

Chapter 4

Evaluating and Assessing Student Performance

Nancy Zeliff

Northwest Missouri State University

The author, editor, and publisher acknowledge and express appreciation for the work done by Marcia Bush and Judith J. Lambrecht, authors of this chapter in the previous edition. Thank you for your contribution to this yearbook and to business education.

Evaluating and assessing student performance is an integral part of teaching. Novice teachers and those outside of education may perceive assessment to be the last step in the learning process. However, evaluating and assessing student performance must be identified and planned for in advance by the teacher in an assessment plan. This chapter explores definitions and examples of evaluation and assessment, discusses validity and reliability, and provides guidelines for developing and using varied assessments such as portfolios, rubrics, selected response test questions, and constructed response questions. How to develop an assessment plan; differentiate assessments for students with special needs, English-language learners (ELLs), and gifted learners; and use assessment data in decision making are also discussed.

PROCESS OF EVALUATION AND ASSESSMENT

Defining the varied terms used in evaluation and assessment is important, as many are used synonymously. Evaluation and assessment are “multidimensional, active processes focused on student learning and instructional improvement” (Musial, Nieminen, Thomas, & Burke, 2009, p. xvii).

Definitions

Although often used interchangeably, evaluation and assessment are different. Bush and Lambrecht (2008) stated that assessment is a more comprehensive term that includes judging, quantifying, and giving feedback. Evaluation is the judgment—or placing of a value or score—on a performance, product, or response. Musial et al. (2009) defined assessment as clarifying what a learner knows and can do. Testing and measurement, too, are often used synonymously with assessment. Linn and Gronlund (2000) and Musial et al. defined testing as the process of assessing student skills and knowledge through an

instrument administered at a fixed time under certain conditions. Measurement is a process that assigns a numerical value to a test or assessment (Linn & Gronlund).

Purposes of Assessment

Musial et al. (2009) outlined five purposes of assessment, which include providing feedback to learners, finding out what other knowledge the learner needs to learn, diagnosing learner difficulties or misconceptions, identifying what learners know when compared with other learners, and improving programs or educational units. Other purposes of assessment include motivation of students, placement of students in appropriate learning environments, evaluation of programs, assignment of grades, and teacher evaluation (Gronlund, 2006).

Validity and Reliability

Two very important elements of assessment are validity and reliability. Validity can be increased by sound construction and administration of assessment methods. Increased reliability can be achieved by the use of assessment instruments, such as rubrics and checklists.

Validity. Validity refers to the appropriateness with which an assessment measures what is intended to be measured (Gronlund, 2006). If an accounting teacher wishes to assess an accounting student on the preparation of a balance sheet, the most valid assessment is for the student to prepare a balance sheet. Musial et al. (2009) defined two types of validity: content and criterion. Content validity is the degree to which the assessment measures what was intended to measure. If the accounting student completed a selected response exam of 20 true-false questions about the balance sheet, this exam would be of low validity. A true-false exam may just measure the student's ability to guess the correct answer to these 20 questions and not the student's knowledge of preparing a balance sheet. Criterion validity measures how an assessment aligns with another assessment. For example, the accounting exam that measures a student's ability to complete a balance sheet may correlate with the advanced placement accounting exam.

Reliability. Reliability is the degree to which an assessment is stable or consistent. Test-retest reliability determines an assessment's reliability when the assessment is administered to the same students at two different times. Time between administrations can negatively affect the assessment's reliability. Too short a time may mean that the students simply recall the same assessment questions. If new information has been presented to students, the reliability of the assessment can also be adversely affected (Musial et al., 2009).

Equivalent form reliability is necessary in standardized assessments, so that multiple users taking the assessments on different dates and at different locations are assessed on the same objectives. Internal consistency reliability ensures that the assessment includes only the intended concepts and content. Herman and Baker (2005) stated that inter-rater reliability is the most recognized type of reliability and refers to the consistency of scoring students by more than one evaluator. Assessments with selected response questions are scored objectively and maintain higher inter-rater reliability, whereas assessments with constructed response questions require subjective grading and can have lower reliability (Musial et al., 2009).

