Universal Design and the Assessment of Student Learning in Higher Education: Promoting Thoughtful Assessment

Nicole S. Ofiesh
California State University, East Bay

Colleen M. Rojas
University of Arizona

Robin A. Ward
University of Arizona

Abstract

Universal design, applied to instructional delivery systems, can be a powerful way to promote greater access to information for a wide range of students. However, the assessment of this information is of equal importance. This article provides a brief review of current assessment practice in postsecondary settings, and explains how universally designed assessments can help to promote a more accurate representation of student learning. Basic recommendations are provided to help faculty who are interested in universally designed assessment.

At institutions of higher education, designing assessments that accurately measure what students have learned in large classes can be daunting. This is especially so for discipline-based faculty who are experts in content but who often have little professional preparation in college instruction and assessment. Couple this with sometimes competing research and teaching demands, and the task has the unfortunate potential of becoming an almost undesirable undertaking for some faculty.

Over the past two decades, educators have sought to improve methods to assess student learning. Fair and accurate assessment of knowledge for many persons, including those with cognitive disabilities, as well as nonnative speakers of English, is a multifaceted challenge: How do faculty members create assessments that are accessible to diverse populations while upholding the same standards for all students, ensuring academic rigor, and assessing the full taxonomy of understanding (Bloom, 1984; Bloom, Madaus, & Hastings, 1981; Guilford, 1959; Krathwohl, Bloom, & Mosia, 1964)? How can assessments reach diverse populations through language (e.g., vocabulary, wording), format (e.g., constructed response, multiple choice), and presentation (e.g., computer-based, paper-pencil, three-dimensional)?

The purpose of this paper is to present recommendations from the field of universal design as they apply to assessment of students at the postsecondary level. The recommendations are gleaned from three models of universal design: Universal Design for Learning (UDL), Universal Design for Instruction (UDI), and the work from the National Center on Educational Outcomes (NCEO). The ideas posited through UDL and UDI are presented earlier in this special issue and will not be further elaborated on, except to illustrate their contributions to assessment of student learning, as each of these models encompasses assessment differently. The recommendations are intended to provide faculty, disability services, and lead instructional personnel (e.g., teaching center staff) with the beginning steps to foster the role of accessible assessment for all learners at institutions of higher education.

Background

At the postsecondary level, course instruction is often traditional in nature: taught through lecture and evaluated using paper-and-pencil, in-class, timed tests. As a result, exams are usually designed so that they are easily

administered and graded, usually in the form of a multiple-choice test (Cantor, 1987; Rodriguez, 2005). These types of tests are typically timed according to the length of the class as opposed to a function of the test (Ofiesh, Mather, & Russell, 2005). Under this assessment format, the degree to which the instructor acquires information regarding how well students understand various concepts is dependent, in part, on the way the test question was designed (e.g., one that measures application, rote memory) (Haladyna, Downing, & Rodriguez, 2002; Rodriguez, 2003). Further, while multiple-choice exams can potentially provide useful assessment data, careful attention must be paid to formatting text such that the tests are accessible to a wider range of students in electronic format (Rose & Meyer, 2002).

When these qualities of test design are taken into consideration, routine multiple choice tests can serve as a valid tool for assessment of student learning. However, in the typical college classroom, little attention is given to the relationship between the test question and quality of response, or alternative ways to assess students. Thus, this conventional approach to instruction and assessment often limits the ways in which students can demonstrate what they know and/or, for students with disabilities, may require a test accommodation(s) in order to ensure fairness. Though the findings are mixed, research on multiple-choice tests has called into question the efficacy of this format in terms of constituting an adequate and fair assessment tool for various populations and under differing situations (Abu-Rabia, 2003; Breland, Yong-Won, & Muraki, 2005; Cohen & Rosenzweig, 2006; Klecker, 2000; McCoubrie, 2004; Rodriguez, 2005; Stuyven, Dochy, & Janssens, 2005; Walker & Thompson, 2001). Moreover, brain research continues to underscore the point that individuals process, retrieve, and produce information in vastly different ways (Demg, Boynton, & Heeger, 1997; Kasniak, 1996; Newman & Kasniak, 2000; Miller et al., 2002; Osaka et al., 2003; Wager, Jonides, Smith, & Nichols, 2005).

