Assessment at a distance: Traditional vs. Alternative Assessments

Semire Dikli
Florida State University

"Assessment performances are day-to-day activities that can also be authentic and engaging demonstrations of students' abilities to grapple with the central challenges of a discipline in real life contexts" (Kulieke, Bakker, Collins, Fennimore, Fine, Herman, Jones, Raack, & Tinzmann, 1990, p.2).

Assessment is one of the crucial components of the instruction. People within the educational community, i.e. policymakers, educators, students, parents, administrators, have different ideas regarding the implementation of assessment strategies (Dietel, Herman, and Knuth, 1991). While some believe traditional assessment methods are more effective, others think that alternative assessment tools are superior. This article is written to inform people particularly in the field of distance education about assessment practices at a distance. However, the content of assessment is not a field specific and it can be applied to various instructional settings (Simonson, Smaldino, Albright, and Zvacek, 2000). Therefore, anyone who is directly or indirectly related to education –distance or face-to-face- might find the information presented in this article useful.

Assessment and testing

Assessment and testing considerably differ from each other. While testing is formal and often standardized, assessment is based on a collection of information about what students know and what they are able to do. In other words, students are given the exact procedures for administering and scoring in testing. In assessment, on the other hand, there are multiple ways and methods of collecting information at different times and contexts (Law and Eckes,1995, p.29).

Dietel, Herman, and Knuth (1991) define assessment as "any method used to better understand the current knowledge that a student possesses" (online document). According to Mitchell (1992, in Law and Eckes, 1995, p. 29) testing can be defined as "single-occasion, unidimensional, timed exercise, usually in multiple choice or short-answer form." For a long time, student learning was measured only by testing in traditional school settings. Currently, it is realized that there is not only one way of gathering information about student learning. Furthermore, testing is seen as only one part of assessment and a broader concept of assessment is being widely used (Kulieke, Bakker, Collins, Fennimore, Fine, Herman, Jones, Raack, and Tinzman, 1990).

Traditional assessment tools:

The most widely used traditional assessment tools are multiple-choice tests, true/false tests, short answers, and essays.

True/false tests: True/false items require students to make a decision and find out which of two potential responses is true. Since they are easy to score, it is easy to administer true/false tests. However, guessing might increase the chance of success by 50%. Especially, when the test item is false, it is quite hard to find out whether the student really knows the correct response. One possible solution is to ask student to provide with an explanation for the incorrect item, or rewrite the statement correctly. However, this affects the ease in scoring negatively (Simonson et al., 2000).

Multiple-choice tests: Multiple-choice tests are commonly utilized by teachers, schools, and assessment organizations for the following reasons (Bailey, 1998, p. 130):

1. They are fast, easy, and economical to score. In fact, they are machine scorable.
2. They can be scored objectively and thus may give the test appearance of being fairer and/or more reliable than subjectively scored tests.
3. They "look like" tests and may thus seem to be acceptable by convention.
4. They reduce the chances of learners guessing the correct items in comparison to true-false items.

Simonson and others discussed the disadvantages of multiple choice tests. They claimed that depending on the level of cognitive effort, they become harder and more time consuming to create. In other words, multiple choice items can be used effectively in testing the items that demand low level of cognitive effort such as recalling previously memorized knowledge, yet items that require students to use higher order thinking skills such as analyzing and synthesizing are more difficult to produce (2000). Similarly, Hughes (in Bailey, 1998) criticizes multiple-choice tests for the following aspects:

"1. the technique test only recognition knowledge, 2.guessing may have a considerable but unknown effect on the test scores, 3.the technique severely restricts what can be tested, 4 it is very difficult to write successful items, 5. backwash maybe harmful, 6. cheating may be facilitated " (p.131).

Essays: Essays are effective assessment tools since the questions are flexible and assess the higher order learning skills. However, they are not very practical due to the fact that it is very difficult and time consuming to score the essays. Moreover, subjectivity might be an issue in scoring. Creating a rubric might be helpful to grade the essays (Simonson et al., 2000). A rubric can be defined as "a criteria-rating scale, which gives the teachers a tool that allows them to track student performance" (Abrenica, online document). Instructors have an option to create, adapt, or adopt rubrics depending on their instructional needs. The templates provided on the web might be helpful for them to adjust the generic rubrics into their own instruction (Simonson et al., 2000).