Musial et al. (2009) discussed how to increase validity and reliability in assessments. Evaluators should consider the following:

- Use longer tests: the more test questions, the more opportunities students have to showcase what they know and can do.
- Write and use sound test questions, ensuring that the test questions match what was taught
- Remove unrelated questions from the test.
- Include clear and concise directions for all sections of the test.
- Use objective scoring over subjective grading to increase reliability, so evaluators have definitive right and wrong answers to use as a basis in scoring the test.

Discerning the differences among key assessment terms is important. Assessment is a more comprehensive term that includes objective and subjective assessment measures. Evaluation assigns a value or judgment on student work. Testing encompasses the use of an instrument at a fixed time, and measurement yields a numerical value assigned to an assessment.

TYPES OF ASSESSMENT

Teachers should use a variety of assessment measures to determine more accurately what a student knows and can do. Varied assessments include formal and informal; formative, interim, and summative; traditional versus alternative; authentic; performance; and norm versus criterion referenced.

Formal Assessment

Formal assessments are planned and prepared for and formalized in delivery, time, and structure. A unit exam on four chapters and topics in a personal finance class is a formal assessment. A three-minute timing in a keyboarding class is a formal assessment. Outside of the classroom, high-stakes testing such as the ACT PRAXIS, CPA, or a state-board exam for health professionals are all formal assessments.

Informal Assessment

Informal assessments can also be planned and prepared for but are less formal in delivery, time, and structure. A teacher can plan to spot-check student proficiency with formulas in an Excel spreadsheet, by asking students to “show formulas” on students’ monitors. The teacher then walks around the computer lab and sees that the correct formulas were entered or uses computer lab monitoring software to see all students’ screens from the teacher’s workstation. Another informal assessment can be the observation of keyboarding technique among keyboarding students, ensuring their hand and arm placement and posture are correct for keyboarding. Informal assessments can also be “on the fly.” At the end of a business law class, the teacher may ask students to complete a three-minute write and pose questions to the teacher that were unanswered or still remain from class discussion. This informal assessment is helpful to assess student comprehension and can lead to the next class presentation.

Formative Assessment

Formative assessment is ongoing assessment as students are learning concepts and skills. Pinchok and Brandt (2009) identified that “formative assessment is a process in which teachers use various tools and strategies to determine what students know, identify gaps in understanding, and plan future instruction to improve learning” (p. 2). Gronlund (2006) stated that formative assessments measure learning outcomes from a segment of instruction, such as a textbook chapter or instructional unit. Any form of assessment can be used in formative assessment, but the assessment should align with objectives and classroom activities. Timely feedback from these assessments is essential to help the student achieve.

Summative Assessment

Summative assessments occur at the end and are a summary or assessment of the end result. The most common instructional summative assessment is a comprehensive final exam at the end of a course or term. Lambrecht (2000) identified that high-stakes testing often occurs at the end of courses and programs. If the results of summative assessments are positive, certification, graduation, admittance, or licensure often results. If the results of summative assessments are negative, then remediation may be the result. Students are the most common recipient of summative assessments; however, school districts and colleges and universities undergo accreditation, which is also a summative assessment. If accreditation is not achieved as a result of the summative assessments, the programs or organizations must enact change and improvements.

Traditional Assessment

Traditional assessment is identified by assessment measures that are commonplace and have been used perhaps for generations within particular curriculum areas. Objective tests consisting of multiple

choice or true-false questions are common assessments for assessing the lower levels of the cognitive domain or a learner's knowledge about a particular subject matter. Spelling exams are common assessments to assess a learner's language arts skills. Speed typing tests are traditional assessments for a keyboarding student. Traditional assessment is appropriate because of its high validity; however, alternative assessments should also be used.

Alternative Assessment

Alternative assessments are often used to assess a learner's affective domain and the higher levels of the cognitive domain. For example, it is better to assess the soft skills or human relations skills through observations by supervisors or teachers, rather than by students taking a paper/pencil exam. A student's behavior and attitude can be assessed by anecdotal records and observations. Peer and self-assessment measures are helpful in identifying skills that need improvement or those at which a student excels.