As awareness of the limitations of traditional assessment grow, a shift toward more inclusive delivery and evaluation of course content is emerging, using pedagogies such as UDL (Rose & Meyer, 2002), UDI (Scott, Shaw, & McGuire, 2005), and universal instructional design (Silver, Bourke, & Strehorn, 1998). In addition to the development of improved multiple-choice tests, individuals who are responsible for assessment have experimented with rubrics, portfolios, projects, combinations of test formats, computer-based tests, and flexible text through assistive technology. The results of research designed to examine the validity of alternative kinds of assessment is emerging, and like that of multiple-choice tests, the findings are mixed (Breland, et al., 2005; Cahalan-Laitusis, Cook, & Aicher, 2004; Klecker, 2000; Lu & Suen, 1995; Stearns, 1996).

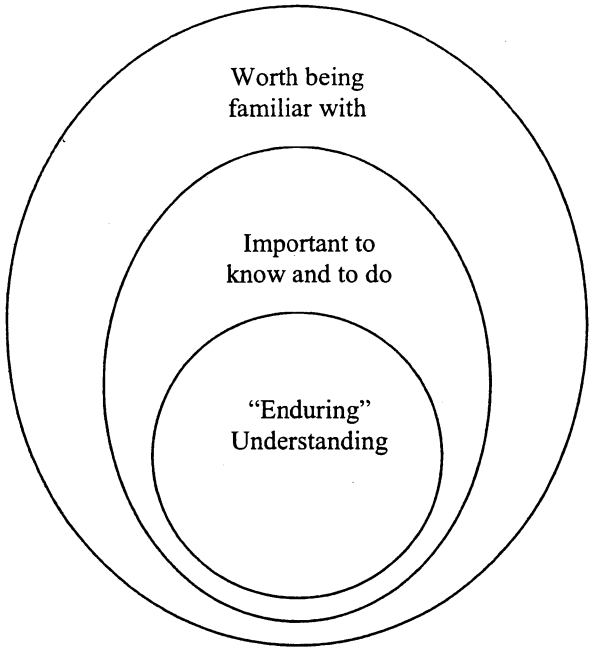
In light of these mixed findings, several points repeatedly emerge that are useful to the development of universally designed assessments. Regardless of the type of assessment a faculty member designs (e.g., project, short answer, oral presentation), effective assessment occurs when faculty: (a) are clear *up-front* about what level of student understanding they seek (e.g., factual knowledge, applied knowledge); (b) identify *upfront* how the assessment relates to the goals and objectives of the teaching/learning dynamic; and (c) consider *up-front* the role of time in the assessment (Rodriguez, 2003, 2005; Sax, 1997; Stiggins & Conklin, 1992). In short, the focus of the assessment can be considered the force behind instruction, as faculty think through what the expectations are for their students in terms of learning outcomes. That is, instruction is planned from the end goal.

Backward Design

Some theorists have referred to the planning of instruction from the end goal as backward design (Wiggins & McTighe, 1998). "Backward design" is not directly aligned with universally designed assessment as it is rooted in the development of curriculum rather than universal access. However, backward design is a critical first step in the analysis of a curriculum to ensure that it is amenable to universal design and access. In backward design, the curriculum becomes a means to an end; that is, the curriculum is analyzed so that faculty members have a clear idea about the goals and objectives (i.e., the final outcome) of the course from the very beginning. Wiggins and McTighe (1998) encourage instructors to think through their course content and identify the desired results (on the part of the students) from the onset. Once the desired results are identified, they are classified in terms of "curricular priorities." These curricular priorities are depicted using the three nested rings shown in figure 1.

The largest ring represents knowledge that may not be covered in depth in class but that is "worth being familiar with." In the middle ring, "important knowledge" (facts, concepts, and principles) and "skills" (processes, strategies, and methods) are identified. According to Wiggins and McTighe (1998), student learning is incomplete without mastery of these essentials. The smallest ring is known as "enduring" understanding. Enduring understanding are the aspects of the course that depict the "big ideas" or guiding forces that students will remember long after the details are lost. Wiggins and

Figure 1. Curricular priorities and assessment.



G. Wiggins & J. McTighe, 1998, Alexandria, VA: *Understanding by Design*, Reprinted with permission. The Association for Supervision and Curriculum Development.

