have been assigned with various essays. Did you give yourself enough time to research, write, and revise your essay in order to meet the requirements? (26).

Exams and essays along with speeches and projects are forms of assessment. Assessment is a critical step in the learning process. It determines whether or not the course's learning objectives have been met. A **learning objective** is what students should know or be able to do by the time a lesson is completed. Assessment affects many facets of education, including student grades, placement, and advancement as well as curriculum, instructional needs, and school funding (26).

A comprehensive teacher assessment and evaluation system should have two distinct components: 1) ongoing, consistent, formative assessments of performance for the sole purpose of fostering professional growth and improved practice; and 2) periodic summative evaluations of teacher performance for the purpose of approving continued employment (27).

Student learning outcome assessment begins with specifying the objectives ... Finally, plans are made for revisions that will improve student outcomes in the ... tests, oral and written reports, individual and group projects, portfolios, and other (27).

An assessment plan details how you will work through the steps of the assessment cycle for the learning outcomes you have identified for your course (19).

3. Determine "learning opportunities" (i.e., where learning takes place). Curriculum mapping is the recommended strategy for doing this (19).

Once you have settled on expected student learning outcomes that reflect your current program, it is necessary to map where students actually have the opportunity to gain the knowledge and learn the skills necessary to meet those outcomes. A curriculum map is a matrix or table that demonstrates alignment of instruction with desired student learning outcomes [SLOs] (19).

Why do We Need an Assessment Plan?

(Patton, 2008).

Just as using a roadmap facilitates progress on a long journey, an evaluation plan can clarify what direction your evaluation should take based on priorities, resources, time, and skills needed to accomplish the evaluation. The process of developing an evaluation plan in cooperation with an evaluation workgroup of stakeholders will foster collaboration and a sense of shared purpose. Having a written evaluation plan will foster transparency and ensure that stakeholders are on the same page with regards to the purpose, use, and users of the evaluation results. Moreover, use of evaluation results is not something that can be hoped or wished for but must be planned, directed, and intentional A written plan is one of your most effective tools in your evaluation tool box (31).

Definitions of Key Terms Assessment, Test, Measurement, and Evaluation

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).

In the most general sense, assessment is the process of making a judgment or measurement of worth of an entity (e.g., person, process, or program). Educational assessment is a method for faculty to collect and analyze data evolving from planned learning activities or programs and its importance in the teaching/learning process. The purpose of **assess**ment is to provide faculty and students with information and insights needed to improve student learning, teaching strategies, and curriculum (17, 28).

Assessment is a way to measure if the learning objectives of a course are being met. Assessment is valuable for both students and teachers in evaluating progress (26).

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 (Akpan, Notar, & Padgett, 2012).

Tests are very systematic, formal procedures for gathering information about students' performance. They are a collection instrument. Tests are a formal set of questions or tasks, often administered to a group of students, that address particular cognitive capabilities learned in a specific course or subject area. A test or quiz is used to examine someone's knowledge of something to determine what he or she knows or has learned. Testing measures the level of skill or knowledge that has been reached (45).

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.

Measurement is concerned with quantification. Language proficiency, like many other constructs and characteristics of persons in social sciences, needs to be quantified before any judgments can be made about it. This process of quantifying is called operationalization in research by which we mean assigning numbers according to observable operations and explicit procedures or rules to measure a construct (Bachman 1990, Ary et al. 1996 as cited in 28).

Measures are specific techniques or instruments used for measurement and generally refer quantitative devices. These are often tests and questionnaires that provide objective, quantifiable data. For example, a specific reading test may be used to provide measurement of reading ability.

Measurement can be defined as the assignment of numbers to indicate different values of a variable; some researchers may also use it to refer to quantitative data collection.

Measurement is used to determine how much of a trait, attribute, or characteristic an individual possesses. Numbers are used to describe and differentiate attributes or characteristics of a person, object, or event.

Evaluation

Evaluation is the process of making judgments based on criteria and evidence (45) using assessment information to make judgments about students, teachers, or educational programs.

Bachman (1990 as cited in 28), quoting Weiss (1972 as cited in 28) defines evaluation as "the systematic gathering of information for the purpose of making decisions." Lynch (2001 as cited in 28) adds the fact that this decision or judgment is to be about individuals. In this conceptualization, both authors agree that evaluation is the superordinate term in relation to both measurement and testing. Assessment is sometimes used interchangeably for evaluation. The systematic information can take many forms, but these forms are either quantitative or qualitative.

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 as cited in 28) 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 as cited in 28). Each method is part of an *assessment system* and should be used comprehensively with other samples of student work both formally and informally.

Additionally, evaluation often requires a substantial degree of professional decision making by classroom teachers. Since this type of decision making has the potential for very important repercussions, it should occur only after adequate samples of assessment information have been collected, analyzed, and synthesized. Only then can teachers make truly informed decisions and judgments.

What Is Evaluation? Evaluation is also used to compare performance with an objective or standard. In this sense student achievement may be evaluated by comparison to national norms or locally set standards of achievement (e.g., the dropout rate will be less than 4%; the number of high school students studying foreign languages will rise to 40%; the mean level of achievement of sixth graders will be above the national mean). Other definitions of evaluation focus on professional judgment or a process in which a judgment is made about something. Such judgments may or may not involve measurement. The distinguishing aspect of evaluation is that data are interpreted and some kind of value is placed on the results.

Summary of Definitions

Assessment System - all the systematic methods and procedures that are used to obtain information about behaviors and upon which educational decisions are based [Planning of measurement and evaluation]

Measurement - a process involving a structured situation that includes samples of particular characteristics or behaviors that results in a numerical or narrative score [Collection of data]

Test - a formal set of questions or tasks, often administered to a group of students, that address particular cognitive capabilities learned in a specific course or subject area [Instrument of measurement]

Evaluation - the use of assessment information to make judgments about students, teachers, or educational programs [Interpretation] (Mertler, 2003).

Assessment vs Evaluation: What's the Difference?

Assessment and evaluation are not the same. But what are the differences between assessment and evaluation in education (24). Depending on the authority or dictionary consulted, assessment and evaluation may be treated as synonyms or as distinctly different concepts. As noted above, if a distinction exists, it probably involves what is being measured and why and how the measurements are made. In terms of what, it is often said that we assess students and we evaluate instruction. This distinction derives from the use of evaluation research methods to make judgments about the worth of educational activities. Moreover, it emphasizes an individual focus of assessment, i.e., using information to help identify a learner's needs and document his or her progress toward meeting goals (34).

Evaluation of teaching means passing judgment on it as part of an administrative process. Ideally, a fair and comprehensive plan to evaluate teaching would incorporate many data points drawn from a broad array of teaching dimensions (35). The significant differences between assessment and evaluation are discussed in the points given below: The process of collecting, reviewing and using data, for the purpose of improvement in the current performance, is called assessment. A process of passing judgment, on the basis of defined criteria and evidence is called evaluation (36).

Evaluations, are a set of statements or activities that seek to point out whether the objectives were met. It is the last part of the inquiry, and involves telling the people whether the solution or the goal has been realized or not (39).

Figure 1 provides a chart to summarize assessment and evaluation.

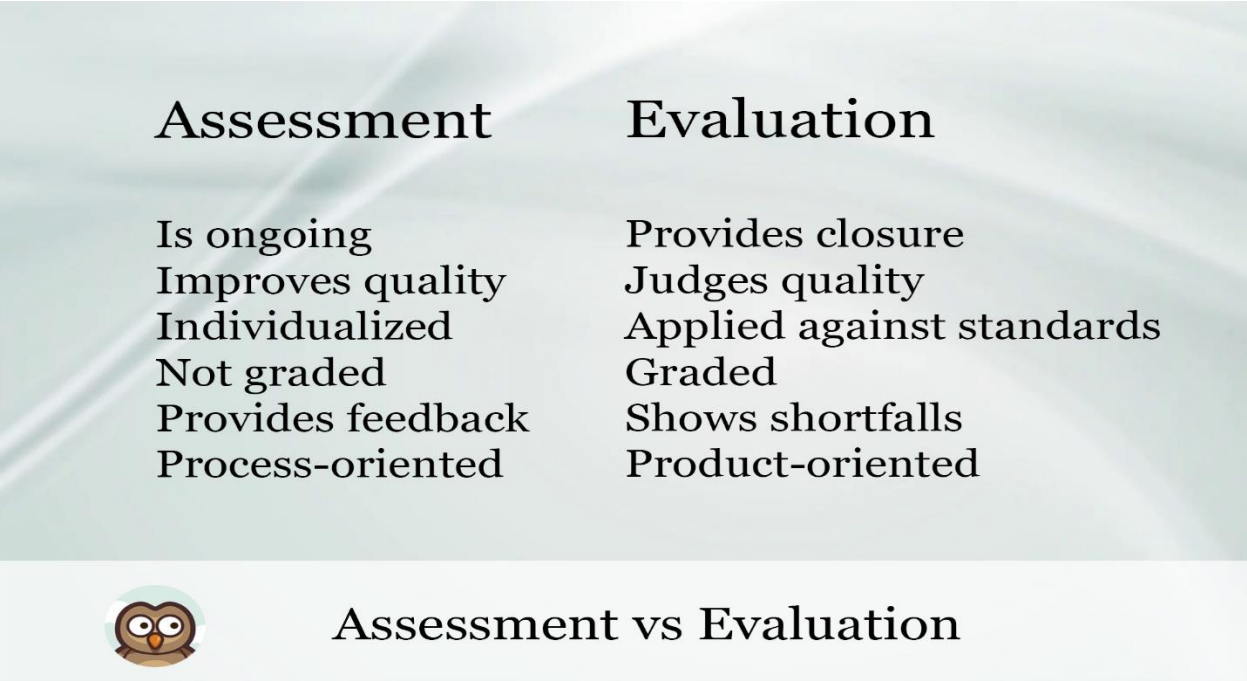


Figure 1: Summary Assessment and Evaluation (19, 24).

Relationship between Assessment and Evaluation

Besides the differences, there are also some similarities (Figure 2) between assessment and evaluation. They both require criteria, use measures and are evidence-driven.

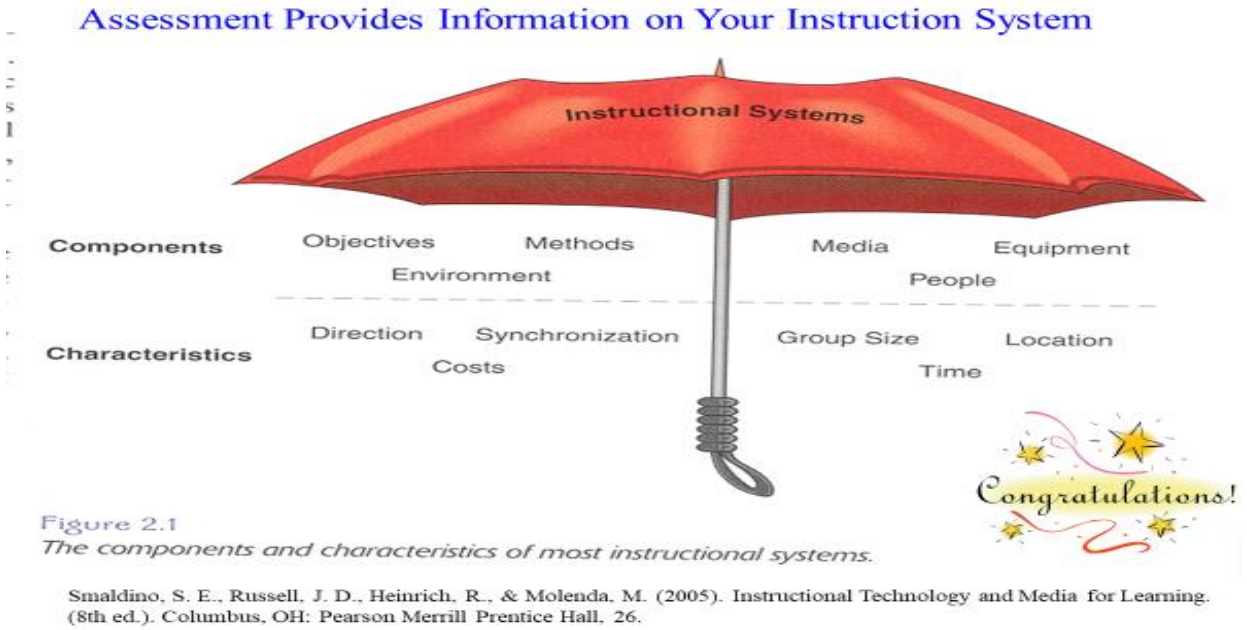


Figure 2: Assement Provides Information on Your Instruction System

When defined within an educational setting, assessment, evaluation, and testing are all used to measure how much of the assigned materials students are mastering, how well students are learning the materials, and how well students are meeting the stated goals and objectives. Although you may believe that assessments only provide instructors with information on which to base a score or grade, assessments also help you to assess your own learning (45).

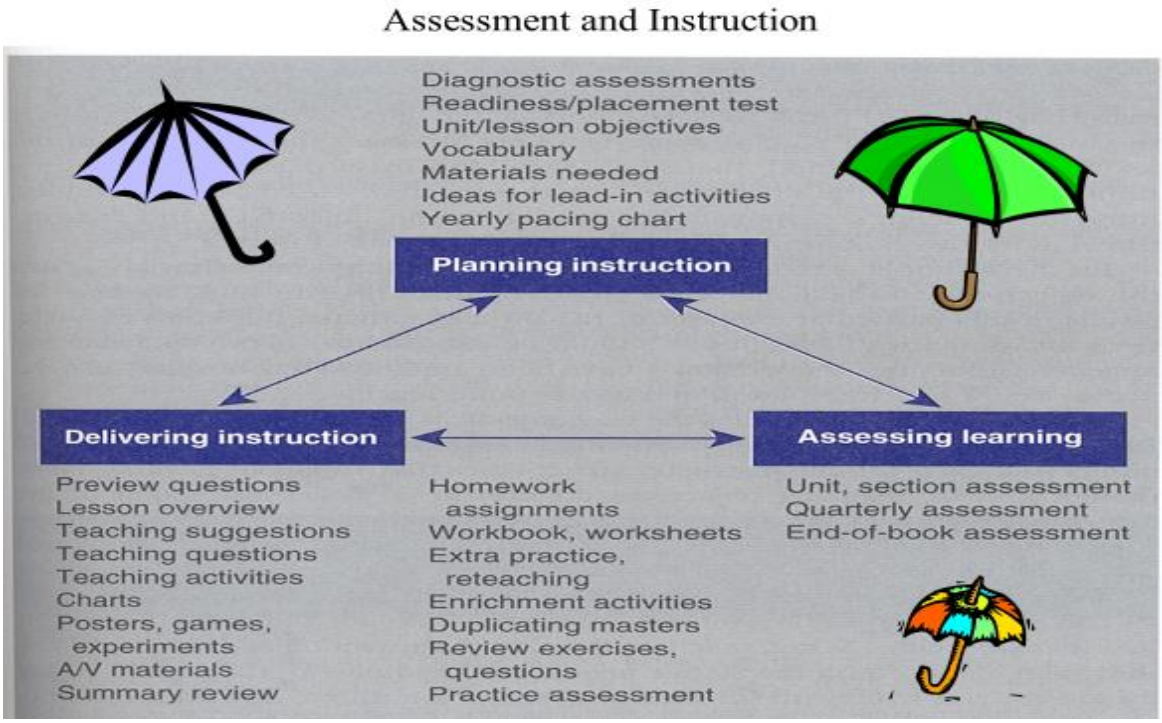


Figure 3: Assessment and Instruction

While figure 3 provides a model of how assessment and instruction are intertwined and figure 4 explains the how one affects the other (47, 48).