Another type of alternative assessment is performance assessment, which more readily measures multiple domains of learning. Tasks are performed in order to measure the student's skills and competence. Musial et al. (2009) stated that performance assessments give teachers "the opportunity to evaluate a cognitive skill while it is being performed" (p. 205). Mock job interviews are an excellent means to measure a student's interviewing skills and communication skills when the student is "performing" a simulated job interview. Although objective or essay exam questions could address job interview skills on a paper/pencil exam, observing and using a checklist to assess job interview skills is more valid and more reliable. Exam questions that require written or essay responses are considered performance assessments, as well as essays, compositions, and written work.

Authentic Assessment

Authentic assessments require students to complete real-world tasks and activities, rather than contrived projects from workbooks and textbooks. Musial et al. (2009) defined authentic tasks as those similar to what professionals in a ~~that~~ discipline ~~daily~~ perform daily.

Norm-Referenced Assessment

Norm-referenced assessments interpret student performance and provide a relative ranking of students (Gronlund, 2006). Standardized examinations that are summative in nature provide norm-referenced results and include the PRAXIS, GRE, ACT, and MFAT. Administrators use the rankings of students from one school district compared with other students as an indicator of achievement and progress by both the student and the program or institution. Norm-referenced assessment is typically used in districtwide or state-level assessments (Lambrech, 2000).

Criterion-Referenced Assessment

In contrast to norm-referenced assessment, Gronlund (2006) stated that criterion-referenced assessment interprets assessment results or what an individual student can do without reference to the performance of other students. Lambrecht (2000) stated that with criterion-referenced assessments, students' results are compared with standards and learning outcomes. Criterion-referenced assessments are typically used for mastery exams.

METHODS OF ASSESSMENT

This section discusses construction suggestions and the advantages and disadvantages of two methods of assessment: traditional paper/pencil exams and performance-based assessments. Also discussed are tools for evaluation, including checklists and rubrics, which add reliability to the assessment process; developing an assessment plan; and assigning course grades.

Traditional Paper/Pencil Exams

Traditional paper/pencil exams are those that include selected and constructed response questions. Objective types of exams, also called selected response questions, fall into this category.

Selected response questions. Selected response questions are objective types of questions that include multiple-choice, true-false, and matching. Students “select” their response from given options of answers. The advantages of selected response questions include the ease and speed in providing formative feedback to the student, parents, and other stakeholders; ease in grading; and ability to compare scores among students, classes, and schools taking the same assessment. Disadvantages of selected response questions include the lack of insight into students' reasoning and learning and the chances that students are guessing the correct answers (Musial et al., 2009).

Multiple-choice questions. Multiple-choice questions can measure the lower levels through the higher levels of the cognitive domain and are the most commonly used type of exam question, especially on standardized tests. These questions are not as easy to construct as other types of questions but are easy to score. Students are less prone to “guessing” with multiple-choice than true-false questions. Good multiple-choice questions should include (a) a stem, which presents an incomplete statement or question, (b) a single correct answer, (c) alternatives to the correct answer, and (d) distracters. Developing the question (stem) and the right answer are relatively easy; the difficult part is writing the alternatives and meaningful distracters so the question is valid and measures what it intends to measure. When writing the alternative answers, one should be clear on the relationship between the alternatives and the question within the stem. When composing distracters, it is important to ensure they are sufficiently incorrect so

prepared students can identify which answer is wrong. Well-written distracters should be plausible to students who did not prepare adequately for the exam (Musial et al., 2009).

Apply caution in using “all of the above” and “none of the above.” When used only because not enough choices of answers were provided, students will recognize this overuse and discount these options as a correct answer. Some students may find the correct answer among the first alternatives and not read further to select these options. When using “all of the above,” be sure that all answers are indeed correct. Furthermore, avoid indirect clues when writing multiple-choice questions. Students will readily identify the correct answer as the most lengthy answer, so use parallel format and length of all choices (Musial et al., 2009).

True-false questions. True-false questions also measure varied levels of the cognitive domain. They are easy to construct and easy to score. When writing sound true-false questions, educators should use important concepts and facts, rather than trivial information. Correct answers must be distinctly true or false and should not include obvious clues. Furthermore, verbatim textbook phrases should not be used in true-false questions but rather more global and factual information (Musial et al., 2009).

