Journal of Applied Research For Business Instruction
A Refereed Publication of Delta Pi Epsilon, Inc. 2008
Volume 6, Issue 4

# Assessment Challenges for Business Education in Changing Times

Sunil Hazari, University of West Georgia
Jorge Gaytan, North Carolina A&T State University
Alexa North, University of West Georgia

In addition to the difficult task of identifying teaching methods that ensure student learning, the American educational system is facing other significant challenges (Bok, 2005; Kendall, 2006), which have also been identified in two governmental reports. The first report, *Rising Above the Gathering Storm: Energizing and Employing America for a Brighter Economic Future,* was published by the Committee on Science, Engineering, and Public Policy (2006). The second one, *A Test of Leadership: Charting the Future of U.S. Higher Education,* was produced by the Secretary of Education's Commission on the Future of Higher Education (2006).

Some of the challenges include keeping college affordable; expanding college access for low-income and minority students, including the selection of proper assessment techniques that address the learning needs of students with Limited English Proficiency (LEP); increasing accountability for educational outcomes; preparing secondary students for higher education; increasing opportunities for lifelong education and workforce training; and internationalizing the student experience (American Council on Education, 2006). Schools are struggling to maintain standards for high-quality teaching while trying to address the learning needs of LEP students. The same struggle is experienced by individuals involved in the preparation of business teachers (U. S. Department of Education, 2006). Adding complexity to this problem is the fact that teacher preparation programs face increased demands for accountability by state and federal organizations, parents, and the community at large (National Council for Accreditation of Teacher Education, 2006). This increased accountability has a direct impact on the assessments being used with LEP students and poses many challenges for business teacher educators throughout this nation.

The purpose of this article is to examine assessment issues in regard to learning needs of LEP students and to discuss the impact of federal regulations that have affected teaching and testing in the classroom. Strategies for addressing these challenges are also provided.

## ISSUES IN ASSESSMENT

In today's classrooms, students' progress greatly depends on their language abilities. Students with Limited English Proficiency (LEP) are often mainstreamed into classrooms where business teachers do not necessarily have the resources or the support to meet students' needs. Without this support, children who are struggling to acquire even basic skills in their second language begin to fall behind academically and create an achievement gap that widens over time (Harris, 2003). Equally important is the need to train teachers in the use of appropriate instructional strategies and the means (such as authentic assessment) by which to assess students (Alcala, 2000).

LEP students represent a fast-growing, diverse student population in the United States. In July 2004, while the overall school population increased by 13%, the population of LEP students increased by approximately 132%, accounting for 6.7 million public school pupils. Most of the LEP students were Hispanic (79%), followed by Vietnamese (2%), Hmong (2%), Cantonese (1%), and Korean (1%) (National Education Association, 2006). As the population of LEP students in the United States continues to increase, the need for support and services to this population skyrockets. One of the most significant areas requiring attention is the monitoring of LEP students' academic performance.

A major goal of assessment is to provide information that can be used to make effective decisions for low- and high-stakes assessment. Low-stakes assessments are those that encourage both student and teacher reflection of learning and are used to inform the teaching and learning process within classrooms (North Central Regional Educational Laboratory [NCREL], Purposes for Assessing, n.d.). This process can include examining reliability, validity, and item analysis of teacher-developed tests that are administered to LEP students. High-stakes assessments provide information that affects sites, districts, schools, programs, and/or students including student retention, promotion, graduation, and assignment to instructional groups.

### *High-Stakes Assessment*

The impact of federal regulations has affected teaching and testing in the classroom. At the high-stakes level, in particular, two federal legislations address assessment of LEP students: The No Child Left Behind Act and the Carl D. Perkins Act.

***No Child Left Behind.*** With the passage of the No Child Left Behind (NCLB) Act of 2001, high-stakes tests are being used more widely than ever before. NCLB required that all children, including LEP students, reach high standards by demonstrating proficiency in English language arts and mathematics by 2014. If schools and districts are unable to demonstrate adequate yearly progress, which is typically measured as a percentage of students who pass standardized tests, corrective actions could be imposed on schools, such as school-wide restructuring or allowing students the option of transferring to other schools (Coltrane, n.d.). Bilingual education and testing have remained a very volatile issue in the educational community. When NCLB was first issued, the National Association for Bilingual Education was in support of its passage. Schools have found, though, with regard to its LEP populations, that the "NCLB is clearly failing to meet its goals" (Crawford, 2004, p. 2). A provision of the NCLB Act allows testing of LEP students in their native language for up to three years.