Relation Between Instruction and Assessment

Instruction **Instruction is most effective when**	**Assessment** **Assessment is most effective when**
1. Directed toward a clearly defined set of intended learning outcomes.	1. Designed to assess a clearly defined set of intended learning outcomes.
2. The methods and materials of instruction are congruent with the outcomes to be achieved.	2. The nature and function of the assessments are congruent with the outcomes to be assessed.
3. The instruction is designed to fit the characteristics and needs of the students.	3. The assessments are designed to fit the relevant student characteristics and are fair to everyone.
4. Instructional decisions are based on information that is meaningful, dependable, and relevant.	4. Assessments provide information that is meaningful, dependable, and relevant.
5. Students are periodically informed concerning their learning progress.	5. Provision is made for giving the students early feedback of assessment results.
6. Remediation is provided for students not achieving the intended learning.	6. Specific learning weaknesses are revealed by the assessment results.
7. Instructional effectiveness is periodically reviewed and the intended learning outcomes and instruction modified as needed	7. Assessment results provide information useful for evaluating the appropriateness of the objectives, the methods, and the materials of instruction.

Figure 4: Relation Between Instruction and Assessment (46)

Before getting into the who, what, why, when and how of the assessment plan a three basics behind an assessment plan need explanation.

With an understanding of assessment and evaluation a teacher's effective student assessment plan needs to be developed. The following are the foundation on which that assessment plan needs to be designed:

1. Must have a clear conception of all intended learning outcomes.
2. A variety of assessment procedures must be used.
3. Assess only instructionally relevant outcomes.
4. An adequate sample of student performance is required.
5. Assessment must be fair. No bias, discrimination, etc.
6. Assessment requires criteria for judging successful performance.
7. Feedback is required on strengths and weaknesses.
8. Assessment must support a comprehensive grading and reporting system.

As you can see education professionals make distinctions between assessment, tests, measurement and evaluation. However, for the purposes of this paper, all you really need to understand is that there are four different terms for referring to the process of figuring out how much you know about a given topic and that each term has a different meaning. To simplify things, the author will use the term "assessment" throughout this paper to refer to this process of measuring what you know and have learned (45).

Introduction to an Actual Example of Assessment Plan

The first section provides background information on the school, its history, its educational philosophy, its distinctive features, its competence-based approach to instruction. The second section seeks to help the instructors "get started." (20).

Assessment in education has become increasingly important as the nation debates educational effectiveness at all levels. Political and public pressure requires schools to be held accountable for both the resources spent in education and the educational outcomes of those resources (21).

In summary, grades do not provide the following:

Specific information about students' performance on discrete tasks

Meaningful data across sections

Objective student data which can be used for improvement of student learning or recognition of student achievement (21).

CAUTION It is critical, however, that students do not approach outcome assessment assignments or exam questions thinking they are of no consequence, as they would likely not take them seriously thus creating a false impression regarding the effectiveness of our courses. Regardless of how instructors grade the instruments, they should communicate to students the value of the outcomes and the instruments used to access them (21).

Assessment is the process of gathering evidence of student learning, reviewing the evidence to determine if students are learning what they are expected to learn, and using this evidence to alter the direction of your course (21).

Every course should have a set of school-wide, common, core expectations for student learning. These expectations are the most important things a student who passes the course should take away from any section of the course. While individual instructors may add to this course, there should be a shared understanding of the core skills and knowledge upon which the course is based. It is these expectations which should be reflected on each course syllabus and which should be used to determine student learning outcomes for the outcome assessment process. Student learning outcomes are statements that specify what you want your students to know and be able to do at the end of the course. For example, student learning outcomes can refer to knowledge, practical skills, critical thinking skills, etc. that students are expected to develop or learn (21).

What is Included in an Assessment Plan?

There are no set rules for how to construct an assessment plan for your department/program (1).
An assessment plan should detail how you will work through the steps of the assessment cycle for the learning outcomes you have identified for your course (19).

A comprehensive teacher assessment and evaluation system should have two distinct components: 1) ongoing, consistent, formative assessments of performance for the sole purpose of fostering professional growth and improved practice; and 2) periodic summative evaluations of teacher performance for the purpose of approving continued employment (27).

Don't forget that assessments are a means, not an end in themselves. That is, don't forget their *purpose*. Student learning outcome assessment begins with specifying the objectives ... Finally, plans are made for revisions that will improve student outcomes in the ... tests, oral and written reports, individual and group projects, portfolios, and other (27). Models of learning have been based on an accumulation of facts until just recently. Currently, problem-solving, decision-making and critical thinking are what we want to promote in the classroom.

Here are some simple guidelines to help you develop an assessment plan (1).

1) *Identify your overall goals for the plan (1).* Decide which outcomes are to be assessed each year and set clear learning objectives.
2) *Identify specific, measurable objectives to be achieved for each goal (1).* Assessments should be appropriate (relevant) for those objectives. Identify, select, and/or develop measures for program assessment (19) and what evidence/samples of student work will be collected.
3) *Identify the activities that you will carry out to achieve each of the objectives (1).* A variety of objectives generally requires a variety of assessment techniques. Know the limitations of each form of assessment (e.g., sampling error, measurement error) so assessments are appropriate (relevant) for those objectives.
4) *For each activity create at least one measure to assess how well it was implemented, e.g., the number of participants who attended a workshop, a survey to measure how satisfied the participants were with the workshop (1).*
5) For each activity, create at least one measure to assess the impact of the activity and determine whether the goal was achieved (1).
6) Create a timeline for evaluation plan (1).
7) Multiple methods of assessment, including direct and indirect
   a. Indirect: Gathering information from students on their thoughts, attitudes, perceptions, in relation to the course/program
   b. Direct: Examining work produced by students to determine whether it meets the course/program outcomes (1).

In addition to the method to be used you determine who will collect student work, analyze data, and summarize results/do reporting (1). This includes when and where student work will be collected and when each outcome is to be assessed. Location in instructional process is key to a valuable assessment.

The following author assessment plan provides the elements that should be in a successful assessment plan.

Components of Assessment Plan

A. Picture of School

1. School

a. Name of School System:

b. School system structure:

Superintendent

Asst. Superintendent(s)

Consultants

c. School system

Mission statement

Philosophy

Aims

Goals

Purpose

d. Name of Your School:

e. Your School Administration

Principal

Asst Principal(s)

Depart chairs, etc.

f. Your School

Mission statement

Philosophy

Aims

Goals

Purpose

g. Size: **(Number of students)**

h. Structure: (Grade structure)

i. Locations: (Demographics—rural, urban, median income, etc. (solid paragraph))

j. Describe the anticipated learner (These are your students that will be doing the unit being developed for the teacher made test (e.g. how many, sex, reading grade level, socio-economics, free lunch, diversity)

2. Class Materials/Classroom Description

2.1 Class Materials

a. When are textbooks selected?

b. Process of textbook selection?

c. Who are selectors at various grades

d. Is the curricula coordinated, math, science, English, social studies or divergent? (coordinated is within discipline and divergent is between discipline)

2.2 Classroom Description

Describe for the anticipated learner the following: (Class periods are 55 minutes long)

a. Objectives/Competencies

b. Content (Subject Matter)

c. Method(s)

d. Text(s)

e. Learning Activities

f. Physical Layout(s) of classroom

g. Class Schedule

3. Assessment Requirements

a. Reporting system

b. Grading system

c. State and local test requirements (e.g. SAT 10, graduation exam)

B. Assessment Plan

1. Annual assessment plan and subset unit assessment plan

2. Unit objectives

3. Unit plan with 7-8 types assessment (formative and summative activities)

4. Grading system

C. Teacher materials

1. "Testwiseness" program

2. How to present assessment to parents

3. Handout to parents on standardized test scores and their meaning

D. Unit test

1. Test/Answer key

2. Table of specifications

3. Test admin manual

4. Remediation plan

5. Rubric for one of the summative activities for unit

6. Accommodations (Notar, 2008).

As you can see there are multiple parts to an assessment plan. The following information will help you in developing the parts (3-14). Figure 5 provides the mind set you as a teacher should have as you develop your assessment plan.

Key Questions for Evaluating Learning and Instruction

	Evaluating Learning	Evaluating Instruction
Before Instruction	Do students have the prerequisite knowledge and skills? Do students already know the content they are slated to learn? What is the student's current level of performance (baseline)?	How well is the instruction likely to work? Will the instruction hold student interest? Is there an alternative way to organize the instruction to make better use of available time and resources?
During Instruction	Are students ready for new content or is additional practice and feedback needed? In what specific areas do students need additional practice and feedback? What types of remediation or enrichment activities may be necessary for students?	What obstacles are students encountering and how can they be overcome? What can be done to maintain student motivation? How can these students be helped to better progress through the instruction?
After Instruction	Have students learned what was intended? Can students be accredited or "passed"? What will be needed to help students generalize what they have learned and transfer it to new situations?	What improvements could be made in the instruction for future use? What revisions have the highest priority? Did students find the instruction interesting, valuable, and meaningful? Were the selected instructional methods, media, and materials effective in helping students learn?

Newby, T. J., Stepich, D. A., Lehman, J. D., & Russell, J. D. (2000). *Instructional Technology for Teaching and Learning: Designing Instruction, Integrating Computers, and Using Media.* (2nd ed.). Columbus, OH: Merrill, 223.

Figure 5: Key questions for evaluating learning and instruction

The key questions and your assessment plan start with objectives. *All learning in the classroom should be based on objectives. Hence, an assessment plan starts with the development of the objectives that will be assessed.*

Objectives

All learning in the classroom should be based on objectives. Hence, an assessment plan starts with the development of the objectives that will be assessed. Once you have settled on expected student learning outcomes that reflect your current program, it is necessary to map where students actually have the opportunity to gain the knowledge and learn the skills necessary to meet those outcomes. A curriculum map is a matrix or table that demonstrates alignment of instruction with desired student learning outcomes [SLOs] (19, 49).

Determine "learning opportunities" (i.e., where learning takes place). Curriculum mapping is the recommended strategy for doing this (19).

Objectives are for students and measuring performance in real world. The action, condition(s), and standard(s) must measure actual performance, not test performance.

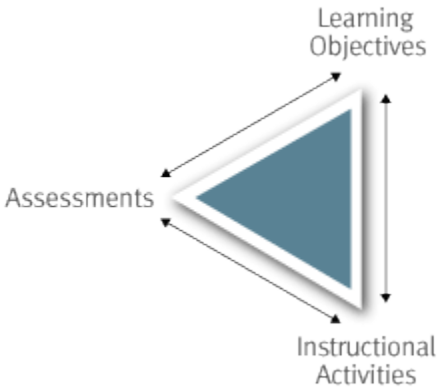
All evaluations are designed to assess whether the objectives of the project have been met.

As the objectives of each project are unique, there is no "pre-packaged" evaluation plan that you can use to evaluate a program. Every evaluation plan is unique and needs to be designed to measure the specific objectives of the project (1).

For the purpose of schematic representation, the three concepts of evaluation, measurement and testing have traditionally been demonstrated in three concentric circles of varying sizes. This is what Lynch (2001 as cited in 28) has followed in depicting the relationship among these concepts.

Assessments should reveal how well students have learned what we want them to learn while instruction ensures that they learn it. For this to occur, assessments, learning objectives, and instructional strategies need to be closely aligned so that they reinforce one another.
To ensure that these three components of your course are aligned, ask yourself the following questions:

- **Learning objectives:** What do I want students to know how to do when they leave this course?
- **Assessments:** What kinds of tasks will reveal whether students have achieved the learning objectives I have identified?
- **Instructional strategies:** What kinds of activities in and out of class will reinforce my learning objectives and prepare students for assessments? (22).



An objective is a simple, clearly written means of describing an observable and measurable student behavior. Objectives are for students and measuring performance in real world. The action, condition(s), and standard(s) must measure actual performance, not test performance.

Therefore, teachers **MUST** show them to the students so that they know what they are to learn and to what degree.

Teachers use objectives to help plan their instruction, give them guidance during presentation, and insures that they have the same goals as the students. They provide a sound basis for selection or design of instructional materials, content or methods. They serve as guidelines for test questions. They provide students with means to organize their efforts toward accomplishment of objectives. They provide instructor with guidelines in developing lesson plans. They control the instructional process. They help in preparing the means of finding out whether instruction has been successful. Select best media. They point out the content and procedures that lead to successful instruction. Select best teaching strategies. Develop ASSESSMENT PLAN.

Table 1 presents examples of the kinds of activities that can be used to assess different types of learning objectives (adapted from the revised Bloom's Taxonomy, 1956).

Type of learning objective	Examples of appropriate assessments
Recall **Recognize** **Identify**	Objective test items such as fill-in-the-blank, matching, labeling, or multiple-choice questions that require students to: o recall or recognize terms, facts, and concepts
Interpret **Exemplify** **Classify** **Summarize** **Infer** **Compare** **Explain**	Activities such as papers, exams, problem sets, class discussions, or concept maps that require students to: o summarize readings, films, or speeches o compare and contrast two or more theories, events, or processes o classify or categorize cases, elements, or events using established criteria o paraphrase documents or speeches o find or identify examples or illustrations of a concept or principle
Apply **Execute** **Implement**	Activities such as problem sets, performances, labs, prototyping, or simulations that require students to: o use procedures to solve or complete familiar or unfamiliar tasks o determine which procedure(s) are most appropriate for a given task
Analyze **Differentiate** **Organize** **Attribute**	Activities such as case studies, critiques, labs, papers, projects, debates, or concept maps that require students to: o discriminate or select relevant and irrelevant parts o determine how elements function together o determine bias, values, or underlying intent in presented material
Evaluate **Check** **Critique** **Assess**	Activities such as journals, diaries, critiques, problem sets, product reviews, or studies that require students to: o test, monitor, judge, or critique readings, performances, or products against established criteria or standards
Create **Generate** **Plan** **Produce** **Design**	Activities such as research projects, musical compositions, performances, essays, business plans, website designs, or set designs that require students to: o make, build, design or generate something new

Table 1: Examples of appropriate assessments 19, 54

Table 1 does not list all possible examples of appropriate assessments. You can develop and use other assessments – just make sure that they align with your learning objectives and instructional strategies! (22). Well developed objectives will be the foundation for dealing with the many questions that will be asked of you as a teacher (figure 6). This information will be also be used when you create your unit plan as seen in Table 8.