Short-answer tests: In short-answer tests "items are written either as a direct question requiring the learner fill in a word or phrase or as statements in which a space has been left blank for a brief written answer" (Simonson et al., 2000, p. 270). Furthermore, the questions need to be precise. Otherwise, the items that are open to interpretations allow learners to fill in the blanks with any possible information (Simonson et al., 2000).

Alternative assessment tools:

According to Simonson and others, there are three approaches in alternative assessment: Authentic assessment, performance-based assessment, and constructivist assessment. Similarly, Reeves (2000) suggests three main strategies to integrate alternative assessment into online learning settings: 1. cognitive assessment, 2. performance assessment, 3. portfolio assessment. Researchers and educators use the term performance-based, alternative, and authentic assessment interchangeably. As Wangsatorntanakhun (1997) states the term, performance-based assessment, embraces both alternative and authentic assessment. Therefore, throughout this article, performance assessment is used to refer to alternative assessment.

There are two major concepts that describe performance assessment: "1. Performance: A student's active generation of a response that is observable either directly or indirectly via a permanent product, 2. Authentic: The nature of the task and context in which the assessment occurs is relevant and represents "real world" problems or issues" (Elliott, 1995). Authentic assessment aims to relate the instruction to the real-world experience of the learners. The task needs to be meaningful in order to be authentic (Simonson et al.). Winking (1997) also points out the role of authenticity and states that alternative assessments require higher order thinking skills so that students can solve real-life related problems. Finally, Bailey (1998) relates the power of the performance tests are not only to their authenticity, but also to their direct and highly contextualized nature.

In order to increase the effectiveness of performance assessment, instructors need to pay attention to the following points (Elliott, 1995):

1. Selecting assessment tasks that are clearly aligned or connected to what has been taught.
2. Sharing the scoring criteria for the assessment task with students prior to working on the task.
3. Providing students with clear statements of standards and/or several models of acceptable performances before they attempt a task.
4. Encouraging students to complete self-assessments of their performances.
5. Interpreting students' performances by comparing them to standards that are developmentally appropriate, as well as to other students' performances (online document).

Alternative assessment strategies include open-ended questions, exhibits, demonstrations, hands-on execution of experiments, computer simulations, and portfolios (Dietel et al., 1991). The two common alternative assessment techniques, portfolios and projects, are discussed below.

Portfolios: Portfolios consist of student work that displays mastery of skill of the task and expression (Kulieke et al., 1990). Paulson, Paulson, and Meyer (in Bailey, 1998) define portfolios as " a purposeful collection of student work that exhibits the student's efforts, progress, and achievements in one or more areas. The collection must include student participation in selecting contents, the criteria for judging merit, and evidence of student self reflection" (p. 216). Because of their cumulative nature, portfolios require a lot of input and responsibility of from the student. Moreover, they demand a great deal of time commitment from the teachers, which yields a practicality problem in assessment (Bailey, 1998).

The benefits of portfolios are pointed out by Arter (1995). Students will:

a. Get a broader, more in-depth look at what students know and can do.

b. Base assessment on a more 'authentic' work.
c. Have a supplement or alternative to report cards and standardized tests.
d. Have a better way to communicate student progress to parents (online document).

One application of portfolio use at a distance is electronic portfolios. An electronic portfolio is "a technology-based form of authentic student-based assessment" (Abrenica, online document). It functions same as a traditional portfolio. The only difference is that the further is technology based. Unlike traditional portfolios, electronic portfolios can take up little space since information can be stored in a computer hard drive, a floppy disc, or a CD rom. The practicality of e-portfolio use is highly dependent on instructor's as well as learner's knowledge of computer technology. The variety of information that could be included to e-portfolios is infinite. It is helpful to use rubrics to assess the quality of the work (Abrenica, online document).