McTighe propose that traditional paper-pencil-based quizzes and tests that use selected response (multiple-choice, true-false, matching) and constructed response (short-answer, essay, list) is best for knowledge that is classified as "worth being familiar with" and "important to know and do." Performance-based tasks and projects (authentic tasks that require production or performance) are best for knowledge that has been classified as "enduring" understanding, as well as "important to know and do."

In order to allow for universally designed assessments to emerge, educators, test developers, and disability service providers must understand the importance of aligning assessment with course objectives at the initial planning stage rather than at the end of instruction. Rather than ask at the end of instruction, "How and what should I assess?" and then adapt or accommodate the assessment, an accessible assessment is designed from the onset in connection with the goals and objectives of instruction. Employing universally designed principles may not eliminate the need for specific accommodations for students with disabilities, but it may minimize the need for certain accommodations (Burgstahler, 2005).Thus, it is at this stage of instructional planning that universally designed assessments can be considered. In fact, research indicates that one of the most important recommendations gleaned from studies on test effectiveness is that *test construction should relate directly to the specific knowledge of behaviors addressed in the instructional objectives* (Hoepfl, 1994). Nitko (1989) noted, "the negative consequences of using tests that are not adequately linked to instruction include an inaccurate understanding about the progress students are making, the use of inappropriate information to make decisions which affect students' welfare, and incorrect assessments about the effectiveness of instruction" (p. 447).

Thoughtful Assessment

Identification of essential knowledge is a key component of all courses in higher education and, as decades of research have shown, it is an intricate part of effective instruction and student engagement (Denham &

Lieberman, 1980; Dolan & Hall, 2001; Howell & Nolet, 2000). This initial planning of assessment in relation to course goals and universal design is what we term "thoughtful assessment." Thoughtful assessment serves both the instructor and the student. When an assessment is carefully aligned to course content, as well as what the instructor plans to achieve in a lesson, unit, or course, the instructor receives valuable feedback on how well the instruction worked to accomplish his or her goals and objectives. With this feedback, instruction can be adjusted, preconceived notions about the students' background knowledge and preparation for the material can be reevaluated, and the appropriateness of the course content and materials can be reconsidered.

For example, upon reflection, an instructor may find that the class was designed very well, but that some students were not engaged for reasons not related to the course. A thoughtfully designed assessment system allows faculty to troubleshoot those reasons related to the course, eliminate ineffective delivery systems, and maximize the opportunity for all students to achieve with equal access to instruction and assessment. Not only does a carefully crafted assessment provide the instructor with the clarity necessary to design an effective and accessible instructional delivery system, it allows the student an opportunity to become engaged in the curriculum with a clear set of expectations and to demonstrate his or her knowledge, and it provides valuable information in terms of which areas are open to professional and educational growth.

The ideas espoused by Wiggins and McTighe (1998) provide a foundation for assessment and universal design, especially for faculty who want to explore possibilities of universal design, but who may feel reluctant to change current practices. Once faculty members identify the essential knowledge to be gained and skills to be acquired by students enrolled in a course, as well as the best methods for assessment, elements of universally designed assessments can be applied.

Universally Designed Assessment

A thoughtful, universally designed assessment consists of a multitude of considerations, including, but not limited to, subject content, electronic flexibility, English language usage, format options (e.g., essay, short-answer), time limits, text characteristics, a direct link from the goals and objectives of the course, instruction, and informational delivery system, and more. The idea of designing educational opportunities that uphold intellectual rigor and that are accessible for the majority of learners holds great appeal to most faculty who embrace teaching. However, without a careful analysis of goals and objectives from the onset, well-meaning faculty can run the risk of applying principles of universal design that may substantially (and undesirably) change the goals or objectives. Once the most critical aspects of the course are clear to both the teacher and the student, access can be successfully built into both instruction and assessment. In this manner goals and objectives essentially serve as a roadmap for the link between the creation of accessible instruction and accessible assessment.

Like universally designed instruction, the purpose of universally designed assessments is to develop assessments that allow participation of the widest range of students and result in valid inferences about their performance (Thompson & Thurlow, 2002). This can be achieved for both large-scale standardized tests and classroom tests spanning from kindergarten to postsecondary levels. The type and scope of the assessment design depends on the purpose of the test and the construct(s) being measured. The overarching goal is always the same, however: design of assessments that allow for meaningful interpretation of performance with the greatest access to information by the widest range of individuals.