Forces Influencing Teacher's Teaching

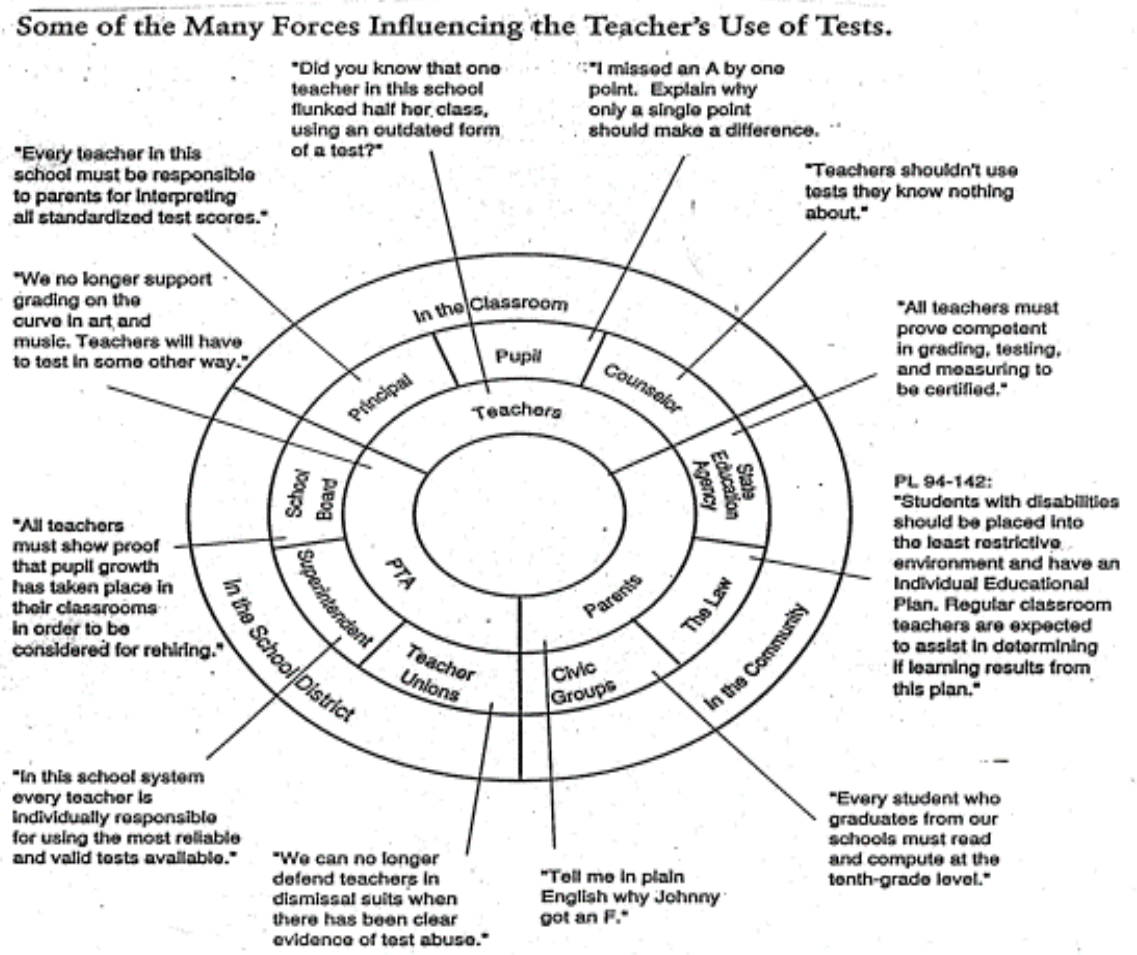


Figure 6: Forces Influencing the Teacher's Teaching

Develop a plan for ongoing learning outcomes assessment ... assessment or instruction process, whereas outcomes are the results of tests or assignments (29).

In education, at its simplest, outcomes assessment has three stages: 1. Defining the most important goals for students to achieve as a result of participating in an academic experience (outcomes) 2. Evaluating how well students are actually achieving those goals (assessment) 3. Using the results to improve the academic experience (closing the loop) (21).

Effective instructional programs have certain common characteristics. First and foremost, effective instructional programs are focused on student learning and employ the best practices in curricular design and instructional modalities. Effective programs require continual assessment of student learning and provide ready feed-back to students about their learning. Effective programs create a learning environment that provides access to learning for all students (21).

Testing

In the assessment plan you develop you come to the part on testing. There have been problems with test construction and grading as cited in the news. There is also a problem where teachers' are teaching the test. In the author's opinion this is caused by the way teachers are being evaluated.

Testwiseness

Before we get into testing a comment needs to be made about teachers teaching the test. Teaching the test is a definite NO-NO. However, there is the issue of testwiseness.

The ability to correctly answer items is often called testwiseness. Testwiseness is the ability to use assessment-taking strategies, clues from poorly written items, and experience in taking assessments to improve your score beyond what you would otherwise attain from mastery of the subject matter itself. When you write classroom assessments, be aware of how students may take advantage of your idiosyncrasies in item writing or flawed items to improve their scores without attaining the desired level of mastery. (Nitko, 2004).

Teachers should create good-quality assessments that minimize any advantage that testwise students have. Research has demonstrated that testwiseness is learned, and it improves with grade level, experience in being assessed, mauration, and motivation to do well on the assessment (Samacki, 1979; Slakter, Koehler, & Hampton, 1970 as cited in Nitko, 2004).

A teacher can ethically teach testwiseness skills, checking answer documents to make sure that, each has been properly completed and motivate parents, students, and teachers for good test results (Nitko, 2004).
Testwiseness is defined as the set of cognitive strategies used by a student that is intended to improve his or her score on a test regardless of the test's subject matter (53).

Testwiseness is any skill which allows a student to choose the correct answer on an item without knowing the correct answer. Students who are testwise look for mistakes in test construction, make guesses based on teacher tendencies, and search for any unintentional clues that can be found in a test. This is an issue of validity because the score on a test should be a reflection of the level of the trait that the test is designed to measure (knowledge, skill, understanding) not a reflection of a general ability to do well on poorly made tests ((Haladyna, & Downing, 1989a; Haladyna, & Downing, 1989b; Haladyna, Downing, & Rodriguez, 2002; Frey, Petersen, Edwards, Pedrotti, & Peyton, 2003).

However, there is an unethical side to testwiseness. Developing a curriculum based on the content of the test and preparing objectives based on items on the test and teaching accordingly are the major no-no's. Presenting items similar to those on the test, using score-boosting activities and presenting items verbatim from the test to be given are also unethical. Dismissing low achievement students on testing day to artificially boost test scores on standardized tests that measure a school or school system is the epidemy of unethical (Nitko, 2004).

Regardless of whether you call it problem solving, thinking skills, test-taking skills, or call it test preparation, but do it every day as part of your regular instruction. Start by making sure your students are learning the curriculum. Expose them to all types of item formats, and be creative to get your students thinking (Nitko, 2004).

One way to reduce the influence of test-taking skills on test performance is to familiarize your students with the general requirements for taking tests prior to the time the test is administered.

Explain what the student should do prior to a test once it has been announced.

1. When a test is announced well in advance, do not wait until the day before to begin studying. Spaced practice is more effective than massed practice.
2. Ask the instructor for old copies of the examination to practice with.
3. Ask other students' what kinds of tests the instructor usually gives.
4. Don't turn study session into social occasion, isolated studying is usually more effective.
5. Don't be too comfortable when studying. Lying down is a physical cue for your body to sleep.
6. Study for the type of test which was announced.
7. If you do not know the type (style) of test, study for a free recall exam.
8. Use SurveyQ3R technique when studying. Survey material, ask yourself questions about the subject material, read for detail, recite the material to yourself, and review material just prior to test.

9. Try to form material you are studying into test questions.
10. Read test directions carefully before beginning exam. Ask administrator if unclear or some details are not included.
11. If essay test, think about question and mentally formulate answer before you begin writing.
12. Pace yourself while taking test. Do not try to be first person finished. Allow enough time to review answers at end of session.
13. If you can rule out one wrong answer choice, guess, even if there is a penalty for wrong answers.
14. Skip more difficult items and return to them later, particularly if there are a lot of questions. It is better to leave two difficult questions unanswered than 10 to 15 easy questions.
15. When time permits, review your answer. Don't be overly eager to hand in your test paper before all the available time has elapsed (50-53).

In classroom or school settings where locally developed tests are administered, teachers should make sure that students have adequate test-taking skills so that their performance on the test reflects what they know about the subject being tested, instead of what they know about how to take a test (50, 51, Kubiszyn, & Borich, 2007).

1. Listen to or read directions carefully.
2. Follow directions carefully in marking the answer sheet (e.g. darken the entire space).
3. Read test items, passages and related information carefully.
4. Work quickly and set a pace that will allow time to complete the test.
5. Manage test-taking time; Attempt easier items first; Answer high-point questions first; bypass time-consuming difficult test items and return to for later review and changing of answers if time permits.
6. Make informed guesses rather that omitting items; eliminate as many incorrect alternatives as possible on multiple-choice items before guessing.
7. Check the item number and the answer match when marking an answer sheet. Proof work.
8. Mark items for later review and changing of answers if time permits
9. Focus on the task, not your feelings about it, focus on the facts; address specific material posed by short-answer and essay questions (50, 51, Kubiszyn, & Borich, 2007).

Also, while these items are not reflected in the scoring of tests, they improve a students' ability to take a test. As a teacher and you are grading a test you should make observations in the following areas and provide feedback to the student. These areas are:

Answering the Question	Using appropriate vocabulary
Sentence structure and language mechanics	Organizational skills
Justification for choices/decisions	Quality of response
Proof the text was read	Following directions

A teacher who has taught their students to be good test takers will have:

1. Demonstrated a positive attitude toward the assessment, regardless of how they may actually feel.
2. Created a partnership with students to demystify tests for them.
3. In class, explained as often as possible why incorrect answers are incorrect.
4. Given practice tests so students can practice with timed activities -- teaching them how to pace themselves.
5. Modeled problem-solving strategies with students.
6. Talked through how to determine what a question is asking.
7. Encouraged explanations of how students choose answers or approach problems.
8. Talk about strategies used to eliminate incorrect answer choices.
9. Established a suitable testing environment
10. Motivated students to do their best
11. Explained why tests are given and how results will be used
12. Told students they probably won't know all the answers
13. Told students not to give up (50, 51, Kubiszyn, & Borich, 2007).

A student with poor test-taking skills will:

Read too quickly
Jump to conclusions
Be confused by different answer formats
Have difficulty adjusting to the structured setting of a test
Miss important words
Make random guesses
See tests as an extra burden
Not understand the concept of "best" answer (50, 51 Kubiszyn, & Borich, 2007).

A student who has good test-taking skills will have learned to:

Underline important words (where permissible)
Analyze items systematically
Make sure they understand all directions
Skip difficult items and return to them if time permits
Eliminating irrelevant information.
Pace themselves during the test
Guess (after eliminating unreasonable options)

Can use number sense, estimation, backwards thinking and other skills useful in test-taking. (50, 51, Kubiszyn, & Borich, 2007).

Advice About Changing Answers

Will students benefit if they change their answers once they have been marked on the answer sheet? Wisdom says most often your first hunch is the right one, and changing answers usually lowers scores. Despite popular opinion, it does pay to change answers if changing them is based on a ***thoughtful reconsideration*** of the item. A summary of the research findings (Pike, 1978 as cited in Nitko, 2004) on this issue follows:

* Most test takers and many educators believe it does not pay to change answers.
* Most students, however, do in fact change their answers to about 4% of the items.
* Research studies show that it does, in fact, pay to change answers. Typically, two out of three answers changed will become correct.
* The payoff for changing answers diminishes as the items become more difficult for the student.
* Lower scoring students benefit less from changing answers than higher scoring students do.

However, psychological research has shown that examinees make higher scores when they reconsider their answers and change those which they have doubts associated with the correctness of their choice. Answers were more likely to be changed from wrong to right than the reverse case (Benjiman et. al. 1984; Geiger, 2010).

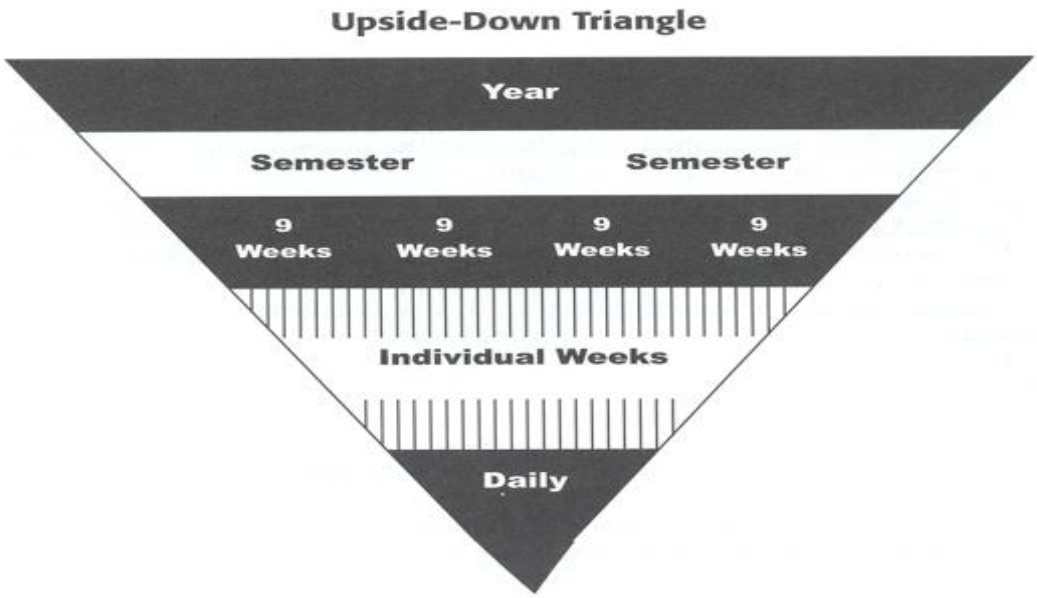
Examinees must be informed whether or not guessing can penalize you (points subtracted for wrong answers). Guessing can inflate scores more on all true-false tests than typical multiple choice format. Guessing usually results in higher scores when examinees can eliminate at least one false answer from the choices before guessing.

A comment on guessing, means random guessing, essentially flipping a coin and choosing an answer. Scores from a student who got lucky and guessed his or her way to a high score are meaningless and not valid. It is important to distinguish between this sort of a guess, which good tests are designed to protect against, and an "educated guess" which is not nearly as harmful to the validity of a test. With educated guesses, students, at least, have some knowledge of the content which has allowed them to narrow their answer options down to a small number of reasonable alternatives. The guidelines on this page are designed to protect against the lucky guess, not the educated guess (Haladyna, & Downing, 1989a; Haladyna, & Downing, 1989b; Haladyna, Downing, & Rodriguez, 2002; Frey, Petersen, Edwards, Pedrotti, & Peyton, 2003).

Guessing can inflate scores more on all straight true-false tests than typical multiple choice format (Benjiman et. al. 1984; Geiger, 1997). The best way to counter that is to use true-false correct questions.

Test Planning and Table of Specifications

An assessment plan is an upside-down triangle as shown in figure 7. Effective program delivery occurs when diagnostic, formative and summative assessment data are correctly interpreted and used to make decisions about initial instruction, intervention, future instructional strategies, evaluation and communication (18).



Blackburn, B. R.(2005). *Classroom Motivation From A to Z: How to Engage Your Students in Learning*. Larchmont, NY: Eye on Learning. 12.

Figure 7: Upside-Down Triangle

Numerous terms are used to describe different types and approaches to learner assessment. Although somewhat arbitrary, it is useful to these various terms as representing dichotomous poles (33). Table 2 and Table 3 are two examples of the different types and approaches.

Table 2: Types and approaches learner assessment

Diagnostic (gathered before new learning)
Formative (gathered throughout the gradual release of responsibility instructional model)
Summative (gathered after significant independent practice and cumulative demonstration of student achievement) (18).