Projects: Projects can be created individually or as a group. They can possess authenticity, real life related concepts as well as prior experience of the learners. Any type of method that display what student know about a specific topic, i.e. development of plans, art work, research proposals, multimedia presentations, is considered as project. Problem-based learning requires learners to use their problem solving skills to respond to a given situation. For instance, they can be presented a scenario and asked to provide strategies or solutions. The task is assigned to either individuals or groups. They present with the findings they come up with in various forms, such as multimedia presentation, role-play, and written report (Simonson et al., 2000).

Traditional assessments vs. Alternative assessments

There has been a movement from traditional assessment toward alternative assessments. Alternative assessment started being used as a means for educational reform due to the increasing awareness of the influence of testing on curriculum and instruction (Dietel, Herman, and Knuth, 1991). Similarly, Reeves stated that traditional assessment, which is generally called testing, is challenged by alternative assessment approaches (2000, p. 103).

According to Bailey (1998), traditional assessments are indirect and inauthentic. She also adds that traditional assessment is standardized and for that reason, they are one-shot, speed-based, and norm-referenced. Law and Eckes (1995) underline the same issue and state that traditional assessments are single-occasion tests. That is, they measure what learners can do at a particular time. However, test scores cannot tell about the progression of child. Similarly, they cannot tell what particular difficulties the students had during the test. Bailey (1998) also mentions that there is no feedback provided to learners in this type of assessment. The projects are mainly individualized and the assessment procedure is decontextualized. Law and Eckes (1995) point out most standardized tests assess only the lower-order thinking skills of the learner. Similarly, Smaldino et al. (2000) state that traditional assessment often focus on learner's ability of memorization and recall, which are lower level of cognition skills. Additionally, traditional assessment tools require learners to display their knowledge in a predetermined way (Brualdi, 1996).

Alternative assessments, on the other hand, assess higher-order thinking skills. Students have the opportunity to demonstrate what they learned. This type of assessment tools focus on the growth and the performance of the student. That is, if a learner fails to perform a given task at a particular time, s/he still has the opportunity to demonstrate his/her ability at a different time and different situation. Since alternative assessment is developed in context and over time, the teacher has a chance to measure the strengths and weaknesses of the student in a variety of areas and situations (Law and Eckes, 1995)).

More authentic assessment tools, such as portfolios, independent projects, journals and so on, let learners express their knowledge on the material in their own ways using various intelligences (Brualdi,1996). According to Gardner, there are eight intelligences (Brualdi): "1.logical-mathematical intelligence, 2.linguistic intelligence, 3. spatial intelligence, 4.musical intelligence, 5.bodily-kinesthetic intelligence, 6.the personal intelligences: a. interpersonal intelligence, b.intrapersonal intelligence, 7. naturalistic intelligence" (1996, online document).

Reeves (2000) believes the emphasis on performance assessment is the ability of learner in applying his/her knowledge and skills to real life simulations. He further states that there are five main points in performance assessment (p. 108): "1. It is focused on complex learning, 2. engages higher-order thinking and problem solving skills, 3. stimulates a wide range of active responses, 4. involves challenging tasks that require multiple steps, 5. requires significant commitments of student time and effort." Similarly, Simonson and others (2000) discuss the several advantages of alternative assessment. First of all, they tend to simulate real-life

contexts. Learners have opportunity to practice the authentic activities that they might encounter in real life. These activities allow them to transfer their skills to various real world related settings. Second, collaborative working is encouraged. Finally, alternative assessments assist instructors to have a better understanding of student learning (Winking, 1997). That is, looking at the student product rather than scores can allow instructor to get further insights regarding students' knowledge and skills (Niguidila, 1993).

Bailey (1998) contrasted traditional and alternative assessment (p. 207):
One-shot tests → Continuous, longitudinal assessment
Indirect tests → Direct tests
Inauthentic tests → Authentic tests
Individual projects → Group projects
No feedback provided to learners → Feedback provided to learners
Speeded exams → Untimed exams
Decontextualized test tasks → Contextualized test tasks
Norm-referenced score interpretation → Criterion-referenced score interpretation
Standardized tests → Classroom-based tests.