Matching questions. Matching items require the learner to identify common characteristics between two or more sets of items, usually vocabulary. Advantages of using matching questions include the ease of creation and scoring and the ability to assess a broad conceptual base. Disadvantages of matching questions stem from the chance of guessing correct answers by the test taker and that these questions can only assess recall of information, not higher-order thinking. When constructing sound matching items, “offering an unequal number of items in [your] two columns is an effective way to reduce the possibility of answering correctly through the process of elimination” (Musial et al, 2009, p. 131).

Constructed response questions. Constructed response questions include questions on exams in which the learner must construct a response. Common examples are short answer and essay.

Short-answer questions. Short-answer questions require the learner to supply the answer, unlike true-false and multiple-choice questions, for which students select an answer from those provided. Short-answer questions should be clear and unambiguous, because students are expected to recall information and remember the context in which the information was presented, read, or studied. Advantages of short-answer questions are that they adequately measure vocabulary and the language of a discipline and are a quick formative type of assessment. A disadvantage is that these questions primarily measure recall and can be time consuming to grade, because of the wide range of potential answers (Musial et al., 2009).

Essay questions. Essay questions allow the student more choices in constructing their answers. Essay questions assess higher-order cognitive domains and enhance a student's exam preparation and study skills. Essay questions are also fairly easy to construct. Two disadvantages to the use of essay questions include the amount of time needed to score the responses and the potential bias or lack of objectivity in grading essay exams (Musial et al., 2009).

Two types of essay questions are the restricted-response essay and the extended-response essay. The restricted essay exam restricts or limits the student response. Recollection of facts and the ability to summarize and defend one's response is often necessary. The extended-response essay question may not have a definitive right or wrong answer but relies on "evaluation, organization, analytical reasoning, or originality" (Musial et al., 2009, p. 145).

Publisher test banks. Publishing companies provide ancillary materials, which include test banks to accompany textbooks. Advantages in using publisher test banks are the large number of questions from which to choose and the ease of use, especially for electronic test banks. The disadvantage of using publisher test banks is often the lack of alignment with standards and objectives of the course (Bush & Lambrecht, 2008).

Teacher-prepared test questions. Teacher developed exams and test questions are advantageous because they are written with specific standards and objectives of the course in mind. Using teacher-prepared test questions can, however, reduce the validity of an assessment if the items are poorly constructed. The use of publishers' test banks can increase the validity of the assessment, if the teacher wisely selects exam questions from those provided in the test bank (Bush & Lambrecht, 2008).

Item analysis. Item analysis enhances the likelihood of better teacher-made tests. In item analysis, teachers use procedures to evaluate the quality of questions used on an assessment. Individual questions are reviewed on item difficulty—the "ratio or percentage of individuals who answered an item correctly" (Musial et al., 2009, p. 322). Computing an item analysis is time consuming; however, online assessments and computerized testing software can compute the analysis automatically, and this analysis can lead to more valid and reliable assessments.

Performance-Based Assessments

Performance-based assessments include those assessments for which the student performs a task, completes a project or written work, or gives an oral performance. To clarify the expectations of these varied assessments to the student and to ensure higher reliability, checklists or rubrics should be used in scoring these performance-based assessments.

Presentations. Presentations by students are an excellent way to assess content-specific knowledge, as well as technology skills and oral communication skills. Presentations can be an extension of case studies, projects, and portfolios. Students not only prepare these performance-based assessments but present information about their work to their peers and teacher (Bush & Lambrecht, 2008).

Written work. Students complete written work through varied assessments. Case studies, research papers, article reviews, blogs, and journals are examples of written assessments. Written work can assess communication skills, personal reflections, and content knowledge. Use of rubrics can minimize bias and subjectivity and increase reliability in the use of written assessments.

Portfolios. Portfolios show student learning to a degree that not many assessments can. Musial et al. (2009) stated that portfolios can reveal a great deal about students and allow students to assume ownership in their assessment. Fernsten and Fernsten (2005) stressed that the key components of portfolios are the reflections written by the students. Students analyze their achievement of the course objectives and evaluate their work and growth through these written reflections.

Portfolios also provide a means for students to showcase their creativity and organizational skills. Portfolios can be time intensive for the student; however, they have a direct carryover to the workplace as the portfolio can be used in the job search process and shared with prospective employers (Bush & Lambrecht, 2008).