Table 3: Types and approaches learner assessment

Formative <----------> Summative

Informal <----------> Formal

Continuous <----------> Final

Process <----------> Product

Divergent <----------> Convergent

Table 2 or 3 both mention formative assessment. However, formative assessment as a Sizing-Up Assessment are discussed. Sizing up is done at the start of the school year. Most teachers can describe the personal, social, and academic characteristics of each pupil and the class as a whole after the first two weeks of school. Sizing up is pupil-centered. Much of the information about pupils and their characteristics are the focus of assessment. Informal observation is used. Much of the information about pupil behavior and performance is collected through spontaneous, informal observations. Observations are synthesized into perceptions. Teachers put together their observations in idiosyncratic ways to form generalized perceptions pupils.

Impressions are rarely written down. Unlike test scores or grades are written down in rank books or report cards, the perceptions formed from sizing-up assessments are unwritten and selectively communicated. Observations are broad and diverse. Teachers attend to broad range of cognitive, affective, and psychomotor characteristics when they size up their pupils. Early impressions tend to become permanent. Teachers are very confident about the accuracy of the sizing-up assessments they perform in the first days of school. Initial perceptions are very stable from the first week of school to the end of the school year.

Definition of Summative and Formative Evaluation

Assessment ***of*** Learning = Summative Evaluation

Assessment ***for*** Learning = Formative Evaluation

Assessment plans are developed to provide baseline data by educators to communicate performance and are an integral part of effective instruction 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. 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).

Summative Evaluations

Summative assessments or assessment of learning provides students with multiple and varied opportunities to demonstrate the full extent of their learning (18).

As teachers, administrators, and counselors you may examine tests on the basis of format, intent, content, and condition. Tests which are given at the end of large blocks of instruction are generally summative in nature. They are used to evaluate student performance and usually occur infrequently. These tests have the disadvantage of identifying problems when it's too late to resolve them. Objectives which were not mastered may be identified without sufficient time left in the course to re-learn them. Also, if the instruction is cumulative all of the following modules may be failed because of failure of a previous module. Summative tests may be criterion tests, but when used in this fashion their ability is minimized.

Summative evaluations are culminating assessments conducted at the end of a unit, course, or grade level to determine the degree of mastery or proficiency according to identified achievement targets. They are used to determine the success of a section of instruction. Final tests for a grading period, assessments at the end of a unit (tests, term reports, summaries of portfolios), and end of the year assessments are all typical sources of summative data. They rely on a broad sampling of the relevant content; they focus more generally on all the objectives of the unit of instruction; and they are often a major part of the data collected for grading (Notar, 2008).

"Summative" assessment is what we normally call "evaluation." Evaluation is the process of observing and measuring a thing for the purpose of judging it and of determining its "value," either by comparison to similar things, or to a standard (25).

Formative Test

A better use of the criterion test is for formative evaluation. Formative evaluation is used to identify learning weaknesses prior to completion of a lesson, module, unit. Therefore, formative evaluation occurs frequently during instruction with the aim of testing enabling objectives before testing the terminal learning objectives. Formative evaluation is a major learning activity in an instructional system. The use of practice exercises and self-tests with student feedback sheets are examples of formative evaluation. In addition to being used as a corrective feedback loop to reinforce the appropriate response it also serves to extinguish incorrect responses and becomes a developmental tool.

Formative evaluations are procedures carried out as the instruction progresses. Formative assessments are typically short, focused exercises. They are aimed at skill development at a given point in a specific unit of instruction to verify that students are keeping pace with the concept and skills presented by instruction. Formative assessments of all kinds inform both teacher and student by providing periodic feedback.

They indicate if the class or an individual student has missed a point(s) in the concept under study. They advise instructors how to pace teaching and learning activities and indicate where difficulties may be occurring.

A better use of the criterion test is for formative evaluation. Formative evaluation is used to identify learning weaknesses prior to completion of a lesson, module, unit. Therefore, formative evaluation occurs frequently during instruction with the aim of testing enabling objectives before testing the terminal learning objectives. Formative evaluation is a major learning activity in an instructional system. The use of practice exercises and self-tests with student feedback sheets are examples of formative evaluation. In addition to being used as a corrective feedback loop to reinforce the appropriate response it also serves to extinguish incorrect responses and becomes a developmental tool (Akpan, Notar, & Padgett, 2012).

Formative evaluations are procedures carried out as the instruction progresses. Formative assessments are typically short, focused exercises. They are aimed at skill development at a given point in a specific unit of instruction to verify that students are keeping pace with the concept and skills presented by instruction. Formative assessments of all kinds inform both teacher and student by providing periodic feedback. They indicate if the class or an individual student has missed a point(s) in the concept under study. They advise instructors how to pace teaching and learning activities and indicate where difficulties may be occurring (Akpan, Notar, & Padgett, 2012).

Formative assessment provides valuable feedback to the students, parents/guardians and teachers outlining strengths, needs and next steps.

Examples of Basic Purposes for Which Classroom Assessment Results are Used

Formative uses	*Summative uses*
A. Size up of students at beginning of year	A. Assigning grades
B. Identifying individual students' learning needs	B. Placement
C. Identifying group learning needs	C. Self-evaluation

D. Designing instructional materials and activities. (Akpan, Notar, & Padgett, 2012).

Formative assessment is designed to assist the learning process by providing feedback to the learner, which can be used to identify strengths and weakness and hence improve future performance. Formative assessment is most appropriate where the results are to be used internally by those involved in the learning process [students, teachers, curriculum developers] (16).

As teachers, administrators, and counselors you may examine tests on the basis of format, intent, content, and condition. Tests which are given at the end of large blocks of instruction are generally summative in nature. They are used to evaluate student performance and usually occur infrequently. These tests have the disadvantage of identifying problems when it's too late to resolve them. Objectives which were not mastered may be identified without sufficient time left in the course to re-learn them. Also, if the instruction is cumulative all of the following modules may be failed because of failure of a previous module. Summative tests may be criterion tests, but when used in this fashion their ability is minimized.

Summative evaluations are culminating assessments conducted at the end of a unit, course, or grade level to determine the degree of mastery or proficiency according to identified achievement targets. They are used to determine the success of a section of instruction. Final tests for a grading period, assessments at the end of a unit (tests, term reports, summaries of portfolios), and end of the year assessments are all typical sources of summative data. They rely on a broad sampling of the relevant content; they focus more generally on all the objectives of the unit of instruction; and they are often a major part of the data collected for grading.

Summative assessment is used primarily to make decisions for grading or determine readiness for progression. Typically, summative assessment occurs at the end of an educational activity and is designed to judge the learner's overall performance. In addition to providing the basis for grade assignment, summative assessment is used to communicate students' abilities to external stakeholders, e.g., administrators and employers (16).

Informal vs. Formal Assessment With informal assessment, the judgments are integrated with other tasks, e.g., lecturer feedback on the answer to a question or preceptor feedback provided while performing a bedside procedure. Informal assessment is most often used to provide formative feedback.

As such, it tends to be less threatening and thus less stressful to the student. However, informal feedback is prone to high subjectivity or bias (16).

Formal assessment occurs when students are aware that the task that they are doing is for assessment purposes, e.g., a written examination or OSCE. Most formal assessments also are summative in nature and thus tend to have greater motivation impact and are associated with increased stress. Given their role in decision-making, formal assessments should be held to higher standards of reliability and validity than informal assessments (16).

Continuous vs. Final Assessment Continuous assessment occurs throughout a learning experience (intermittent is probably a more realistic term). Continuous assessment is most appropriate when student and/or instructor knowledge of progress or achievement is needed to determine the subsequent progression or sequence of activities. Continuous assessment provides both students and teachers with the information needed to improve teaching and learning in process. Obviously, continuous assessment involves increased effort for both teacher and student (16).

Final (or terminal) assessment is that which takes place only at the end of a learning activity. It is most appropriate when learning can only be assessed as a complete whole rather than as constituent parts. Typically, final assessment is used for summative decision-making. Obviously, due to its timing, final assessment cannot be used for formative purposes (16).

Characteristics of Formative and Summative Assessments

Characteristic	*Formative Assessment*	*Summative Assessment*
Purpose	To improve teaching and learning	Evaluation of learning outcomes
	To diagnose student difficulties	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	Does student understand the material
	What needs to be improved	Is the student prepared for next level of activity
	How can it be improved	
Examples	Observations	Projects
	Homework	Performance assessments
	Question and answer sessions	Porfolios
	Self-evaluations	Papers
	Reflections on performance	In-class examinations
	Curriculum-based measures	State and national tests

Table 4: Characteristics of Formative and Summative Assessments (54-58; Dixson & Worrell, 2016)

Summary of Definitions

Assessment System - all the systematic methods and procedures that are used to obtain information about behaviors and upon which educational decisions are based (Planning of measurement and evaluation)

Measurement - a process involving a structured situation that includes samples of particular characteristics or behaviors that results in a numerical or narrative score (Collection of data)

Test - a formal set of questions or tasks, often administered to a group of students, that address particular cognitive capabilities learned in a specific course or subject area (Instrument of measurement)

Evaluation - the use of assessment information to make judgments about students, teachers, or educational programs (Interpretation)

Writing test plan

Assessment Methods

A Test plan reflects your entire test schedule and approach. The easiest way to do this is to look at your program map, see where assessment of SLOs is already taking place, and select measures will provide insight into whether students are learning or developing the SLOs. This is called *course embedded* program assessment. Course embedded assessments can include (but are not limited to) questions embedded in exams, as well as course essays, papers/theses, problem sets, interviews, oral exams, case study reports, presentations, capstone projects, portfolios, etc. Note: If there are two or more sections of a given course, the course embedded program assessment for that course would be the same for both sections (19).

What are assessment methods? Assessment methods are the strategies, techniques, tools and instruments for collecting information to determine the extent to which students demonstrate desired learning outcomes. Several methods should be used to assess student learning outcomes. See the Assessment Methods Table for an overview of some commonly used direct and indirect methods of assessment (30).

What are direct and indirect methods of assessment? Direct methods of assessment ask students to demonstrate their learning while indirect methods ask students to reflect on their learning. Tests, essays, presentations, etc. are generally direct methods of assessment, and indirect methods include surveys and interviews (30).

What are embedded assessment methods? Embedded assessments utilize existing student course work as both a grading instrument as well as data for assessing SLO. Embedded assessments are also referred to as "classroom-based" or "continuous" assessments. Embedded assessments can assess individual student performance, the course, or the program if the information is aggregated; they can be formative or summative, quantitative or qualitative. If embedded assessments are properly designed, students should not be able to tell whether they are being taught or assessed. For example, as part of a course, each student completes a research paper that is graded for content and style, but is also assessed for advanced ability to locate and evaluate Web-based information [as part of a college-wide outcome to demonstrate information literacy] (30).

Process vs. Product Assessment Process assessment focuses on the steps or procedures underlying a particular ability or task, i.e., the cognitive steps in performing a mathematical operation or the procedure involved in analyzing a blood sample. Because it provides more detailed information, process assessment is most useful when a student is learning a new skill and for providing formative feedback to assist in improving performance (34).

Product assessment focuses on evaluating the result or outcome of a process. Using the above examples, we would focus on the answer to the math computation or the accuracy of the blood test results. Product assessment is most appropriate for documenting proficiency or competency in a given skill, i.e., for summative purposes. In general, product assessments are easier to create than product assessments, requiring only a specification of the attributes of the final product (34).

Divergent vs. Convergent Assessment Divergent assessments are those for which a range of answers or solutions might be considered correct. Examples include essay tests, and solutions to the typical types of indeterminate problems posed in PBL. Divergent assessments tend to be more authentic and most appropriate in evaluating higher cognitive skills. However, these types of assessment are often time consuming to evaluate and the resulting judgments often exhibit poor reliability (34).

A convergent assessment has only one correct response (per item). Objective test items are the best example and demonstrate the value of this approach in assessing knowledge. Obviously, convergent assessments are easier to evaluate or score than divergent assessments. Unfortunately, this "ease of use" often leads to their widespread application of this approach even when contrary to good assessment practices. Specifically, the familiarity and ease with which convergent assessment tools can be applied leads to two common evaluation fallacies: the Fallacy of False Quantification (the tendency to focus on what's easiest to measure) and the Law of the Instrument Fallacy [molding the evaluation problem to fit the tool] (34).

Most assessment is thought to be centered on the individual student. As Figure 8 shows that is not strictly true. A great deal of assessment is group centered.

Matrix for Classifying Assessment Options

	Formal instruments and activities: Options	Informal instruments and activities: Options
Group assessment	1. Text-embedded tests 2. Curriculum-embedded tests 3. Commercial criterion tests 4. Commercial normed tests 5. Rating scales 6. Performance tests 7. Questionnaires	1. Oral questions 2. Writing samples 3. Seatwork 4. Homework 5. Paper–pencil tests 6. Rating scales 7. Exihibitions 8. Portfolios 9. Demonstration 10. Peer assessment 11. Interviews
Individual assessment	1. Performance assessment 2. Standardized norm-referenced tests 3. Standardized criterion-referenced tests 4. Curriculum-embedded tests	1. Observation 2. Oral questioning 3. Writing sample 4. Homework 5. Seatwork 6. Paper–pencil tests 7. Portfolios 8. Interviews 9. Self-assessment 10. IEP monitoring 11. Error analysis

Figure 8: Matrix for Classifying Assessment Options (Notar, 2005)

Group Evaluation

Group work often is evaluated according to results, products, or presentations. In addition, teachers sometimes attempt to evaluate each group on its ability to cooperate, stay on task, fulfill individual responsibilities, and convey positive attitudes when groups or committees were meeting. Cooperative learning, for example, tends to encourage students to help one another meet their responsibilities. Stronger students often tutor weaker students in the various aspects of the assignment. In a cooperative learning activity, the teacher might evaluate the performance of a randomly selected group member; therefore, every student must be prepared to represent the group as a whole. Evaluation of group panel or committee reports can provide assessment difficulties. Some teachers give individual grades on oral performance, and others also give a group grade. To more fairly judge individual contributions to the group's presentation, some mentors require written material and reports from each student to help reveal the amount and quality of individual efforts (59-63).

Alternative Assessment

The term alternative assessment, and particular testing practices associated with it, have recently come into vogue in language testing. The movement is directed at establishing qualitative, more democratic, and task-based methods of evaluation in testing a learner's language proficiency (Brown & Hudson, 1998; Aschbacher, 1991; Herman, Aschbacher, & Winters, 1992; Huerta-Macías, 1995). It contrasts with traditional methods of testing by involving the learners in the evaluation process, and having the tendency to locate evaluation in a real-life context and, as result of these two features, being longitudinal. Thus, the insights emanating from these methods, alongside being used for decision-making about the future of learners, contribute to and furnish additional instructional purposes. As McNamara (2000) points out:

"This approach stresses the need for assessment to be integrated with the goals of the curriculum and to have a constructive relationship with teaching and learning" (as cited in 32).

The procedures used within this paradigm include checklists, journals, logs, videotapes and audiotapes, self-evaluation, teacher observations, portfolios, conferences, diaries, self- assessments and peer-assessments (Brown & Hudson, 1998). These procedures have been diversely called alternative or performance assessment as opposed to traditional assessment techniques such as multiple choice, cloze test, dictation, etc.

While the new movement promises more humanistic and rewarding methods of testing and thus has a lot to offer, most teachers are not quite familiar with the new concepts and practices within the emerging paradigm. To enlighten the views of interested teachers, it can be a good start to answer a basic question about the so-called alternative methods of testing which may have occupied their minds. This question is concerned with the relationship of these other methods with the traditional methods normally used within classrooms. Or to put the question another way, how can we place both traditional and alternative assessment methods in perspective to get a panoramic view of both in the pieced together jigsaw of language testing? To this purpose, it seems necessary to draw on the concepts of testing, measurement and evaluation (32). Figure 9 shows the major shifts in assessment practices.