***Carl D. Perkins.*** The Carl D. Perkins Career and Technical Education Improvement Act IV of 2006 made changes to specific performance indicators that states and local programs have to report. At the secondary level, academic attainment is measured by NCLB state-approved academic assessments. Technical proficiency should include student achievement on technical assessments that are aligned with industry-recognized standards when possible. Further, performance measures for each indicator must be valid and reliable and to the extent possible, aligned with other state and federal programs so that similar information can be gathered to reduce administrative burdens (Association for Career and Technical Education, n.d.).

In addition to federal legislation, national accreditation boards such as the National Council for Accreditation of Teacher Education (NCATE) and the Teacher Education Accreditation Council (TEAC) also play key roles in the overall assessment of post-secondary educational institutions and their faculty members. Professional organizations such as the National Business Education Association (NBEA) and the International Accounting Standards Board (IASB) address assessment in a variety of ways. For example, NBEA addresses assessment through guidelines for the content areas of business education, through policy statements, and through yearbooks designed to prepare teachers to instruct students with diverse needs (National Business Education Association, n.d.). Similarly, IASB is committed to developing a single set of high-quality, understandable and enforceable global accounting standards (International Accounting Standards Board, 2008).

### *Low-Stakes Assessment*

Writing quality teacher-made tests for low-stakes assessment is a skill requiring mastery by teachers whether they utilize curriculum-based assessment, criterion-referenced testing, or mastery learning techniques. Research indicates that while most teachers rely on a student's performance on tests to determine grades, many do not feel competent in developing reliable and valid test questions, and few believe they have received sufficient training to do so (Popham, 2005). The goal of low-stakes classroom assessment is to improve student learning and also to provide feedback to the teacher about his/her instructional effectiveness. Formal and informal assessment practices can guide teachers in developing an informed, complete, and fair evaluation of student knowledge and understanding. Incorporating a variety of assessment practices allows LEP learners to demonstrate both concrete and higher-order thinking skills.

Several procedures can be used to improve teacher-made tests. Improvement of tests usually relates to improving reliability and validity as well as conducting item analyses to identify test items that do not fit well due to being improperly phrased, biased, or ambiguous. Test reliability refers to the ability of the instrument to measure a construct consistently; test validity refers to the instrument's ability to measure what it is intended to measure (Invernizzi, Landrum, Howell, & Warley, 2005). In a classroom of diverse students with different language skills and cultures, creating a non-biased assessment is the key to producing valid results. Joseph and Ford (2006) stated that bias is a function of differences in experience that is due to factors involving many variables, including culture and language. If reliability of a test is found to be problematic, item analysis can be used to further investigate the problem in test construction.

Item analysis uses a combination of methods to evaluate characteristics of test items, such as difficulty, discrimination, and distractibility of items (Burton, 2001). Although most teachers do not conduct an item analysis on their assessments, Strauss (2007) believed that the issue of what tests measure has become more important than ever with the onset of the NCLB legislation. Because of the accountability involved in the legislation, more school systems and teachers should revisit assessment issues of LEP students to ensure that the tests they construct are reliable and valid.

## ASSESSMENT STRATEGIES

Popham (2005) described the case of a junior high school math teacher who saw an increase in the Hispanic student population in his classroom. Because English was a second language for these students, the teacher wondered if his mathematics examinations were biased against these students and was contemplating providing a Spanish-language version of the test to reduce bias in assessment procedures. Furthermore, the teacher realized that there were increasing numbers of Southeast Asian students in his school district, so he was perplexed if translations of tests needed to be provided in different languages. Teachers in many schools today are faced with this scenario and may wonder about the use and selection of assessment strategies. Several factors should be considered, such as the limitations of language translation.