Tools for Assessment

Use of checklists and rubrics can increase the inter-rater reliability of assessments. These tools of assessment clearly define the criteria that will be assessed for both the student and teacher.

Checklists. A checklist includes a list of criteria to be met or applied in completing an exercise or learning activity. Bush and Timms (2000) stated that checklists are an excellent organizer for the student, who can check that all components of the assignment are completed. Whether the checklist is a simple list or a more complex matrix of components of a presentation or project, the teacher can check off components that are present. Figure 1 provides an example of checklist for a Windows Movie Maker project in a digital media class.

Figure 1. Digital Media: Windows Movie Maker Project

Project component	Criteria present (check if present)
Introductory title at beginning of movie	
Video clip present	
5–8 digital pictures in .jpg format	
Transitions and effects as appropriate	

Title for each picture: on or before/after picture	
Voice recording: good quality and volume	
Music clip present	
Credits at end	
Sufficient timing to view/read/hear	
Color scheme appropriate	
Copyright laws followed	
Effective and entertaining	
Comments:	

Rubrics. A rubric or scoring guide takes the criteria from the checklist even further. Musial et al. (2009) described a rubric as a scoring guide with specific performance levels used to assess a research paper, project, or performance. Two specific types of rubrics are useful in assessing student performance: analytic and holistic.

Analytic rubrics identify specific knowledge and skills that are expected in a performance or product. The criteria are assigned performance levels and weighted separately. Specific feedback on each of the criteria provides the learner very specific information. A final score is the total of all the criteria. Analytic rubrics are very effective with formative assessments (Bush & Timms, 2000).

Holistic rubrics are used to assess the performance or product as a “whole” and judge the overall quality of that performance or product. Rather than list separate criteria and expected performance levels, three to four written statements are made to describe “exceptional” to “inadequate” performances or products (Musial et al., 2009). Bush and Lambrecht (2008) indicated that holistic rubrics are useful as summative assessments.

The first step in developing a rubric is writing the descriptive criteria. The second step is to determine how many performance levels are needed: three or four performance levels work best. If more than four levels are used, differentiating among the criteria is more difficult. The third step in rubric development is describing the criteria for the various performance levels. Begin with the highest performance level, and use action verbs to describe the performance or appearance. Avoid using nonspecific adverbs (i.e., nearly or occasionally) and differentiating performance levels by quantity (i.e., four of five components are present). The final step is the assignment of a value or points for each performance level and criterion. Points assigned can be equal or weighted, depending on the importance of the criteria in the overall performance or product (Zeliff & Schultz, 1998). Many online interactive tools are also available. An example of an analytic rubric for creating a logo in Photoshop can be found online at

http://rubistar.4teachers.org/index.php?screen=PrintRubricDownloadFile&rubric_id=2206693& (Zeliff, 2012). Figure 2 provides an example of rubric for a portfolio made with Adobe Acrobat Professional in a multimedia class.

Figure 2. Digital Media Electronic Portfolio

	Commendable	Acceptable	Not acceptable
Presentation and content	Designed professionally. Layout design used represents portfolio well. No misspelled words. Content of portfolio should be helpful to owner in job search and helpful to prospective employer to see skills of applicants. (5 points)	Layout and design acceptable. No misspelled words. Portfolio could be used in job search. (2–4 points)	Poor layout and design. One or more misspelled words. Do not recommend using this portfolio in the job search. (0–1 points)
Required items <ul style="list-style-type: none"> • Introduction • Reflection • Audio folder with audio file • Video folder with video file 	Purpose of portfolio evident. Text-based introduction present. Two required folders are present. All other required items present. (8–10 points)	Purpose of portfolio clear. Text-based introduction present but not strong. One required folder missing. One other required item missing. (4–7 points)	Required items missing. Not clear that portfolio would be used in future by owner. No introduction. (0–3 points)
Video and audio files	Audio recording in mp3 and video in .swf format present. Audio recording clear, inviting, with good volume and voice. (5 points)	Audio and/or video formats not correct. Audio recording not of best quality for voice, volume, clarity, and message. (2–4 points)	Either video or audio file are not present. Audio recording of poor quality and distracts from portfolio. (0–1 points)
Overall effect	Pride in work evident; exhibits technical skill in Adobe Acrobat. Color scheme, fonts, and layout effectively present portfolio in professional manner. (5 points)	Pride in work is evident; Adobe Acrobat skills apparent. Color scheme, fonts, and layout are appropriate. (2–4 points)	Evidence of little pride in work completed. Color scheme, fonts, and layout used are not appropriate for professional portfolio and may make “reading” portfolio difficult. (0–1 points)