Major Shifts in Assessment Practices

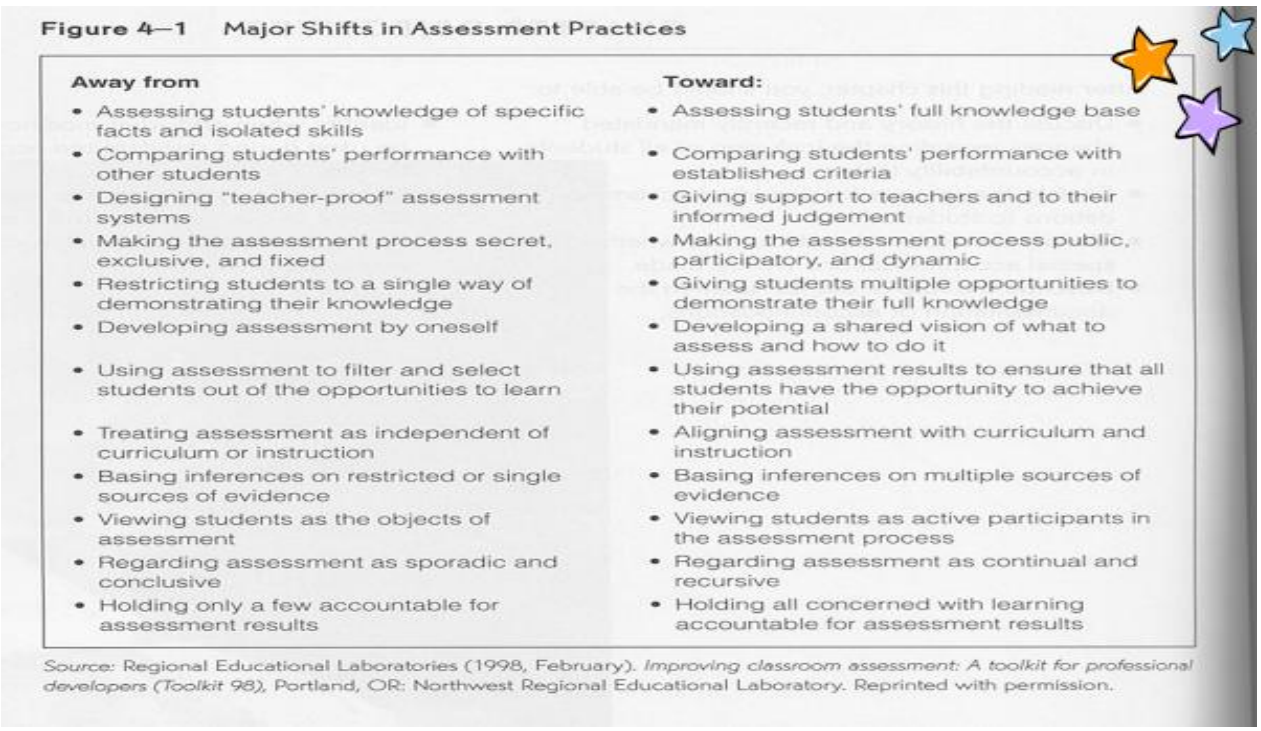


Figure 4—1 Major Shifts in Assessment Practices

Away from	Toward:
• Assessing students' knowledge of specific facts and isolated skills	• Assessing students' full knowledge base
• Comparing students' performance with other students	• Comparing students' performance with established criteria
• Designing "teacher-proof" assessment systems	• Giving support to teachers and to their informed judgement
• Making the assessment process secret, exclusive, and fixed	• Making the assessment process public, participatory, and dynamic
• Restricting students to a single way of demonstrating their knowledge	• Giving students multiple opportunities to demonstrate their full knowledge
• Developing assessment by oneself	• Developing a shared vision of what to assess and how to do it
• Using assessment to filter and select students out of the opportunities to learn	• Using assessment results to ensure that all students have the opportunity to achieve their potential
• Treating assessment as independent of curriculum or instruction	• Aligning assessment with curriculum and instruction
• Basing inferences on restricted or single sources of evidence	• Basing inferences on multiple sources of evidence
• Viewing students as the objects of assessment	• Viewing students as active participants in the assessment process
• Regarding assessment as sporadic and conclusive	• Regarding assessment as continual and recursive
• Holding only a few accountable for assessment results	• Holding all concerned with learning accountable for assessment results

Source: Regional Educational Laboratories (1998, February). *Improving classroom assessment: A toolkit for professional developers (Toolkit 98)*. Portland, OR: Northwest Regional Educational Laboratory. Reprinted with permission.

Figure 9: Major shifts in assessment practices

Performance Objectives *(Also Specific Learning Objectives, Learning Objectives, Behavioral Objectives)*

Performance-based Assessments (PBA)

Performance objectives are usually written in a behavioral objective format and contain three specific components: a clearly stated action, conditions, and a measurable standard. The action statement ordinarily indicates what a learner should be able to do after instruction. The key element is a word, ordinarily a verb, that clearly communicates a specific behavior. For instance, if students are to "know" the first 10 amendments to the U.S. Constitution in order to more completely understand the court system and thereby become better citizens, they might be asked to write the first 10 amendments and explain their meaning, along with an example and explanation of how each protects individuals' rights. Notice that students are asked to give written examples, make written descriptions, and provide written explanations.

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 (Pellegrino, 1999; 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. Figure 10 provides Dale's Cone of Experiences provides a bases for why Performance-based Assessments (PBA) and Authentic Assessment are seen as more reliable and valid assessments of learning. Figure 11 provides a pictorial view of Dale's Cone.

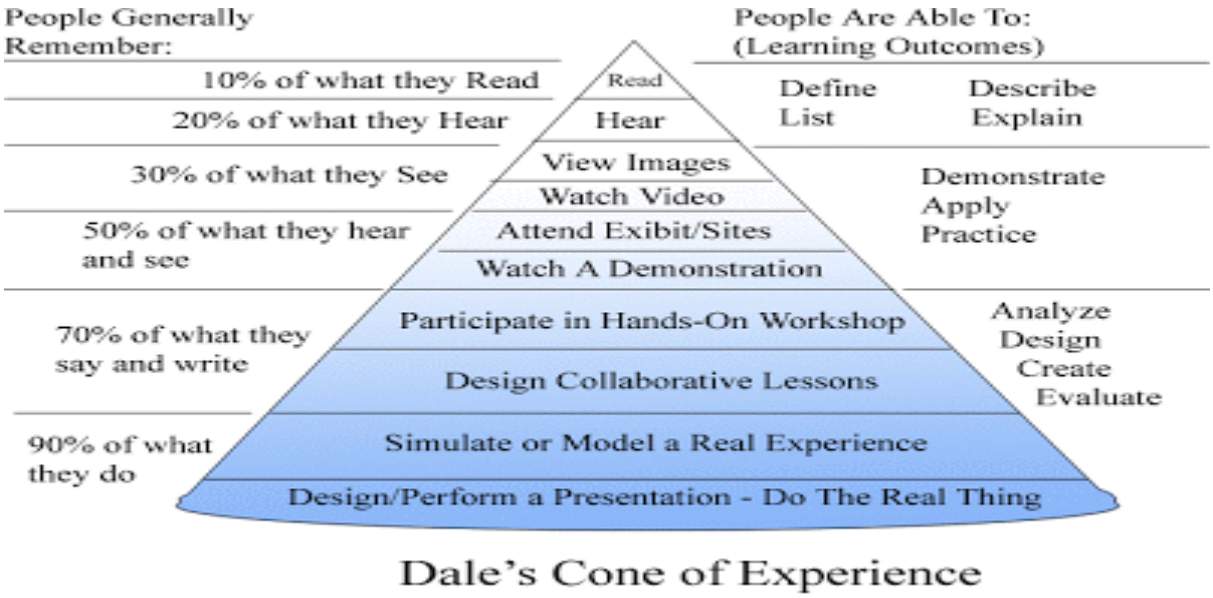


Figure 10: Dale's Cone of Experience (2).

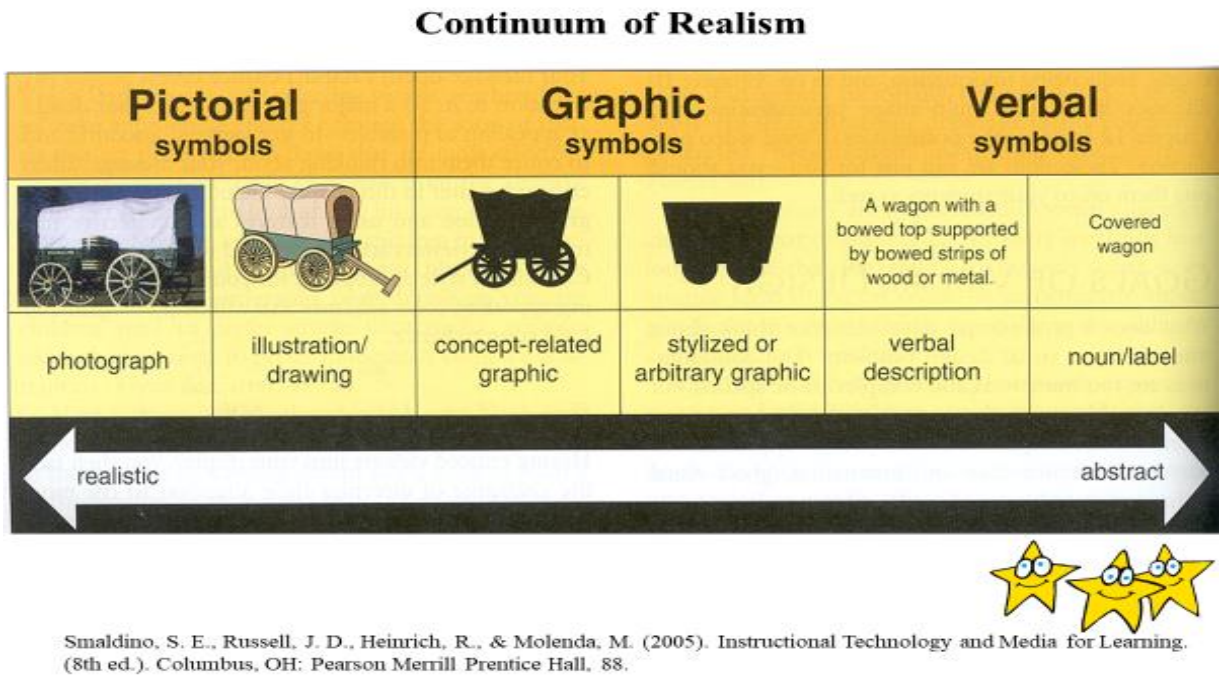


Figure 11: Continuum of realism

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. 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.

Table of Specifications

Consideration in a table of specifications include the final distribution of items to reflect the emphasis given during the instruction. Objectives considered more important by the teacher should be allotted more test items. Similarly, areas of content receiving more instruction time should be allotted more test items. Although the decisions involved in making the table are somewhat arbitrary and the process is time consuming, the preparation of the table of specifications is one of the best means for ensuring that the test will measure a representative sample of instructionally related tasks. It takes your objectives and build a pyramid with them.

You will see a layering effect similar to the steps in the domain you are teaching. This will insure that you are following the steps in the domain and also help you combine some objectives. It will ensure Time on Objective +Total Time Spent Teaching = Test emphasis. Time emphasis is defined as Direct + Integrated = (Time is in minutes). Therefore, a table of specifications is a two-dimensional table that relates the instructional objectives to the course content. The table makes it possible to classify each test item in terms of both objectives and content. A completed table describes the number of test items needed to obtain a balanced measure of the instructional objectives and the course content emphasized in the instruction.

Table of Specifications Components

The following items should be in a Table of Specifications:

1. Heading
2. Objectives
3. Level in taxonomy of the action verb
4. Levels of taxonomy being used
5. Time spent on objective in classroom (direct and integrated)
6. Percent of time of each objective v.v total instructional time
7. Types of questions to be asked for each objective

8. Number of each type question per objective and domain level of questions
9. Total number of questions on test and for each domain level
10. Total value of test
11. Value of points per objective and questions for objective
12. Grade level of test and subject matter

The Table of Specifications is used to show two things: first, the emphasis of the test item is equal to the emphasis of the instructional time. Instructors are testing what they taught. The second thing is a Table shows is the test is assessing at the appropriate level(s). If there are constraints, always test at the highest level. If an individual can perform the most difficult aspects of the objective, the instructor can (assume the lower levels can be done. However, if testing at the lower levels, the instructor cannot "assume" the individual can perform the higher levels. If there are no constrains, testing across levels can be conducted so as to indicate where a student or class erred when they did not perform at the highest level. The Table of Specifications ties teacher made tests and accountability together so the test meets validity and reliability requirements (Notar, Zuelke, Wilson & Yunker, 2004, p.128). For specific information on completing a Table read *The Table of Specifications: Insuring Accountability in Teacher Made*

Tests by Notar, Zuelke, Wilson and Yunker, 2004, *Journal of Instructional Psychology, 31*(2), 115-129. Tables 5-7 take you through examples of the upside-down triangle (figure 7) from yearly, semester, and unit.

Yearly Assessment Plan 11th Grade American History

Assessment instruments	Semester 1	Semester 2	Semester 3	Semester 4	Semester 5	Semester 6
Summative Evaluation						
Class groups		1 x 50	1 x 50	2 x 50	1 x 50	
Demonstrations			P1,3,5	P2,4,6		
			1 x 40	1 x 40		
Exams, Unit	3 x 100	2 x 100	2 x 100	2 x 100	2 x 100	2 x 100
Exams, Semester	1 x 100	1 x 100	1 x 100	1 x 100	1 x 100	1 x 100
Games		3 x 10	3 x 10	2 x 25	2 x 25	
Homework, Graded	6 x 10	6 x 10	5 x 10	4 x 10	3 x 10	2 x 10
	2 x 25	2 x 25	2 x 25	1 x 25		1 x 50
Oral presentation, Graded		1 x 15	1 x 15	1 x 25	1 x 25	1 x 50
Projects		P1,3,5	P2,4,6	P1,3,5	P2,4,6	
		1 x 40	1 x 40	1 x 100	1 x 100	
Portfolios		P2,4,6	P1,3,5	P2,4,6	P1,3,5	
		1 x 40	1 x 40	1 x 100	1 x 100	
Questioning (Oral)	10 x 5	10 x 5				
Quizzes, Graded	5 x 5	5 x 5	5 x 5			
Research projects			P1,3,5	P2,4,6		
			1 x 50	1 x 50	1 x 100	1 x 200
Seat work, Graded	8 x 5	5 x 5	5 x 5			
	4 x 10	4 x 1	3 x 20			
Simulations						1 x 200
Speaking		1 x 10	2 x 15	3 x 15	2 x 25	1 x 100
Writing	4 x 10	4 x 20	3 x 30	2 x 50	1 x 75	
TOTAL	**705**	**755**	**705**	**975**	**950**	**900**
Tests	**425 (60%)**	**325(42%)**	**325(43%)**	**300(31%)**	**300(32%)**	**300(33%)**
Classwork	**280 (40%)**	**430 (58%)**	**380 (57%)**	**675(69%)**	**650(68%)**	**600(67%)**
Formative Evaluation						
Anecdotal records	x	x	x	x	x	x
Checklist	4	3	2			
Homework	5	5	4	3	2	
Individual/Group/Teacher						
Meeting	x	x	x	x	x	x
Observation	x	x	x	x	x	x
Oral presentation	3	2	1			
Questioning (Oral)	x	x	x	x	x	x
Quizzes, Ungraded	5	4	3			
Seat work, Ungraded	5	4	3			
Speaking	3	2	1			
Writing	5	4	3			

Table 5: Yearly Assessment Plan 11th Grade American History (Notar, 2008).