### *Limitations of Language Translation*

In a classroom with LEP students, teachers are experiencing ever-increasing demands to demonstrate student success, so construction of a reliable and valid test is critical to truly determine a student's ability. In the Popham (2005) case, although a literal translation of the test into another language may appear to be a simple solution to a teacher, psychometrically it would raise more problems than it would solve. For example, a simple question such as "What is seven times seven" would be better phrased as "What is seven multiplied by seven" for easier comprehension. Other issues would emerge from the translation that would cause reliability and validity problems. For example, Huempfner (2004) explained that there are many dialectal differences among individuals from different Spanish speaking countries. The teacher would not be able to perform a thorough check for cultural and dialectal biases. There may be a bias in the translated text that stems from the inclusion of culturally rich reading comprehension texts that represent specific populations within the Spanish speaking world. Furthermore, issues of content, capitalization, and word analysis differences exist between languages. All of these would directly affect the reliability and validity of the translated test; therefore, language modifications would be an inherent limitation for assessing teacher-made tests.

### *Empirical Analysis of Teacher-Made Tests*

In cases such as the one described above, empirical item improvement procedures involving quantitative techniques can be implemented. Although empirical techniques may appear complicated at first glance, techniques and strategies, such as use of pre-designed templates, can be used without a teacher having to know complicated formulas and perform complex calculations. Two such templates are available from Siegle (n.d.) and Elvin (2003).

Siegle (n.d.) describes a spreadsheet template that helps teachers calculate the different types of reliability index. To calculate the index, a teacher would enter scores for each student in the template. Based on results generated by the spreadsheet template, the teacher would then be able to do a "what-if" analysis by removing individual questions to see how it affects the reliability index of the entire test. Problematic items can then be identified and subjected to further item analysis.

Elvin (2003) provided a similar template for conducting item analysis. This template is also based on an Excel spreadsheet that can be used by business education teachers to refine test items and make decisions about individual items. The basic format of a spreadsheet template is to enter scores to determine what proportion of students answered each item, how many items were answered correctly, and the efficiency of distracters. Culver and McBride (2006) demonstrated use of this template for a checking account lesson in a Banking and Finance class. Using the template, the spreadsheet calculated not only the score for each student but also the parameters, such as the reliability index, average score, standard deviation, and standard error of measurement for the test.

The Item Facility (IF), which refers to the difficulty index (the proportion of students who answered the question correctly), is also included. An acceptable IF range is between .3 and .7. The item discrimination (ID) score is the difference between the item facility for the top third of test takers and the bottom third of test takers for each item on the test. Ideally this number would be greater than .2. For example, if the ID is less than .2 for any item, this would be an indication that there may not have been a high enough spread among the top third students and the bottom third students, thereby indicating that a high percentage of both groups answered these questions correctly. These questions may have been "too easy" or the instructional strategies used to teach the items in this unit may have been extraordinary. It may not be necessary to completely throw out these items or to avoid using them in future tests; however, because these items fall outside of the acceptable range, it may be a good idea to reevaluate them. This strategy of using empirical techniques for calculating reliability and performing an item analysis can make teachers more confident in the quality of their tests when used for LEP students. Since teachers who work closely with LEP students know their capabilities, they are able to monitor their students' understanding by using these empirical assessment strategies (Tanner & Jones, 2006).

***Alternative Assessments***

In general, many teachers know their students' capabilities better than any test. By including alternative assessments, results can provide a window into the classroom performance of the child that a test may not illustrate (Pierce, 2007). Teachers must realize that the LEP student population generally finds most aspects of the school experience alien: language, culture, socioeconomic levels, schedules, procedures, and building facilities (Alcala, 2000). Therefore the instructional strategies used for delivery of instruction can play an important part in LEP students' success. Students from different backgrounds have different learning styles and may do better if exposed to an alternative assessment tool. Using alternative assessment as part of the teaching process places the problem in a different perspective that is clearer to the student, which helps the student identify the problem with issues related to his or her background, which in turn can promote further learning. Alternative assessment should be an ongoing, interactive process engaging both teacher and student in monitoring the student's performance. It involves measuring student learning in forms other than traditional pencil-and-paper testing and includes any type of assessment in which students create a response to a question or task rather than choosing a response (Alternative Assessment, n.d.). Examples of alternative assessments include demonstrations, essays, exhibits, journal writing, open-ended questions, oral presentations, performance-based assessments, portfolios, and graphic representations.