According to the information provided above, traditional assessments seem to have no positive characteristics at all. However, this is not true. There are advantages of traditional tests just like there are disadvantages of alternative tests. To begin with, traditional assessment strategies are more objective, reliable and valid. This is especially true for standardized tests and other types of multiple choice tests (Law and Eckes, 1995). Alternative assessments, on the other hand, carry some concerns in terms of subjectivity, reliability and validity. Ecke and Law express their concerns by stating " coaching or not coaching, making allowances, or giving credit where credit is not due are critical issues that have yet to be addressed; we simply do not have answers yet" (1995, p.47). While Bailey (1998) agrees with Law and Ecke about the reliability issue, she argues about the high validity in alternative assessments. She gives the portfolio example and claims that the wide variety in student products might cause reliability problems. However, the positive washback they provide to the learner as well as validity let portfolios be a widely used effective assessment tool (1998). Similarly, Simonson et al. claim that "proponents of alternative assessment suggest that the content validity of "authentic" tasks is ensured because there is a direct link between the expected behavior and the ultimate goal of skill/learning transfer" (2000, p. 275).

As Law and Ecke (1995) mention, alternative assessments can be laborious in terms of time and energy spent by the teacher. For example, the diversity of products in portfolios, which is viewed as one of the most important strengths, can lead problems for the teacher in terms of practicality (Bailey, 1998). They might be harder to score and quite time consuming to evaluate the learner's performance (Simonson et al., 2000). Rentz (1997) claims that unlike multiple-choice tests, which are practical to score, performance assessments are viewed quite time consuming to grade. While the firstr is machine scorable, the latter relies on human judgment.

Assessment and distance education

As mentioned previously, the content of assessment can be applied to any instructional setting, no matter the instruction is given at a distance or face-to-face. Nevertheless, there are some essentials of assessment, which are particularly important in distance education (Simonson et al., 2000). Nouwens and Towers (1997) claim that the assessment strategy to be employed is determined by the delivery media, resources and the time available. Similarly, Jones (2002) underlines that the media for providing tests is heavily dependent on the availability and the accessibility of the resources to the distance instructor. Some of the assessment strategies used in distance education are the following:

- individual works developed individually and sent by regular mail or by email
- assessment based on contributions for group discussions
- tests (automatically handled by computer program)
- term papers (analyzed by professor or assistants)
- oral or written tests conduced in the presence of the instructor (some times through videoconference) or with a remote assistant (Tarouco and Hack, online document).

Jones (2002) argues about two types of assessment options, distributed and on-line, in distance education. Distributed assessment option requires distance learners to use specific software, which can be downloaded from the Internet or mailed on a floppy, CD-ROM etc.; whereas, on-line assessment option occurs by directing the computer browser to the given web page and no installment takes place. Additionally, Tarouco

and Hack (online document) discuss some systems in the market that give information on the progress of the distant student. The examples of this type of system mediated by Internet are CyberQ, WebCT, and AulaNet, which include testing tools such as multiple choice quiz and term papers.

With the rapid improvement in technology, large-scale testing has become quite common. Among computer-based tests are GRE, SAT I: Reasoning test, TOEFL (Bennett, 1997). One of the benefits of the computer-based tests is the learner receives immediate feedback. Bennett further discusses how computer-based tests are implemented:

The computer selects questions based in part on previous responses, tailoring the test to individual skill levels. Depending on the testing program, individuals can register by phone or e-mail; pay by credit card; test by appointment in a relatively small, comfortable center; and receive the scores at the conclusion of the session.

Testing organizations can electronically exchange questions and examinee responses with test centers, and send scores to institutions in the same way (1997, p.3).