Semester 6 Assessment Plan 11th Grade American History

Assessment instruments	Semester 6
Summative Evaluation	
Exams, Unit	2 x 100
Exams, Semester	1 x 100
Homework, Graded	2 x 10
	1 x 50
Oral presentation, Graded	1 x 50
Research projects	1 x 200
Simulations	1 x 200
Speaking	1 x 100
TOTAL	**900**
Tests	**300 (33%)**
Classwork	**600 (67%**
Formative Evaluation	
Anecdotal records	x
Individual/Group/Teacher Meetings	x
Observations	x
Questioning (Oral)	x

Table 6: Semester 6 Assessment Plan 11th Grade American History (Notar, 2008).

Unit Contemporary Issues Semester 6 Assessment Plan 11th Grade American History

Assessment instruments	Unit Contemporary Issues
Summative Evaluation	
Exams, Unit	1 x 100
Research projects	1 x 200
Speaking	1 x 100
TOTAL	**400**
Tests	**100 (25%)**
Classwork	**300 (75%)**
Formative Evaluation	
Anecdotal records	x
Individual/Group/Teacher Meetings	x
Observations	x
Questioning (Oral)	x

Table 7: Unit Contemporary Issues Semester 6 Assessment Plan 11th Grade American History (Notar, 2008).

How do you fix having the feeling of anticipation when you are about to take a quiz or test? Did you take detailed class notes and study enough? Have you studied what will be tested? When the test comes are you shocked that what was in the book or in class is not being tested? The unit organizer will fix those problems for you and your students. The unit organizer is based on your objectives and time spent on those objectives. Table 8 shows a blank unit organizer and table 9 shows a completed unit organizer.

Unit Organizer for Assessment Plan											
Learning Objective				**Item Type**	**Bloom's Taxonomy/Congruency**						
					Know	**Comp**	**Appl**	**Anal**	**Syn**	**Evl**	**Total**
No	**Level**	**ime***	**Q/P/%**								Q/P
1											
2											
3											
4											
5											
6											
7											
8											
9											
10											

MC=Multiple Choice; SA = Short Answer Q= Questions; P = Points

CPO= Class participation, Observation time on objective/task points

*Instructional time: total time spent on objective such as direct instruction, class activities, homework, integrated presentation, reading, etc. (Notar, 2008).

Table 8: Unit Organizer for Assessment Plan

Learning Objective				**Item Type**	**Bloom's Taxonomy/Congruency**						**Total**
No	**Level**	**Time***	**Q/P/%**		**Know**	**Comp**	**Appl**	**Anal**	**Syn**	**Eval**	**Q/P**
1	Appl	95 16%	9/16/16%	Match	2(1)	2(2)	5(2)				9/16
2	Comp	55 9%	5/9/9%	MC	1(1)	4(2)					5/9
3	Appl	50 9%	5/8/8%	MC Essay	2(1)		2(2) 1(2)				5/8
4	Appl	35 6%	3/6/6%	MC Essay	1(1)	1(1)	 1(4)				3/6
5	Syn	45 8%	5/8/8%	MC SA Essay			2(1)		2(1) 1(4)		5/8
6	Know	60 10%	6/10/10%	T/F MC	4(2) 2(1)						6/10
7	Appl	85 14%	8/14/14%	MC SH		2(1)	5(2) 1(2)				8/14
8	Anal	60 10%	6/10/10%	SA Essay				5(1) 1(5)			6/10
9	Comp	70 12%	7/12/12%	Match MC SA	2(1)	2(2) 2(2) 1(2)					7/12
10	Eval	40 7%	4/7/7%	SA Essay			1(1)	1(1)	1(1)	1(4)	4/7
Total		600/100	58/100/100		14/18	14/25	18/35	7/11	4/7	1/4	58/100

Table 9: Unit Organizer for Assessment Plan (Notar, 2008).

You have been taught that a heading comes first. However, in developing a unit test the heading is filled in last after you have completed your unit organizer for the test. Table 10 provides you with a completed heading.

Heading
Course Title: Art III Grade level: 6, 7, 8,9, 10, **11**, 12 (Circle as appropriate) Periods test is being used: 1 2 **3** 4 5 6 7 (Circle as appropriate) Date of test: April 15, 2019 Subject matter digest: 19th and 20th Century Art. Includes artists from around the world. Oils and Water colors primary medium. Identify major works, styles, and schools. Type test: Power, Speed, **Partially Speeded** (Circle One) Test Length: 45 minutes Test Value: 100 points Base Number of Test Questions: 75 Constraints: Test time, quantity of art available for test items

Table 10: Heading (Notar, 2008).

Within the heading you will come across the terms power, speed or partially speeded test. These terms have special meaning that you may or may not have come into contact with. Therefore, to save you the headache of the students telling you they did not have enough or too much time on a test the terms will be explained to your benefit. The author uses the word benefit because we have experienced the problem of too much or not enough in our years as a teacher. Main focus of a power test is in testing the cognitive ability the student possesses. Often there is less concern about the rapidity of a student's responses to questions than about the content of those responses. Accordingly, time limits on such tests are very generous, allowing all students enough time to consider each question and attempt to answer it. Test items tend to be difficult. These tests are called power tests.

The items of a power test have different levels of difficulty usually arranged in the order of increasing difficulty. A power test should be timed that a very large percentage of the pupils for whom it is designed will have ample time to attempt all of the items. Although estimates vary as to the proper size of this percentage, 90% seems to be an acceptable minimum. Hence, the test is slightly speeded (Notar, 2008).

Speed tests are used when it is the speed with which student's perform tasks or answer questions, rather than content of students' responses, is of prime interest. Tests with time limits so strict that no one is expected to complete all items. Speed tests usually contain questions or tasks that are relatively easy. A speed test is one in which a pupil must, in a limited amount of time, answer a series of questions or perform a series of tasks of a uniformly low level of difficulty. The near-constant level of difficulty of the questions or tasks is such that, if the pupil had unlimited time, he could easily answer each question or perform each task successfully. The intent of a speed test is to measure primarily the rapidity with which a pupil can do what is asked of him rather than whether he can do it at all (Notar, 2008).

A partially speeded tests has a purpose of testing may change over the course of instruction, even though the items on the test remain essentially the same. Early in the course of learning, emphasis is usually not on speed of performance. Speed of performance frequently becomes important after students have mastered tasks. Two examples are the use of manipulatives and phonics. Tests for many students are a mixture of speed and power, even when the test developer intends only power. Such tests are called partially speeded tests. When constructing tests, the teacher will have to check the time limits carefully to be sure that all students have the opportunity to consider each test question before the time is up (Notar, 2008).

Test Construction

As discussed earlier everything starts with objective. This is true with test construction. Figure 12 provides a relationship between types of test items and what level of learning to be tested. The term congruence comes into play. Congruence means agreement or harmony; compatibility.

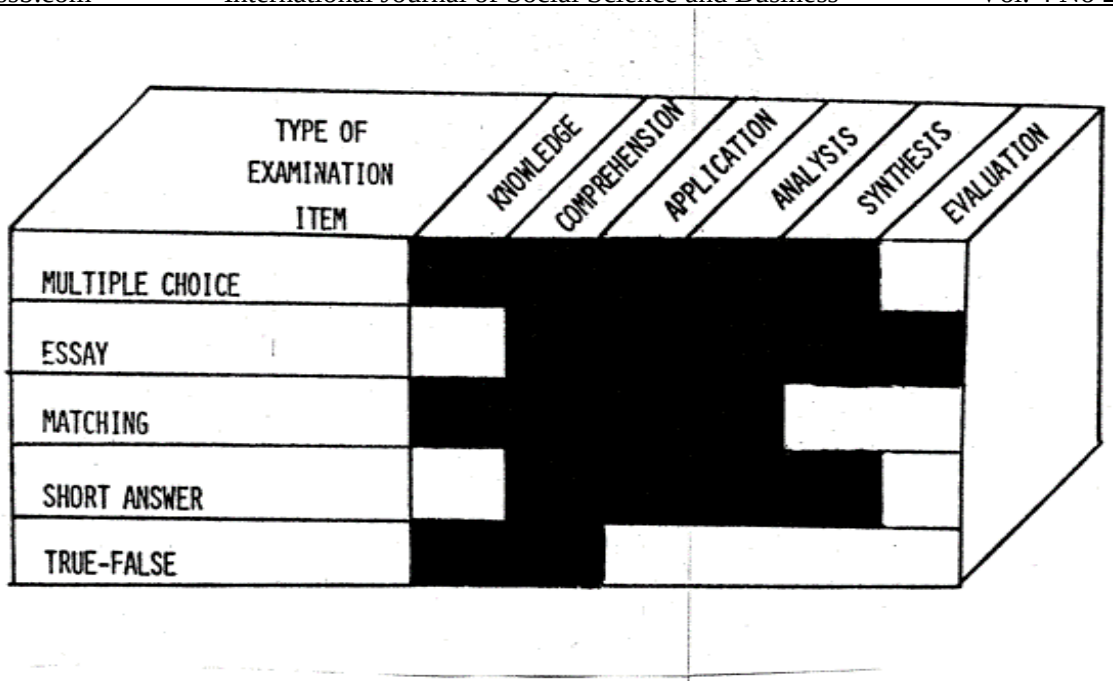


Figure 12: Types of test items and levels of learning (Notar, 2008).

Two examples of congruence are provided:

Example 1

Teaching Knowledge
Learning Knowledge
Testing Knowledge

Example 2

Teaching Application
Learning Application
Testing Knowledge, Application

Knowledge, Comprehension, Application
Comprehension, Application
Application

Table 11 provides a number of tools to assess student performance. This list is not all inclusive. A number of the tools provide outcomes beyond paper and pencil test and will be discussed in the next section.

Anecdotal records Achievement test Aptitude test Attitude test Authentic assessment Behavior counts Checklist Class groups Conference/interviews Daily record Diagnostic testing Demonstrations Documents (homework, seat work, etc.) Error analysis Exams Functional/adaptive assessments Grades Hands on testing (Performance) Interviews Journals (response/learning) Observation Oral presentation Peer evaluations Process assessment	Product assessment Performance assessments Projects Portfolios (documentation of progress, showcase) Progress chart/reports Profiles Questionnaires Questioning (Oral) Quizzes Records (behavior, class, conduct, daily, group, student, test) Report cards Research projects Rubric Self-evaluations Simulations Speaking Standardized tests (Achievement, Aptitude, and Attitude test) Teacher made tests Think-aloud techniques Work-sample analysis Writing

Table 11: Assessment, evaluation, measurement tools (Notar, 2008).

Once you have developed your Table of Specification you will need to develop your paper and pencil tests if you selected those tools. Table 12 provides guidance on selecting what type of test item to use once you have determined the level of learning to be tested. Figure 13 shows differences between selection and supply type test items and Figure 14 helps determine which type essay measured when using restricted and extended response essay questions should be used when supply test items are selected.

Comparison of Test Item Characteristics between Selection and Supply Type Items

Test Item Characteristics	Selection Test Items	Supply Test Item (Short Answer)	Supply Test Item (Essay)
Measures factual information	Yes	Yes	Yes
Measures understanding	Yes	No	Yes
Measures synthesis	No	No	Yes
Easy to Construct	No	Yes	Yes
Samples broadly	Yes	Yes	No
Eliminates bluffing	Yes	No	No
Eliminates writing skill	Yes	No	No
Eliminates blind guessing	No	Yes	Yes
Easy to Score	Yes	No	No
Scoring is objective	Yes	No	No
Pinpoints learning errors	Yes	Yes	No
Encourages originality	No	No	Yes

Table 12: Comparison of test item characteristics between selection and supply type items (Notar, 2008).

Selection Type Test Item Value	**Supply Type Test Item Value**
Measures factual information	Measures factual information
Measures understanding	Measures understanding
Sample's broadly	Measures synthesis
Eliminates bluffing	Easy to construct
Eliminates writing skill	Eliminates blind guessing
Easy to score	Encourages originality
Scoring is objective	
Pinpoints learning errors	

Figure 13: Differences between selection and supply type test items (Notar, 2008).

Restricted response essay questions	**Extended response essay questions**
explain cause-effect relationships	produce, organize, and express ideas
describe applications of principles	integrate learning from different content areas
present relevant arguments	evaluate the worth of ideas
formulate tenable hypotheses	
formulate valid conclusions	
state necessary assumptions	
describe the limitations of data	
explain methods and procedures	

Figure 14: What is measured when using restricted and extended response essay questions (Notar, 2008).

The terms power, speed or partially speeded were introduced above. Table 13 provides you the test writer time ranges for the various types of test items you will write. In a 55 minute class how long do you actually have for a student to take a test. Once you have decided if you are doing a power, speed or partially speeded test you will need to figure out exact questions that will be on the test.

Completion Time Requirements for Different Types of Test Items.

Test Item Type	Completion Time
True-false items	20-30 seconds
Multiple-choice (factual)	40-60 seconds
Multiple-choice (complex)	70-90 seconds
Multiple-choice (w/calculations)	2-5 minutes
Matching (5 stems/6 choices)	2-4 minutes
One-word fill-in	40-60 seconds
Short-answer	2-4 minutes
Word problems (simple arithmetic)	5-10 minutes
Short essays	15-20 minutes
Data analyses/graphing	15-25 minutes
Drawing models/labeling	20-30 minutes
Extended essays	35-50 minutes

Table 13: Completion Time Requirements for different types of test items (Notar, 2008).

Figure 15 provides criteria for where item types should be in the test.

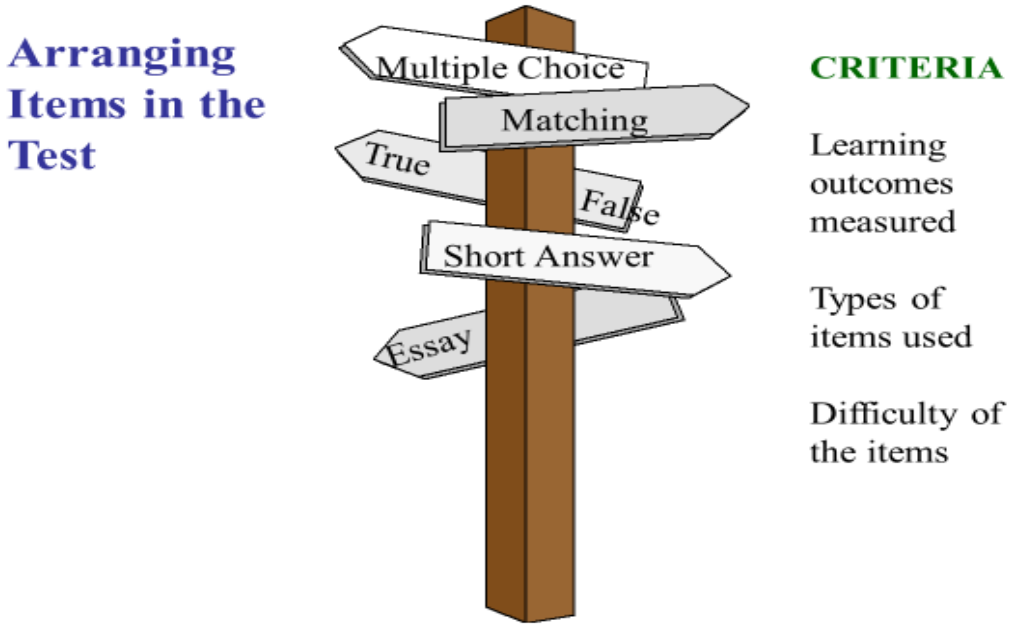


Figure 15: Arranging test items in a test (Notar, 2008).