National Center for Education Outcomes (NCEO)

Once clarity regarding instructional goals and objectives has been established, one way to begin to create an accessible assessment is by applying some of the principles from the National Center for Educational Outcomes (NCEO). Researchers at NCEO have developed a list of elements that comprise "universally designed assessments" (Thompson, Johnstone, & Thurlow, 2002) (see Table 1). While these elements were designed to apply to large-scale tests at the K-12 levels, four of the elements, principles 4-7, can readily serve as a guide for the development of classroom assessment, especially the majority of tests which have traditionally been presented via printed text.

Amenable to accommodations (principle 4). Universally designed assessments should start with the notion of inclusivity and accessibility. However, there will still be a need for accommodations on some tests, and the test results should still indicate a valid representation of a student's performance. Faculty test developers can increase ease of accommodation use and alternate format by employing the following recommendations:

1. Refrain from using irrelevant graphs or pictures, as well as vertical or diagonal text.
2. Place keys and legends directly under the text where they are to be applied.
3. Include verbal/textual descriptions that can be translated into Braille with pictures or graphics.

Table 1

Elements Comprising Universally Designed Assessments

1. Inclusive assessment population
2. Precisely defined constructs
3. Accessible, non-biased items
4. Amenable to accommodations
5. Simple, clear, and intuitive instructions and procedures
6. Maximum readability and comprehensibility
7. Maximum legibility

Simple, clear, and intuitive instructions and procedures (principle 5). Students from diverse backgrounds bring to the assessment situation a variety of characteristics that should not be evaluated as part of the test. These include a student's life experience, knowledge, language skills, or current concentration level. Therefore, assessment instructions and procures need to be simple, clear and consistent, and include sample items, practice questions, and scoring criteria when appropriate (Thompson et al., 2002). Further, the test should measure what the test developer (i.e. the instructor) intends the test to measure.

Maximum readability and comprehensibility (principle 6). A variety of factors can impact the readability and comprehensibility of text. The wording of test items should test the content presented as part of a course, as opposed to the student's reading ability. As noted in NCEO's synthesis report on universally designed assessments, Gaster and Clark (1995) recommended eight readability guidelines for all print materials:

1. Use simple, clear, commonly used words, eliminating any unnecessary words.
2. When technical terms must be used, be sure they are clearly defined.
3. Break compound complex sentences down into several short sentences, stating the most important idea first.
4. Introduce one idea, fact, or process at a time; then develop the ideas logically.
5. Make all noun-pronoun relationships clear.
6. When time and setting are important to the sentence, place them at the beginning of the sentence.
7. When presenting instructions, sequence the steps in the exact order of occurrence.
8. If processes are being described, make sure they are simply illustrated, labeled, and placed close to the text they support. (See Thompson, et al., (2002) for a complete list of characteristics of legible text.)

Maximum legibility (principle 7). Legibility refers to the appearance of print, and includes text, graphs, tables, illustrations, and response formats (e.g., bubble forms, short-answer space). A variety of factors contribute to the legibility of print, such as contrast, font type and size, leading (i.e., the amount of vertical space between lines), line length/width, and blank space. Some recommendations for maximizing legibility include:

1. Avoid gray scale and shading, particularly where pertinent information is provided.
2. To increase the readability for a wider range of persons, increase font size to 14-point (Fuchs, et al., 2000); at least 18 point for students with visual impairments.
3. Make sure type size for captions, footnotes, keys, and legends is at least 12 point.
4. Use standard typeface (Standard Typeface) or boldface, as opposed to all caps (STANDARD) or italics (*standard*).
5. Increase leading (see Arditi, 1999; Fenton, 1996; Gaster & Clark, 1995; Schriver, 1997; Worden, 1991)
6. Avoid font styles that are decorative or cursive.

The recommendations noted above are just a few examples of how assessments can be improved to increase accessibility. "Universal Design can help to ensure that assessments themselves do not produce barriers to learning" (Thompson, et al., 2002, p. 2). When coupled with instruction that is universally designed, thoughtfully designed assessments and backward designed course content provide a complete approach to rethinking learning and assessment in higher education for all students.