Reeves (2000) highlights the distinct nature of assessment in traditional classes compared to its embedded nature in online environments. In traditional learning environments students are usually assessed after they completed certain number of classes in a semester, i.e. mid-term exams and finals in essay, short answer or/and multiple choice formats. In online learning environments, on the other hand, assessment and instruction are integrated through interactive media simulations. For example, in a web-based simulation program, learners can encounter several problems where the program provides them with feedback. Here, while the computer finds out the improvement in the student performance, it provides the instructor with performance assessment data as well.

There are several factors that must be taken into consideration in designing as well as implementing assessment procedures in distance education. Simonson and others (2000) underline the essential role of fairness of the assessment activity. They suggest that instructors should avoid punishing or rewarding distant learner because of their location, i.e. setting different times to submit assignments for students who take the course at a distance or in class. Jones (2002) also argues that in order to ensure singularity, it is necessary that learners respond to the test items at the same time despite their location. Furthermore, Nouwens and Towers (1997) point out the limited opportunities for dialogue between learners and instructors in distance education. They suggest the followings to increase the effectiveness of the assessment process at a distance (online document):

"1. develop the learners' independent study skills, 2. promote educational dialogue between the lecturer and student, 3. help identify and deal with students' misconceptions, 4. give direction to learning in key subject areas, 5. relate learning to student work and experiences, 6. permit students to assess their progress, 7. provide fair, valid and reliable assessment, indicate to lecturers the quality of teaching in a subject, 8. provide feedback about strengths and weaknesses of the study materials."

Tarouco and Hack (online document) also discuss the role and quality of interaction regarding assessment at a distance. They state that instructors use other mechanisms such as body language, participation etc. besides formal mechanisms in face-to-face instruction. In distance education contexts, on the other hand, only formal mechanisms generally take place. Nevertheless, with the advancement in computer technology, this is no longer an issue and networks and the Internet are likely to fill in this gap.

According to Simonson and others (2000), one of the responsibilities of the distance instructor is to make sure that learners are familiar with the technological tools for class as well as the assessment strategies that will be utilized. It is essential to allow students to practice 'digital dropbox' for assignments, online chats, web-based quizzes etc. Another important issue to be discussed is diversity. As Wangsatorntanakhun (1997) states one of the goals of performance assessment is to pay attention to student diversity in terms of learning styles, cultural backgrounds, and proficiency levels. Most distance education programs address a highly diverse group of learners in terms of age, race, socioeconomic status etc. Instructors need to be aware of the fact that besides enriching the socio-cultural interactions among learners, diversity might cause some problems as well. For example, while younger adults might perceive alternative assessment strategies useful, older adults might find traditional assessment tools more effective (Simonson et al., 2000).

Conclusion

Which type of assessment is more appropriate for distance learner, traditional or alternative one? When we consider the main characteristics of distance learner, alternative assessment seem to be more beneficial. In spite of the fact that distance education is being implemented at many elementary, middle, or/and high schools,

the major population in this type of education is adults. It is possible that these adult learners are full time working people who have family or other social responsibilities. Although they might be away from school environment for a long time, most of the times they are highly motivated. The main characteristics of alternative assessments are great benefit to distant student. Ongoing assessment activities as well as self-based assessment tools remove the time pressure on the learner (Simonson et al., 2000). However, it might not be possible to implement alternative assessment strategies all the time. A good example for that is Anadolu university, which is a well-known open university in Turkey. Distant students are assessed at specific locations at the same time by showing their picture IDs. The mid-term and final exams are standardized tests in multiple-choice format. This is believed to be an effective assessment strategy to make sure that all learners meet the standards determined by the university. There are no specific instructors for particular students. In other words, the instruction is delivered via TV to all students and the students are expected to revise and practice the information by reading their textbooks. Therefore, it is not feasible to implement alternative assessment activities at this context unless some major changes done in development and delivery of the instruction.

On the other hand, using multiple-choice tests in an online instructional system design course will not be helpful for the distant students–even for on site learners. When employing alternative assessment strategies, it is necessary to increase the effectiveness of this type of instruction since the course content requires learning by doing. It is an instructor's own preference to include some traditional assessment tools to add variety. Similarly, when we consider a geography course at a distance, both types of assessment strategies can be implemented successfully. Multiple choice or short answer format can be used for items that require retention, i.e. name of the capitals. Moreover, projects or portfolios are effective strategies to employ in assessing the student over time.