Outcomes Beyond Paper and Pencil Test.

In table 11 the author mentioned outcomes beyond paper and pencil test and said explanations for some these outcomes would be forthcoming. Functional/adaptive assessments look at how a child actually functions at home, at school, and in the community. This type of assessment shows what the student can do or needs to learn that is not being reflected in test scores. Functional assessments for some students include looking at reading, writing, and math skills. For others, evaluating whether they are able to ride the city bus, dress independently, or handle money might be considered more appropriate activities to evaluate (Adaptive Behavior Assessment System-II, 2008).

Error Analyses

Error analyses allow teachers to examine student's responses on work samples to identify areas of difficulty and patterns in how students approach a task. Error analyses usually focus on identifying errors related to inappropriate applications of rules and concepts, rather than careless, random errors or errors caused by lack of instruction. An important aspect of error analyses is the students' explanation of their responses (i.e., think-aloud technique), which can help the teacher pinpoint faulty conceptual or procedural knowledge for developing remedial programming.

Interviews

An interview is an interaction, generally face to face, in which the participants verbally share information about the student. Generally, the interviewer follows a prescribed set of questions in a personal and informal atmosphere that encourages the sharing of perspectives, experiences, observations, or background information about the child. Interviews can be conducted with parents, colleagues, other professionals, and students. Interviews may be more appropriate than a questionnaire when there are literacy barriers, when the directions are complex, when questions need to be explained, when further probing is needed, or when the person being questioned needs to be reassured or encouraged to respond.

Peer Evaluations

Peer evaluations help students apply criteria to samples of work in a manner that is less threatening than self-evaluation might be. They teach respect for the ideas of others and positive methods of interacting by requiring confirming statements about each other's work, as well as constructive criticism that is use" for revision. Peer evaluations promote positive learning characteristics, while increasing student motivation, responsibility, self-direction, success, and self-esteem. They also help students gain insight into the thinking and reasoning processes of their classmates.

It is important that students be involved in the development of scoring criteria for self-evaluations and peer evaluations. Doing so helps to ensure that they are committed to and invested in the evaluation procedure and that they are clearly aware of the standards being used. Involving them in this process also gives them a sense of ownership in the grading system, which helps them to understand and value the learning process. in this way, the development of scoring criteria becomes part of the learning process and reinforces student knowledge of key concepts as expectations are delineated.

Some research shows that students work better in pairs when they are allowed to choose their own partners (Meisinger, Schwanenflugel, Bradley, Kuhn, & Stahl, 2004). Allowing students to choose partners tends to result in fewer squabbles between partners and more time spent on task (Lehr, Osborn, & Hiebert, 2004).

Questionnaires

A questionnaire is a group of questions that allow teachers to elicit information from parents, students, or other professionals in more detail than can be elicited from a checklist or rating scale. Questionnaires can be used for face-to-face interviews, or they can be mailed to the respondent to be filled out and mailed back. The format can be open-ended, which enables respondents to share their opinions, express their concerns or feelings, or to take their time to respond thoughtfully and comprehensively. This toot is especially beneficial when respondents need to gather information, such as developmental, medical, or school history data. Other questionnaire formats include multiple-choice, true-false, and fill-in-the-blanks, and response forms, on which respondents mark the appropriate picture or icon. Questionnaires may not be appropriate for parents who find reading and/or writing difficult.

Response Journals (Response/Learning)

Response journals, or learning logs allow students to keep a personal record of their work, including what they learned, how they learned it, what they did not understand, why they are confused, and what help they need. Students can also use this tool to maintain a personal journal in which they can reflect on, describe, analyze, and evaluate their learning experiences, successes, and challenges, as well as to write about the conclusions they draw from these events (Stiggins, 1997). Student journal entries can include assignments they have and have not mastered, information or strategies they have found useful, questions they want to ask, ideas for future projects, steps in planning an assignment, reflections on their work, and documentation of their progress (Kulm, 1994).

Students with poor writing skills may need to maintain an audio journal by recording their responses on an audiocassette.

Self-Evaluations

Self-evaluations are a method of collecting data by having students report on their feelings, activity level, or knowledge. This method can take many forms, including self-rating scales, attitude or interest inventories, portfolios, and journals or logs. It allows students to reflect on their learning, to directly apply grading standards to their work, and to contemplate on their personal strengths and areas that need remediation or m enforcement. It also helps students to project future goals and to develop strategies for achieving them. The self-evaluation process can be a powerful tool for life-long learning because it helps to promote metacognitive skills, ownership in learning, self-monitoring, and independence of thought (National Council of Teachers of Mathematics, 1991).
There are two basic requirements for making self-assessments accurate and reliable for students: (1) Students must be able to express themselves well enough verbally or nonverbally to make their desires understood; and (2) students must be reliable sources of information. As a result of these requirements, self-reports are most commonly used with older students and adults (Seaman, DePauw, Morton, & Omoto, 2003, p. 92).

Think-Aloud Technique

The think-aloud technique is a type of assessment in which students verbally explain the cognitive processes and steps they use while working on a task. This technique includes having students orally explain how they solve a math problem, outline a social studies chapter, use metacognitive skills when reading, plan for long-term assignments, and conduct science experiments. This technique helps teachers to understand how their students approach learning tasks, thus providing insight into any confusion or inaccuracies that are occurring, so that instructional objectives or interventions can be modified. To effectively employ this procedure, teachers need to be (1) astute observers of student performance, (2) knowledgeable about the scope and sequence of the curriculum, and (3) familiar with cognitive strategies (McLoughlin & Lewis, 2005). The think-aloud technique also helps students to gain insight into their own ability to organize, analyze, process information, and solve problems.

Since this may be a new experience for most students, the teacher may need to talk through the steps involved as students solve a simple problem and give them opportunities to practice this technique.

Work-Sample Analyses

Work-sample analyses involve reviewing student's work products by focusing on the quality and quantity of their output. These work samples-which are analyzed to determine areas of success and areas that require review or remediation-can include essays, homework assignments, lab reports, tests, and audio- and videotape recordings of a class discussion. Teachers can use error analysis to analyze the type and frequency of correct and incorrect responses students make on everyday assignments. Also, teachers can focus on other aspects of the work product, such as whether students followed directions; answered questions completely; produced a sufficient amount of work; worked in a sequential, organized manner; used adequate motor planning; copied accurately from the board or textbook; and demonstrated adequate penmanship skills.

Test **A**dministrative **M**anual

Once you have your Table of Specifications and your test written it will be time to administer your assessment. A **T**est **A**dministrative **M**anual (TAM) needs to be prepared. Assessment depends on many factors. A TAM helps you evaluate constraints (Figure 6) Time, Personnel, Cost, Equipment, Facilities, Realism, Logistics, Communications, and Others.

Some factors are not under the control of the administrator are how fatigued a test taker is, motivation level of the test taker, physical discomfort, and test anxiety (53). Because we know these factors exist to affect test scores for reasons other than ability, teachers must seek to standardize all aspects of the test under their control to minimize variability due to factors other than ability. Also, teachers should not test children when possible during typical lunch or playground time or immediately after holidays or exciting events. Don't test longer than 1 hr. (30 min attention span for preschool and elementary school children) or longer than 90 min. for secondary school children.

The TAM is to insure we are testing the student's knowledge and not if they can take a test. The TAM tells students if the test is open or closed book. Cheating will be dealt with according to school policy. State what the school policy is … zero on test, note to parents, suspension, etc.

- The TAM is to insure we are testing the student's knowledge and not if they can take a test.
- Test admin instructions are uniform for all persons tested.
- Contingency planning
- Absence of other test administrators
- Power failure
- Examinee no shows
- Medical emergency
- Equipment breakdown
- Fire in classroom/building

The TAM is a three phases checklist for test administrator and students: Before, During, and After.

Before the Test

Group together all items of similar format; form sections by item type.

Arrange test items from easy to hard; arrange both the sections and the items within sections in an ascending order of difficulty.

Group the items within each section by the learning outcomes measured.

Use subject matter groupings only when needed for some specific purpose.

Space the items for easy reading.

Keep items and options on the same page.

Positions illustrations near descriptions.

Check your answer key.

Determine how students record answers.

Provide space for name and date.

Check test directions.

The numbers of items to which they apply.

How to record answer.

The basis on which to select answer.

Criteria for scoring.

Proofread the test.

During the Test

Maintain a Positive Attitude
Motivation
Avoid Surprises
Rotate Distribution
Monitor Students
Give time Warnings
Maximize Achievement
Equalize Advantages
Clarify the Rules
Remind Students to Check Their Copies
Minimize Distractions
Collect Tests Uniformly

After the Test

All tests have been handed in

All answer sheets have been handed in with names or other identification indicating whose test paper was whose.

Examinees know when grades or test papers can be collected or will be posted.

Return the test room to its pre-test set up (53).

Critique

Now that the assessment is over you need to provide the students with results. The results are feedback that needs to be immediate, not days after the assessment. A critique is a learning activity.

Every time you stop and look at students work and give them feedback you are doing a critique. A test critique is the same thing. However, to make sure it is not a gripe session it must be focused and organized. The following are requirements for a successful critique.

Acceptability

Establish rapport and mutual respect. Students must accept you before they can willingly accept a critique. They must have confidence in your qualifications, teaching ability, sincerity, competence, and authority. Your manner, attitude, and knowledge of the subject, along with conviction and sincerity, will be accepted far more than your rank or position.

Objectivity

To be effective, a critique should focus on the student's performance; it should not reflect your personal opinions, likes, dislikes, and biases. The student's personality and opinion are not at stake nor should they have to agree or disagree with your beliefs. Although you need to be open and honest, you do not have license to ridicule or show anger.

Constructiveness

To be constructive, criticism should point toward improvement or a higher level of performance. When you identify a fault or weakness, accompany it with positive guidance for improvement.

Flexibility

Critiquing someone's performance calls for flexibility. Critiques should never be so rigidly designed or executed that you cannot allow for variables. Avoid mechanical, predetermined techniques and preconceived opinions regarding content, subject matter, and student capability. Quite often you will be confronted with the problem of selectivity, i.e.; what to say, what to omit, what to stress, and what to minimize. Vary your organization and method of critiquing according to the situation. Adapt your tone, technique, method, organization, and content of the critique to the occasion and to the student. To determine your, approach, consider the class situation, student ability, subject matter, and the time allotted for the critique.

Organization

Unless the critique follows some pattern of organization, valid comments may lose their impact. Almost any pattern is acceptable if both you, and the students, find it logical and easy to understand. An example would be to start with general comments, continue with a group critique, and finally request individuals to comment on their performances. Whatever the organization of the critique you should be prepared to change if the students cannot follow the critique.

Comprehensiveness

Although the critique should be comprehensive, it doesn't necessarily have to be long. It depends on whether you want to discuss major or minor points. To be effective however, feedback has to include both strengths and weaknesses and you will have to determine the best balance between the two. It is a disservice to students if you confine your comments to the excellence of their performance; you must also discuss those areas that need improving. Be specific with your comments and recommendations. When you are finished with the critique session students should have no doubts concerning what they did wells what they did poorly and, most importantly, how they can improve.

Assessment as Reflection

Teachers must be reflective. In this reflection mode beliefs are carefully examined, including values, visions, biases, and paradigms. Beliefs, which result primarily from experiences, influence how we think and behave. They form the lens through which we view the world and acquire an understanding of it.

Goals are also analyzed. In this effort, aims, outcomes, and intentions are compared. Thoughtfully considered differences are used to clarify and refocus educational efforts. Reflecting on practice refers to an analysis of dispositions, behaviors, and skills related to teaching performance. This process involves designing instruction and assessment strategies, interacting with and influencing students, developing family relationships, collaborating with colleagues, and initiating school-wide reforms.

Students' learning is linked to the learning of their teachers (Richardson, 1997). When teachers reflect on their practice, conditions are created to improve teaching because reflection accomplishes the following:

- It creates opportunities for continuous learning.
- It provides a variety of perspectives to draw on in addressing the many challenging and complex components of teaching.
- It provides new knowledge and understandings that have immediate applications because they are created within the context of teaching.
- It provides teachers a sense of personal responsibility for learning and improvement.

Reflection provides an avenue for strengthening relationships with other teachers and developing shared goals and developmental activities. It helps teachers build bridges between theory and practice. It reduces the likelihood that external mandates will be imposed because schools based on reflective practice already show improvements (York-Barr, & Duke, 2004; Taylor, & Nolen, 2005).

Whether you reflect or not assessment applied to your teaching will improve your teaching in several ways. Knowing how to choose or to craft quality assessments increases the quality of your teaching decisions. By assessing how your students use their knowledge and skills, you are able to monitor and evaluate their progress. This allows you to plan better teaching.

What and how you assess communicates in a powerful way what you really value in your students' learning. For example, you may tell your students how important it is for them to be independent and critical thinkers, but your words are empty if your assessments consist of only a few matching exercises based on facts from the text- book or handouts. On the other hand, if your assessments require students to integrate their knowledge and skills to solve "real-life" problems, they learn that you expect them to develop integrating and problem-solving abilities.
When you carefully define assessment tasks, you are clarifying what you want students to learn. When you craft assessment tasks, you learn how to create situations in which students can demonstrate their achievement. These skills apply directly to your teaching, because to teach effectively you must have clearly in mind how students should demonstrate their achievement.

You use your knowledge of how to craft quality assessment tasks when you evaluate assessment materials available from other sources. Your knowledge of the craft will also help you evaluate and become a critical consumer of assessment procedures, whether they are part of your curriculum materials or are imposed by or on your school district, such as standardized achievement tests and state-mandated assessments.

Learning to craft assessment tasks increases your freedom to design lessons. Knowing how to assess students validly, especially in relation to higher-order thinking skills, means that you are no longer chained to the assessment procedures already prepared by textbook publishers and others. You can use a wider variety of teaching strategies because you are able to assess students using your own assessment procedures.

You will improve the validity of your interpretations and uses of assessment results. Research shows that teachers who have studied assessment, either through coursework or in-service training, are able to recognize and produce better assessments.

You will improve your appreciation of the strengths and limitations of each type of assessment procedure. As a professional, you must have the knowledge and skill to independently evaluate proposed assessment approaches. You must evaluate their general educational value and their technical quality. Without both evaluations, you cannot determine whether these approaches can make good on the promises implied by their promoters. To fulfill your professional responsibility, you need a solid foundation in the basic principles of assessment development and in the criteria for valid use of the assessment results.

Perhaps you have heard that the global work culture is changing. Unlike your grandfather, you will probably have a number of different jobs and careers during your lifetime. In order to be successful, you will need to have confidence in your ability to learn and you will need to become a lifelong learner. Assessment plays a key role in developing your confidence in your ability to learn, as well as in developing your lifelong learning skills (41).