Universal Design for Learning

The NCEO recommendations can be extremely useful to faculty members who want to ease into the idea of universal design and to get away from current assessment practices in a step-by-step progression. These text-embedded ideas can serve as the first steps into the evolution of universally designed instruction and assessment. Faculty members can employ these techniques in all written documents that are part of instruction and assessment (e.g., quizzes, multiple choice exams, essay exams).

Employing aspects of universal design in the development of instruction and assessment not only promotes access to information, it also provides opportunities for students to demonstrate knowledge in a multitude of ways, in addition to the traditional methods of tests and papers. For many university-or college-level students, written text is one of the greatest frustrations and hindrances to academic success. Many students discover that there is significantly more to read in college than there was in high school. In order to become efficient readers, some students have learned to annotate the text as they read. Their books will contain underlining, have margin notes, and/or may be full of highlights. However, many other students have not mastered these skills and have not developed any other coping mechanism for reading and understanding their texts. It is this type of learner that requires an alternative when working with the written format.

However, text is flexible. This is a key component of the work disseminated by the Center for Applied Special Technology (CAST). CAST has introduced a set of "text transformations" that represent such an alternative for the diverse learner. "Text flexibility or transformations are simply *text modifications* and innovative *technology tools* that alter or add to the features of printed text" (CAST). These text transformations offer multiple means for students to access the information provided in their texts and to demonstrate their knowledge. In addition to utilizing traditional testing and paper evaluation techniques, teachers and professors may integrate text transformation techniques to allow for a richer and more meaningful learning experience. These practices would incorporate electronic, text-to-speech, video, videodiscs, hypertext and hypermedia in all forms of teaching, learning, and performance (CAST). In effect, this experience will generate a more accurate evaluation of the knowledge and growth gained by a student versus that which is provided in the more traditional method of a written exam.

As described earlier, many students have developed strategies that enable them to experience, understand and recollect the written word. However, these skills elude many other students at the university level: Speech recognition or print-to-speech software allows students, with a variety of learning styles, to gather, express and/or present information and knowledge gained through the text. This type of software can assist students in visually scanning a printed document, as well as enabling a student to listen to the document out loud while the computer highlights the image of the print as it is being read (Kurzweil Educational Systems, Inc.). Clearly "text transformation" provides the tools or paths with which students can gain access to the skills necessary to achieve at the postsecondary level. Providing students with flexible access to information, the ability to fully participate and engage in instruction, as well as a variety of means to express knowledge are key objectives in the goal of effective teaching and evaluation.

Universal Design for Instruction

Universal design for instruction (UDI) is a concept that was initially applied by researchers at the University of Connecticut when using universal design principles in the postsecondary environment (Johns, 2003). UDI focuses on accessibility for all students and includes accommodations that already exist and benefits many types of learners. The "universal" in UDI does not imply that one size fits all; instead, it stresses the need for flexible, customizable content, assignments, and activities. The UDI model centers on the importance of incorporating aspects of universal design into college instruction and assessment practices.

The UDI framework is unique in that it not only builds upon the basic seven principles of universal design as delineated by researchers at North Carolina State University, but also identifies two factors that focus specifically on the learning environment and its impact on the educational environment. These factors, or principles, are driven expressly by the interaction and communication among students and between students and faculty (Scott, McGuire, & Embry, 2002). Student learning is viewed as a collaborative process, in which learning is enhanced through the support of many people, including faculty, staff, and, most important, other students. Within the UDI framework this collaboration is referred to as a

community of learners. At first glance, these terms seem contrary to each other: "community," understood as a population that interacts in a common location, independent of one another and "learners," understood as an individual and his or her ability to gain knowledge or understanding by study, instruction, or experience. However, within the UDI framework, students, faculty and staff work together for the benefit of all students. Students are not only able to improve their own understanding and learning, but they are also able to work together to maximize each other's learning.

The principles of UDI have allowed assessments to become more flexible. This in turn has provided students with an equivalent, rather than identical, opportunity to demonstrate their understanding of the material thereby achieving academic success. Assessments, as they are defined through the principles of UDI, are becoming ever more inclusive. Students are encouraged to demonstrate their knowledge through the use of multiple methods, while the reliability and dependability of the measurement stay intact. At the postsecondary level, accommodations for students with disabilities are most often the method utilized to provide these types of multiple opportunities. Integrating the principles of UDI (as outlined earlier) into one's pedagogy and classroom practices will not only appropriately challenge all students, it will also provide support for increased learning through the provision of imbedded flexibility in design. It is through these diverse strategies that an adaptable assessment format can become integral to the course content.