Brualdi (1996) stated that traditional tests require students to show their knowledge in a predetermined manner. However, Rudman (1989) stated that no one source of data can be sufficient to assess what a student knows about school-related content. Use of Gardner's (1983) Theory of Multiple Intelligences (linguistic, logical-mathematical, musical, spatial, bodily-kinesthetic, interpersonal, and intra-personal) can be a better approach for assessment of LEP students because it allows students to explain the material in a different context. Quite often, textual material alone is complex for LEP students (Popham, 2005). However, use of other intelligence in alternative approaches during instruction and assessment can make the content information more relevant to LEP learners. For example, use of *spatial (visual)* intelligence can be encouraged by Graphic organizers such as flowcharts, webbing, Venn diagrams, and "Know-Want-Learned" (KWL) charts that not only promote active learning, but also help students interpret and summarize textual material (Tannenbaum, 1997). These visual tools may enable LEP students to demonstrate higher levels of achievement where *linguistic intelligence* (comprehension of meaning of words) may not be the best assessment tool because of language barriers. Computer simulation programs, such as the ones used in teaching personal finance topics, can be used to tap into *logical-mathematical intelligence*, which according to Tannenbaum can also be a means of tracking students' progress and performance over a period of time. This can further be enhanced by having students work in groups to tap into *interpersonal intelligence*, which helps students understand and learn from reactions, responses, and feelings of other members in the group who may come from different backgrounds. Feedback and questions from LEP students should be encouraged when using alternative assessment because these students may be shy due to a lack of self-efficacy about their English language speaking skills. By seeking feedback, a teacher can become more aware of students' reasoning skills in relation to their responses, and the teacher can scaffold the students' learning process in a non-threatening and helpful manner (Hazari & Schnorr, 1999).

***Other Test Accommodations***

Abedi, Hofstetter, and Lord (2004) recommended the following four themes when teachers decide on assessment options for LEP students: (1) Validity: translation must not affect the content of the test; (2) Differential Impact: some accommodations may be more effective with certain groups of LEP students than with others; (3) Effectiveness: the assessment must enable LEP students to demonstrate their content knowledge; and (4) Feasibility: the accommodation must be practical and affordable. Providing students with extra time to complete assessments has shown to benefit LEP students and does not require changes in the test itself (Abedi, Lord, Hofstetter, & Baker, 2000). The use of dictionaries has also been found to benefit LEP students (Albus, Bielinski, Thurlow, & Liu, 2001). In addition, oral administration of tests in LEP students' native language has produced positive effects (Abedi, Hofstetter, & Lord, 2004). These types of accommodations provide a fair opportunity to assess knowledge of LEP students without providing unfair advantage over other student groups.

**CONCLUSIONS**

As state and national standardized testing becomes the primary measure for assessing students' performance irrespective of students' background, it is important that the results obtained through this assessment accurately reflect the knowledge of a student. As Darling-Hammond (2006) wrote, "No society in a knowledge-based world can long prosper without supporting a thinking education for all of its people… or we will, within a short time, witness the contemporary equivalent of the Fall of Rome" (p. 15). An increase in

awareness and focus on changing demographics in schools are an essential component of understanding how to create successful learning environments and assessment methods for students. Business education researchers also need to assist with development of valid and reliable assessment measures for LEP students. Today, schools are held accountable by the government for achieving certain levels of performance as indicated primarily by test scores. The limitation of this accountability is that it does not consider culturally and linguistically diverse populations who may be better assessed using alternative reliable, valid, and multi-level assessment methods. This article focused on assessment issues faced by business education teachers and provided information on assessment techniques that can be used to address the learning needs of LEP students with the goal of assisting educational stakeholders. If business education teachers are aware of issues specific to LEP student assessment, they will be better prepared to teach and test in the diverse classroom.

## REFERENCES


Abedi, J., Hofstetter, C. H., & Lord, C. (2004). Assessment accommodations for English language learners: Implications for policy-based empirical research. *Review of Educational Research, 74*(1), 1–28.

Abedi, J., Lord, C., Hofstetter, C. H., & Baker, E. (2000). Impact of accommodations strategies on English language learners' test performance. *Educational Measurement: Issues and Practices, 19*(3), 16–26.