As a teacher this is especially true in your professional role and responsibilities for student assessment. You should be skilled on choosing assessment methods appropriate for instructional decisions and in developing assessment methods appropriate for instructional decisions. Teachers skilled in administering, scoring, and interpreting the results of both externally-produced and teacher-produced assessment methods and using assessment results when making decisions about individual students, planning teaching, developing curriculum, and school improvements (42-44).

Developing valid pupil grading procedures which use pupil assessments and communicating assessment results to students, parents, other lay audiences, and other educators is and art and skill you as a teacher must continual develop (42-44).

As professional you must be able to recognize unethical, illegal, and otherwise inappropriate assessment methods and uses of assessment information (42-44).

Summary

Assessment is designed to highlight important ideas of instruction. Assessment needs to be interwoven throughout instruction, gathering evidence that reflects and supports the goals of instruction. Multiple assessments need to be completed over time throughout the instruction. Assessments need to sample what students know and are able to do within the framework of the instruction. Each assessment provides some evidence about what and how students learn during the instruction (Niess, Lee, & Kajder, 2008).

Conclusion

This whole article has been focused on assessment for learning. Assessment is used for a multitude of educational purposes: person's aptitude to learn something; motivation to achieve in school; self-concept; achievement level in a scholastic area; and environmental factors that affect how much a person learns. However, a caution must be made for when assessment does not promote learning and the reason is the teacher uses assessment only to assign grades rather than to help determine what to teach.

Effective assessment is a continuous process. It's not simply something that's done at the conclusion of a unit of study or at the end of a lesson. Effective assessment and evaluation are integrated into all aspects of the curriculum, providing both teachers and students with relevant and useful data to gauge progress (37).

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.

Teachers, administrators, and counselors must use both summative and formative types of tests, however, the more formative testing is used the better the chances that the students will do well on the summative tests.

References

1 retrieved from https://www.northwestern.edu/searle/docs/developing-evaluation-plan.pdf
2 retrieved from http://rinajanetabinas.blogspot.com/2010/10/edgar-dales-cone-of-experience.html
3 retrieved from https://www.cmich.edu/office_provost/AcademicAffairs/CAA/Assessment...
4retrieved from https://www.softwaretestinghelp.com/test-plan-sample-softwaretesting-and-quality-assurance-templates/
5 retrieved from https://klariti.com/test-plan-template/
6 retrieved from https://www2.cdc.gov/.../templates/CDC_UP_Test_Plan_Template.doc · Web view
7 retrieved from https://www.guru99.com/what-everybody-ought-to-know-about-test...

8 retrieved from https://reqtest.com/testing-blog/how-to-write-a-test-plan-2
9 retrieved from https://www.guardengine.com/Template For Test Plan/results
10 retrieved from https://www.softwaretestinghelp.com/how-to-write-test-plan...
11 retrieved from https://www.guru99.com/test-plan-for-project.html
12 retrieved from https://www.opencodez.com/software-testing/test-planning.htm
13 retrieved from https://reqtest.com/testing-blog/how-to-write-a-test-plan-2/
14 retrieved from https://www.education.gov.za/Portals/0/Plan%20and%20conduct%20assessment.pdf?ver=2009-10-14-124958-510
15 retrieved from http://www.clark.edu/tlc/outcome_assessment/documents/CreatingaComprehensiveAssessmentPlan.pdf
16 retrieved from Compiled by Dr JJ, FSG, UiTM, Shah Alam jjnita@gmail.com, jjnita@salam.uitm.edu.my http://drjj.uitm.edu.my Mobile: +60193551621 Page 7 of 43
17 retrieved from drjj.uitm.edu.my/DRJJ/OBE FSG Dec07/OBEJan2010/DrJJ-Measure-assess...
18 retrieved from www.bgcdsb.org/UserFiles/Servers/Server_5912063/File/NTIP/Ministry...
19 retrieved from www.simmons.edu/.../developing-and-implementing-an-assessment-plan
20 retrieved from https://eric.ed.gov/?id=ED178109
21 retrieved from www.pdc.edu/.../2012/02/Student-Outcomes-Assessment-Plan_1-15-09.pdf
22 retrieved from https://www.cmu.edu/teaching/assessment/basics/alignment.html
23 retrieved from https://www.onlineassessmenttool.com/knowledge-center/assessment-knowledge-center/assessment-vs-evaluation/item10642
24 retrieved from https://keydifferences.com/difference-between-assessment-and...
25 retrieved from http://www.itlal.org/?q=node/93
26 retrieved from https://study.com/.../the-importance-of-assessment-in-education.html
27 retrieved from www.nea.org/assets/docs/HE/TeachrAssmntWhtPaperTransform10_2.pdf
28 retrieved from Compiled by Dr JJ, FSG, UiTM, Shah Alam jjnita@gmail.com, jjnita@salam.uitm.edu.my http://drjj.uitm.edu.my Mobile: +60193551621 Page 13 of 43
29 retrieved from https://www.vetmed.wsu.edu/docs/.../Docs.../program_assessment_process.pdf?...0
30 retrieved from https://www.wssu.edu/about/assessment-and-research/niloa/_files/documents/assessmentmethods.pdf
31 retrieved from https://www.cdc.gov/obesity/downloads/cdc-evaluation-workbook-508.pdf
32 retrieved from Compiled by Dr JJ, FSG, UiTM, Shah Alam jjnita@gmail.com, jjnita@salam.uitm.edu.my http://drjj.uitm.edu.my Mobile: +60193551621 Page 12 of 43
33 retrieved from Compiled by Dr JJ, FSG, UiTM, Shah Alam jjnita@gmail.com, jjnita@salam.uitm.edu.my http://drjj.uitm.edu.my Mobile: +60193551621 Page 6 of 43
34 retrieved from Compiled by Dr JJ, FSG, UiTM, Shah Alam jjnita@gmail.com, jjnita@salam.uitm.edu.my http://drjj.uitm.edu.my Mobile: +60193551621 Page 8 of 43
35 retrieved from www.itlal.org/?q=node/93
36 retrieved from https://keydifferences.com/difference-between-assessment-and...
37 retrieved from https://www.teachervision.com/assessment-vs-evaluation
38 retrieved from https://www.cdc.gov/obesity/downloads/cdc-evaluation-workbook-508.pdf
39 retrieved from www.differencebetween.net/miscellaneous/difference-between...
40 retrieved from forum.wordreference.com › English Only
41 retrieved from http://tutorials.istudy.psu.edu/testing/testing2.html
42 retrieved from https://www.aft.org
43 retrieved from www.oalib.com/references/5625964
44 retrieved from https://eric.ed.gov/?id=ED323186
45 retrieved from https://www.linkedin.com/pulse/differences-between-testing..
46 retrieved from https://www.bing.com/images/search?q=chart%20instruction%20is%20most%20effective%20when%20...%20Assessment%20is%20most%20effective%20when%20...&cbir=sbi&imageBin=&qs=n&form=QBIR&sp=-1&pq=chart%20instruction%20is%20most%20effective%20when%20...%20assessment%20is%20most%20effective%20when%20...&sc=0-82&sk=&cvid=E2AA05BF434D4B91B3CE4EDAE119A7F2

47 retrieved from https://www.bing.com/images/search?q=Location%20of%20assessment%20in%20instructional%20process%20%20%20&cbir=sbi&imageBin=&qs=n&form=QBIR&sp=-1&pq=&sc=0-0&sk=&cvid=9A283735DC1A4E4897C19E35A71AA407
48 retrieved from https://www.bing.com/images/search?q=Seven%20purposes%20for%20assessment%20&cbir=sbi&imageBin=&qs=n&form=QBIR&sp=-1&pq=&sc=0-0&sk=&cvid=4200FC2A30A1406DB81BA99092DA42C7
49 retrieved from http://www.adprima.com/objectives.htm
50 retrieved from https://msu.edu/~dwong/StudentWorkArchive/CEP900F00-RIP/Haniza-TestWiseness.htm
51 retrieved from http://www.specialconnections.ku.edu/?q=assessment/quality_test_construction/teacher_tools/testwiseness_and_guessing
52 retrieved from http://homepages.rpi.edu/~verwyc/Chap3tm.htm
53 retrieved from https://www.researchgate.net/publication/309881618_Examining_the_effects_of_testwiseness_in_conceptual_physics_evaluations
54 retrieved from http://www.brilliant-insane.com/2014/04/7-characteristics-of-quality-summative-assessments.html
55 retrieved from https://www.bartleby.com/essay/Characteristics-of-Formative-and-Summative-Assessment-in-FKZ6GFPBYRPS
56 retrieved from https://halfbaked.education/characteristics-of-effective-summative-assessment/
57 retrieved from https://link.springer.com/chapter/10.1007%2F0-306-47227-9_4
58 retrieved from https://www.aeseducation.com/blog/formative-vs.-summative-assessments-what-do-they-mean
59 retrieved from https://k12teacherstaffdevelopment.com/tlb/assessment-strategies-for-classroom-group-work/
60 retrieved from https://www.teachervision.com/professional-development/cooperative-learning
61 retrieved from https://www.k12irc.org/development/collaboration2/group.php
62 retrieved from https://cirt.gcu.edu/teaching3/tips/groupwork
63 retrieved from https://www.brighthubeducation.com/teaching-methods-tips/19619-advantages-of-group-work-in-the-classroom/
64 retrieved from https://keydifferences.com/difference-between-assessment-and...

Adaptive Behavior Assessment System-II. (2008). retrieved from https://www.sciencedirect.com/book/9780123735867/adaptive-behavior

Akpan, J., Notar, C., & Padgett, S. (2012). Formative assessment: A basic foundation for teaching and learning. *National Teacher Education Journal, 5*(1), 83-98.

Aschbacher, P. (1991). Performance assessment: State activity, interest, and concerns. *Applied Measurement in Education, 4*, 275-288.

Benjiman, et al. (1984). retrieved from http://www.nflrc.hawaii.edu/prodev/SI99/test-taking/multchoicetips.htm

Blackburn, B. R. (2005). *Classroom motivation from A to Z: How to engage your students in learning.* Larchmont, N.Y: Eye on Learning, 12.

Brown., J. D., & Hudson, T. H. (1998). The alternatives in language assessment. *Tesol Quarterly, 32*(4), 653-675.

Dixson, 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

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. & Schmidtt, V. L. (2007). Coming to terms with classroom assessment. *Journal of Advanced Academic, 18*, 402-423.

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., Petersen, S. E., Edwards, L. M., Pedrotti, J. T. & Peyton, V. (2003). *Toward a consensus list of item-writing rules.* Presented at the Annual Meeting of the American Educational Research Association.

Geiger, M. A. (1997). An examination of the relationship between answer changing, testwiseness, and examination performance. *The Journal of Experimental Education, 66*(1), 49-60.

Haladyna, T. M., & Downing, S. M. (1989a). A taxonomy of multiple-choice item-writing rules. *Applied Measurement in Education, 2*(1), 37-50.

Haladyna, T. M., & Downing, S. M. (1989b). Validity of a taxonomy of multiple-choice item-writing rules. *Applied Measurement in Education, 2*(1), 51-78.

Haladyna, T. M., Downing, S. M., & Rodriguez, M. C. (2002). A review of multiple-choice item-writing guidelines for classroom assessment. *Applied Measurement in Education, 15*(3), 309-334.

Herman, J. L., Aschbacher, P. R., & Winters, L. (1992). *Practical guide to alternative assessment.* Alexandria, VA: Association for Supervision and Curriculum Development.

Huerta-Macías, A. (1995). Alternative Assessment: Responses to Commonly Asked Questions. *TESOL Journal, 5* (1), 8-11.

Kizlik, B. (2014). *Measurement, assessment and evaluation in education.* Retrieved from http://www.pdfmode.com/ed/education-evaluation-pdf

Kubiszyn, T., & Borich, G. (2007). *Educational Testing and Measurement: Classroom Application and Practice.* (8th). Hoboken, NJ: John Wiley & Sons, Inc., 48-52.

Kulm, G. (1994). *Mathematics assessment: What works in the classroom*. San Francisco: Jossey Bass Publishers.

Lehr, F., Osborn, J., & Hiebert, E. H. (2004). *Research-Based Practices in Early Reading Series: A Focus on Vocabulary*. ED483190

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.

McLoughlin, J. A., & Lewis, R. B. (2008). *Assessing Students with Special Needs*, (7th Ed.). New York, NY: Pearson.

McNamara, C. (2000). *Field Guide to Nonprofit Strategic Planning and Facilitation*. Authenticity Consulting LLC, Minneapolis, MN

Meisinger, E. B., Schwanenflugel, P. J., Bradley, B. A., Kuhn, M., & Stahl, S. A. (2004). Interaction Quality during Partner Reading. *Journal of Literacy Research,36*(2), 111-140.

Mertler, C. A. (2003). *Classroom assessment: A practical guide for educators.* Los Angeles, CA: Pyrczal Publishing.

National Council of Teachers of Mathematics. (1991). retrieved from www-personal.umich.edu/~hybass/prof_standards_2-4&6.pdf

Niess, M. L., Lee, J. K., & Kajder, S. B. (2008). *Guiding Learning with Technology.* Hoboken, NJ: John Wiley & Sons, Inc. 282.

Newby, T. J., Stepich, D. A., Lehman, J. D., & Russell, J. D. (2000). *Instructional technology for teaching and learning: Designing instruction, integrating computers, and using media.* (2nd ed.). Columbus, OH: Merrill, 223.

Nitko, A. J. (2004). *Educational Assessment of Students* (4th ed.). Columbus, OH: Pearson/Merrill Prentice Hall, p. 307-308.

Notar, C. E. (2005). *Social Studies Methods.* (JSU ESE 421)

Notar, C. E. (2008). *Assessment in Education.* (JSU EFD 632).

Notar, C. E., Zuelke, D. C., Wilson, J. D., & Yunker, B. D. (2004). The table of specifications: Insuring accountability in teacher made tests. *Journal of Instructional Psychology, 31*(2), 115-129.

Pellegrino, J. W. (1999). *The evolution of educational assessment: Considering the past and imagining the future.* Educational Testing Service Policy Evaluation and Research Center, Princeton, NJ.

Regional Educational Laboratories (1998). *Improving classroom assessment: A toolkit for professional developers.* (Toolkit 98). Portland, OR: Northwest regional Educational Laboratories.

Richardson, V. (1997). Constructivist teaching and teacher education: Theory and practice. In V. Richardson (Ed.), *Constructivist teacher education: Building new understandings.* (pp. 3-14). Washington, DC: Falmer Press.

Seaman, J. A., Depauw, K. P., Morton, K. B., &Omoto, K. (2003). *Making connections: From theory to practice in adapted physical education*. Scottsdale, AZ: Holcomb Hathaway Publishing.

Smaldino, S. E., Russell, J. D., & Molenda, M. (2005). *Instructional technology and media for learning.* (8th ed.). Columbus, OH: Person, Merrill, Prentice, Hall, 26.

Stiggins, R. J. (1997). *Student-centered classroom assessment*. Portland, OR: Assessment Training Institute, Inc.

Taylor, C. S., & Nolen, S. B. (2005). *Classroom assessment: Supporting teaching and learning in real classrooms.* Columbus, OH: Pearson/Merrill Prentice, Hall, 190.

York-Barr, J., & Duke, K. (2004). What do we know about teacher leadership? Findings from two decades of scholarship. *Review of Educational Research, 74*, 255-316).