Comments:

+ _____ /25 points for Portfolio

Advantages of using rubrics are many. Communication of expectations is clear between teachers and students, and the rubrics provide specific feedback to students. Reliability and objectivity in grading are increased because the teacher clearly sees both the criteria and the varied performance levels to be met. Time spent by the teacher in grading is also reduced, because the rubric is an organizer of what is to be graded and already includes printed feedback for the student (Bush & Timms, 2000).

Developing an Assessment Plan

An assessment plan includes the types of assessments and the grading system that will be used. The plan covers course objectives; a crosswalk of those objectives to national, state, or industry standards; and the specific assessment that will measure each objective. The plan should also include an example of each assessment used, which can include tests, portfolios, projects, modules, performances, case studies, interviews, observations, student self-assessment, and peer assessment. Specific checklists or rubrics used to score the varied assessments are also included in the assessment plan. The grading system used must also be discussed in an assessment plan. The responsibility belongs to the teacher to communicate clearly the grading policy to students. ~~students enrolled in a course.~~

ASSESSMENTS FOR EXCEPTIONAL STUDENTS

Students with exceptionalities include students with special needs, ELL students, and gifted students; these students may require accommodations to assessments in the location, time, format, and type of assessment given them. Public and private school districts and colleges and universities vary in their policies regarding the modification of assessments for students with exceptionalities. State and federal laws and regulations influence the accommodations that must be granted.

Special Needs Students

Special needs students who require assessment accommodations may have an individualized education plan or assessment accommodation plan. Elliott, Thurlow, Ysseldyke, and Erickson (1997) described several recommendations for assessment accommodations that may be necessary:

- Modify the *setting* in which the student takes the assessment to create an environment free from noise and other distractions.

- Alter the *presentation* of the assessment to a manner that fits the needs of the student.
- Read the assessment to the student or translate it into another language.
- Make *time* adjustments to an assessment as needed.
- Modify the *response* that the student provides to an assessment.
- Computerize the assessment so that a pencil/pen does not have to be used or vice versa.

Furthermore, the *scheduling* of the assessment is a factor. For students with documented test anxiety, accommodations are made to reduce that anxiety. Students who have medications given at a particular time of day may be scheduled to take exams at certain times of day that are better for them due to their medication schedule.

English-Language Learners

ELL students are not disabled but have limited language proficiency and have different cultural backgrounds and educational experiences that may affect their performance on assessments. The following instructional elements addressing the needs of ELL students should be considered:

- Incorporate diversity into classroom instruction by recognizing and validating cultural perspectives.
- Ensure all students, including ELL students, understand the assessment directions.
- Enact structured rules for the classroom and assessments so all students know the teacher's expectations during assessments (Elliott, et al., 1997).

Gifted Students

Unlike students with special needs, students with exceptional gifts and talents (the gifted) are not included in federal legislation for special education. Not all states and school districts include gifted students in their accommodations. However, teachers need to assess gifted students fairly. Modifications in assessments for gifted students should not include more of the same problems or assignments that nongifted students receive. Rather, Lidz and Elliott (2006) encouraged the use of dynamic assessment, authentic assessment, and problem-based assessment for gifted students. An example of a dynamic assessment in personal finance would include a unit exam of selected and constructed response questions about savings and investments. Following the initial exam, additional instruction would be provided to the gifted student in areas that need improvement, and a follow-up assessment would then take place. Authentic and problem-based assessments should be offered to all students but are best suited for gifted students.

USING ASSESSMENT DATA FOR IMPROVEMENT

Teachers are largely held accountable for the learning of their students and should strive to help them to continually improve. One way to demonstrate this accountability is through the use of assessment data to improve instruction and programs. Teachers should carefully review student assessment data and make decisions about the improvements to be made. They should implement, monitor, and continually adjust the improvements to ensure continued learning gains.