Summary

The models of universal design presented here each contributes to the notion of universally designed assessments. Universally designed assessments are not intended to eliminate the need for accommodations by students with disabilities. Instead, it is hoped that these models can be used in combination to create more efficient and accessible assessment tools with the ultimate goal of serving all students best. Some universally designed assessments are more amenable to accommodation(s) than traditionally designed assessments. The purpose of this article was to raise awareness of the need to ensure that assessments fairly represent the goals and objectives of a course, and to point out that these assessments can be created to allow a valid representation of student performance with respect to the diversity that is part of classrooms in higher education today.

It is our hope that the recommendations presented herein, might serve as part of universal design workshops sponsored by offices of disability services or university teaching centers, as well as university offices dedicated to student assessment. Collaborative efforts across campuses are beginning to emerge as administrators in higher education recognize the growing diversity of students in higher education (see University Symposium: Rethinking Our Strategies for Assessing Student Learning [www.arizona.edu]). As the efforts continue, universal design offers new ways of thinking not only in terms of architectural access, but access to information as well.

References

Abu-rabia, S. (2003). Cognitive and social factors affecting arab students learning English as third language in Israel. *Educational Psychology, 23*(4), 347-361.

Arditi, A. (1999). *Making print legible.* New York: Lighthouse.

Bloom, B. S. (Ed.). (1984). *Taxonomy of educational objectives: The classification of educational goals book I: Cognitive domain.* New York: Longman.

Bloom, B. S., Madaus, G. F., & J., Hastings, J. (1981). *Evaluation to improve learning.* McGraw-Hill, Inc.

Breland, H., Yong-Won, L., & Muraki, E. (2005). Comparability of TOEFL CBT essay prompts: response mode analyses. *Educational & Psychological Measurement, 6,* 577-595.

Burgstahler, S. (2005). Universal design of instruction. Retrieved October 17, 2005, from http://washington.edu/doit/Faculty/Strategies/Universal/.

Cantor, J. (1987). Developing multiple choice tests items. *Training & Development Journal, 41*(5), 85-89.

Demg, J. B., Boynton, G. M., & Heeger, D. J. (1997). Brain activity in visual cortex predicts individual differences in reading performance. *PNAS, 94,* 13363-13366.

Denham, C., & Lieberman, A. (Eds.). (1980). *Time to learn.* Washington, DC: National Institute of Education.

Dolan, R. P., & Hall, T. E. (2001). Universal design for learning: Implications for large-scale assessment. *IDA Perspectives, 27*(4), 22-25.

Fenton, E. (1996). *The Macintosh font book: Typographic tips, techniques and resources* (3rd ed.) Berkeley, CA: Peachpit Press.

Fuchs, L., Fuchs, D., Eaton, S., Hamlett, C., Binkley, E., & Crouch, R. (2000). Using objective data sources to enhance teacher judgments about test accommodations. *Exceptional Children, 67*(1), 67-81.

Gaster, L., & Clark, C. (1995). *A guide to providing alternate formats.* West Columbia, SC: Center for Rehabilitation Technology Services. (ERIC Document No. ED 405689)

Guilford, J. P. (1959). *Personality.* New York: McGraw-Hill.

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), 309-333.

Hoepfl, M. C. (1994, April). Developing and evaluating multiple choice test. *The Technology Teacher,* 25-26.

Howell, K. W. & Nolet, V. (2000). *Curriculum-based evaluation: Teaching and decision-making* (3rd ed.). Belmont, CA: Wadsworth/Thomson Learning.

Jacobsen, R. (1993). What is good testing? Perceptions of college students. *College Teaching, 41,* 153-156.

Johns, L. (2003). Universal design for instruction. *Retention Pieces, 1(*1), 2-3.

Klecker, B. M. (2000). Assessing students in a graduate tests and measurement course: Changing the classroom climate, *College Student Journal, 34,* 155-161.

Krathwohl, D., Bloom, B. S., & Masia, D. (1964). *Taxonomy of educational objectives: The classification of educational goals, book II: Affective domain.* New York: Longman.

Kurzweil Educational Systems, Inc. Available on Kurzweil website, http://www.kurzweiledu.com.