As a result, there is no best way to assess distant learners. As discussed earlier, there are pros and cons of both types of assessments. A balanced approach between traditional and alternative assessment is critical. While deciding what assessment strategy to use, instructors need to consider the issues such as content, context, audience. Having clearly defined the objectives, appropriate assessment tools need to be utilized. Depending on the nature of the instruction, a combination of both assessment techniques might be useful.

Reference:

Abrenica, Y. Electronic portfolios. Available online: http://edweb.sdsu.edu/courses/edtec596r/students/Abrenica/Abrenica.html

Arter, J. A., Spandel, V., & Culham, R. (1995). Portfolios for assessment and instruction. ERIC Digest. EDRS NO: ED388890.

Bailey, K. M. (1998). Learning about language assessment: dilemmas, decisionjs, and directions. Heinle& Heinle: US.

Bennett, R. E. (1997), Reinventing assessment: Speculations on the future of large-scale educational testing. ETS (Educational Testing Service). Policy Information Center: Princeton, NJ.

Brualdi, A. (1998). Implementing performance assessment in the classroom. Practical Assessment, Research &Evaluation, 6(2). Available online: http://ericae.net/pare/getvn.asp?v=6&n=2

Dietel, R. J., Herman, J. L., & Knuth, R. A. (1991). What does research say about assessment? NCREL, Oak Brook. Available online: http://www.ncrel.org/sdrs/areas/stw_esys/4assess.htm

Elliott, S. N. (1995). Creating meaningful performance assessments. ERIC Digest E531. EDRS no: ED381985.

Jones, T. (2002). Options and Considerations for Distance Education: Leraner Assessment and Self-assessment, (3)3. TOJDE (Turkish Online Journal of Distance Education). ISSN 1302-6488 Available at: http://tojde.anadolu.edu.tr/tojde7/articles/Jonestxt.htm

Kulieke, M., Bakker, J., Collins, C., Fennimore, T., Fine, C., Herman, J., Jones, B.F., Raack, L., & Tinzmann, M.B. (1990). Why should assessment be based on a vision of learning? [online document] NCREL, Oak Brook: IL. Available online: http://www.ncrel.org/sdrs/areas/rpl_esys/assess.htm

Law, B. & Eckes, M. (1995). Assessment and ESL. Peguis publishers: Manitoba, Canada.

Niguidula, D. (1993). The digital portfolio: a richer picture of student performance [online document]. *CES National*. Available online: http://www.essentialschools.org/cs/resources/view/ces_res/225

Nouwens, F. & Towers, S. (1997). Assessment in distance education. Cited in S. Dixon, G. Lefoe, F. Nouwens and S. Wills (eds) Teaching at a distance [CD-ROM]. PAGE: Melbourne. Available online: http://cedir.uow.edu.au/programs/flexdel/resources/AssessmentDistEd.html

Reeves, T. C. (2000). Alternative assessment approaches for online learning environments in higher education. Educational Computing Research, 3(1) pp. 101-111.

Simonson M., Smaldino, S, Albright, M. and Zvacek, S. (2000). Assessment for distance education (ch 11). *Teaching and Learning at a Distance: Foundations of Distance Education.* Upper Saddle River, NJ: Prentice-Hall.

Tarouco, L. & Hack, L. New tools for assessment in distance education [online document]. Available online: http://www.pgie.ufrgs.br/webfolioead/artigo1.html

Wangsatorntanakhun, J. A. (1997). Designing performance assessments: challenges for three-story intellect [online document]. *Teaching with internet*. Available online: http://www.geocities.com/Athens/Parthenon/8658/

Winking, D. (1997). Critical issue: Ensuring equity with alternative assessments [online document]. NCREL (North Central Regional Educational Laboratory), Oak Brook: IL. Available online: http://www.ncrel.org/sdrs/areas/methods/assment/as800.htm