Data from criterion-referenced exams and teacher-constructed assessments provide those teachers valuable information about the specific knowledge and skills students have obtained through instruction. The use of pre-tests and post-tests in units of instruction can reveal to teachers that specific objectives have or have not been met (White, 2007).

Assessments provide feedback to students on their progress toward meeting specific learning objectives of a course or program and provide teachers with feedback on their teaching effectiveness. Data from the varied assessments are useful to educational stakeholders as evidence of effective instruction in courses and programs. In addition, these data are critical to the continued improvement of program effectiveness and student learning.

SUMMARY

Evaluating and assessing student performance is not the end result of teaching but a major component that must be planned for and organized. Teachers should plan for multiple assessments, including a variety of types, methods, and tools for each business education course. Alternative, authentic, and soundly constructed performance assessments that are both valid and reliable are necessary to determine if students meet the course or program objectives. Checklists and rubrics used in assessing performances and products help improve the reliability of those assessments. Formative and summative assessments, both formal and informal, given during and after instruction provide valuable feedback to not only students, but also teachers and administrators. Teachers can use data from these assessments for instructional improvements and other decision making.

REFERENCES

- Bush, M., & Lambrecht, J. (2008). Evaluating and assessing student performance. In M. Rader, G. A. Bailey, and L. A. Kurth (Eds.), *Effective methods of teaching business education: NBEA yearbook* (Vol. 46, pp. 80–93). Reston, VA: National Business Education Association.

- Bush, M., & Timms, M. (2000). Rubric- and portfolio-based assessment: Focusing on student progress. In J. Rucker (Ed.), *Assessment in business education: NBEA yearbook* (Vol. 38, pp. 103–120). Reston, VA: National Business Education Association
- Elliott, J., Thurlow, M., Ysseldyke, J., Erickson, R., & National Center on Educational Outcomes. (1997). *Providing assessment accommodations for students with disabilities in state and district assessments*. NCEO policy directions. Retrieved from ERIC Document Database (ED416628).
- Fernsten, L., & Fernsten, J. (2005). Portfolio assessment and reflection: Enhancing learning through effective practice. *Reflective Practice*, 6(2), 303–309. doi:10.1080/14623940500106542
- Gronlund, N. E. (2006). *Assessment of student achievement* (8th ed.). Boston, MA: Pearson Education Inc.
- Herman, J. L., & Baker, E. L. (2005, November). Making benchmark testing work. *Educational Leadership* 63, 48–54. Retrieved from http://www.kckps.org/teach_learn/pdf/group3/t_112_making.pdf
- Lambrech, J. J. (2000). Characteristics of good assessment. In J. Rucker (Ed.), *Assessment in business education: NBEA yearbook* (Vol. 38, pp. 25–38). Reston, VA: National Business Education Association.
- Lidz, C. S., & Elliott, J. G. (2006). Use of dynamic assessment with gifted students. *Gifted Education International*, 6(2–3), 151–161. doi:10.1177/026142940602100307
- Linn, R. E., & Gronlund, N. E. (2000). *Measurement and assessment in teaching* (8th ed.). Upper Saddle River, NJ: Prentice-Hall, Inc.
- Musial, D., Nieminen, G., Thomas, J., & Burke, K. (2009) *Foundations of meaningful educational assessment*. New York, NY: McGraw-Hill.
- Pinchok, N., & Brandt, W. (Learning Point Associates). (2009). Connecting formative assessment research to practice: An introductory guide for educators.. Retrieved from ERIC Database (ED509943).
- White, R. (2007). Data-driven instruction. In M. Bush and K. Schultz (Eds.), *Assessment for an evolving business education curriculum: NBEA yearbook* (Vol. 45, pp. 117–129). Reston, VA: National Business Education Association.

Zeliff, N. (2012). 44-626 Photoshop Logo. Retrieved from

http://rubistar.4teachers.org/index.php?screen=PrintRubricDownloadFile&rubric_id=2206693&.

Zeliff, N., & Schultz, K. (1998). *Authentic assessment in action: Preparing for the business workplace*.

Little Rock, AR: Delta Pi Epsilon.