Lu, C., & Suen, H. K. (1995). Assessment approaches and cognitive styles. *Journal of Educational Measurement, 32,* 1-18.

McCoubrie, P. (2004). Improving the fairness of multiple-choice questions: a literature review. *Medical Teacher, 26*(8), 709-712.

Miller, M. B., van Horn, J. D., Wolford, G. L., Handy, T. C., Valsangkar-Smyth, M., Inati, S., Grafton, S., Gazzaniga, M. S. (2002). Extensive individual differences in brain activations associated with episodic retrieval are reliable over time. *Journal of Cognitive Neuroscience, 14,* 1200-1214.

Newman, M. & Kasniak, A. (2000). Spatial memory and aging: Performance on a human analog of the MorrisWater Maze. *Aging, Neuropsychology & Cognition, 7,* 86-93.

Nitko, A. J. (1989). Designing tests that integrated with instruction. In R. L. Linn (Ed.), *Educational measurement,* (3rd ed., pp. 447-474). New York: Macmillan Publishing Co.

Ofiesh, N., Mather, N., & Russell, A. (2005). Using speeded cognitive, reading, and academic measures to determine the need for extended test time among university students with learning disabilities, *Journal of Psychoeducational Assessment, 23,* 35-52.

Osaka, M., *Osaka,* N., Kondo, H., Morishita, M., Fukuyama, H., Aso, T., & Shibasaki, H. (2003). The neural basis of individual differences in working memory capacity: An fMRI study. *NeuroImage, 18*(3), 789-797.

Rodriguez, M. C. (2003). Construct equivalence of multiple-choice and constructed-response items: A random effects synthesis of correlations. *Journal of Educational Measurement, 40*(2), 163-185.

Rodriguez, M. C. (2005). Three options are optimal for multiple-choice exams: A meta-analysis of 80 years of research. *Educational Measurement: Issues and Practice, 24*, 3-13.

Rose, D. H., & Meyer, A. (2002). *Teaching every student in the digital age*. Alexandria, VA: Association for Supervision and Curriculum Development.

Sax, G. (1997). *Principles of educational and psychological measurement and evaluation* (4th ed.). Belmont, CA: Wadsworth.

Schriver, K.A. (1997). *Dynamics in document design.* New York: John Wiley & Sons, Inc.

Scott, S., McGuire, J. M., & Embry, P. (2002). *Universal design for instruction fact sheet.* Storrs: University of Connecticut, Center on Postsecondary Education and Disability. Retrieved October 17, 2005, from http://www.facultyware.uconn.edu/home.cfm.

Stearns, S. A. (1996). Collaborative exams as learning tools. *College Teaching, 44*, 111-112.

Stiggins, R. J., & Conklin, N. F. (1992). *In teachers' hands: Investigating the practices of classroom assessment.* Albany: SUNY Press.

Struyven, K., Dochy, F., & Janssens, S. (2005). Student's perceptions about evaluation and assessment in higher education: a review. *Assessment and Evaluation in Higher Education, 30*(4), 325-341.

Thompson, S. J., Johnstone, C. J., & Thurlow, M. L. (2002). *Universal design applied to large scale assessments* (Synthesis Report 44). Minneapolis: University of Minnesota, National Center on Educational Outcomes.

Thompson, S. J., & Thurlow, M. L. (2002). Universally designed assessments: Better tests for everyone! (Policy Directions, No. 14). Minneapolis: University of Minnesota, National Center on Educational Outcomes.

Tittle, C.K. (1994). Toward an educational psychology of assessment for teaching and learning: Theories, contexts, and validation arguments. *Educational Psychologists, 29*, 249-162.

Wager, T. D. Jonides, J., Smith, E. E., & Nichols, T. E. (2005). Toward a taxonomy of attention switching: Individual differences in fMRI during multiple shift types. *Cognitive, Affective, & Behavioral Neuroscience, 5,* 127-143.

Walker, D. M., & Thompson, J. S. (2001). A note on multiple choice exams, with respect to students' risk preference and confidence. *Assessment & Evaluation in Higher Education, 26*(3), 261-267.

Wiggins, G. & McTighe, J. (1998). *Understanding by design.* Alexandria, VA: Association for Supervision and Curriculum Development.

Worden, E. (1991). *Ergonomics and literacy: More in common than you think.* Indiana. (ERIC Document Number 329 901)