Albus, D., Bielinski, J., Thurlow, M., & Liu, K. (2001). *The effect of a simplified English language dictionary on a reading test* (LEP Projects Rep. 1). Minneapolis: University of Minnesota, National Center on Educational Outcomes.

Alcala, A. (2000). A framework for developing an effective instructional program for limited English proficient students with limited formal schooling. *Practical Assessment, Research & Evaluation*, *7*(9). Retrieved April 18, 2007, from http://PAREonline.net/getvn.asp?v=7&n=9

*Alternative assessment*. (n.d.). Retrieved April 8, 2007, from http://www.ncrel.org/sdrs/areas/issues/methods/assment/as8lk30.htm

American Council on Education. (2006). *Addressing the challenges facing American undergraduate education*. Retrieved October 1, 2006, from http://www.acenet.edu/AM/Template.cfm?Section=Home&CONTENTID=18299&TEMPLATE=/CM/ContentDisplay.cfm

Association for Career and Technical Education (ACTE). (n.d.). *Summary and analysis of major provisions and changes "Career and technical improvement act of 2006."* Retrieved April 8, 2007, from http://www.icsps.ilstu.edu/info/newsroom/PerkinsArchive.htm

Bok, D. (2005). Are colleges failing? Higher ed needs new lesson plans. *The Boston Globe*. Retrieved October 1, 2006, from http://www.boston.com/news/education/higher/articles/2005/12/18/are_colleges_failing/

Brualdi, A. (1996). Multiple intelligences: Gardner's theory. *Practical Assessment, Research & Evaluation*, *5*(10). Retrieved February 22, 2008, from http://PAREonline.net/getvn.asp?v=5&n=10

Burton, R. (2001). Do item-discrimination indices really help us to improve our tests? *Assessment & Evaluation in Higher Education*, *26*(3), 213–220.

Committee on Science, Engineering, and Public Policy. (2006). *Rising above the gathering storm: Energizing and employing America for a brighter economic future.* Washington, D.C.: National Academies Press.

Coltrane, B. (n.d.). *English language learners and high-stakes tests: An overview of the issues*. Retrieved April 8, 2007, from http://www.ericdigests.org/2003-4/high-stakes.html

Crawford, J. (2004). *No Child Left Behind: Misguided approach to school accountability for English language learners*. Retrieved February 4, 2007, from http://www.nabe.org/

Culver, A., & McBride, Y. (2006). *Assessment of checking account unit for banking and finance class*. Retrieved May 21, 2008, from http://www.sunilhazari.com/education/abed6106

Darling-Hammond, L. (2006). Securing the right to learn: Policy and practice for powerful teaching and learning. *Educational Researcher*, *35*(7), 13–24.

Elvin, C. (2003). *Test item analysis using Microsoft Excel spreadsheet program*. Retrieved March 16, 2007, from http://www.eflclub.com/elvin/publications/2003/itemanalysis.html

Gardner, H. (1983). *Frames of mind*. New York: Basic Books Inc.

Harris, J. L. (2003). Toward an understanding of literacy issues in multicultural school-age populations. *Language, Speech and Hearing Services in School*, *34*(1), 17–19.

Hazari, S. I., & Schnorr, D. (1999). Implementation and outcomes of interactive web course module. *Journal of Computing in Teacher Education*, *15*(3), 8–16.

Huempfner, L. (2004). Can one size fit all? The imperfect assumptions of parallel achievement tests for bilingual students. *Bilingual Research Journal*, *28*, 379–399.

International Accounting Standards Board (2008). Retrieved February 21, 2008, from http://www.iasb.org/home.htm

Invernizzi, M. A., Landrum, T. J., Howell, J. L., & Warley, H. (2005). Toward the peaceful coexistence of test developers, policymakers, and teachers in an era of accountability. *The Reading Teacher, 58*(7), 610–618.

Joseph, L. M., & Ford, D. Y. (2006). Nondiscriminatory assessment: Considerations for gifted education. *The Gifted Child Quarterly*, *50*(1), 42–53.

Kendall, K. (2006). *Are U.S. colleges keeping up*? Retrieved October 1, 2006, from http://encarta.msn.com/encnet/departments/elearning/?article=collegeskeepingup

National Business Education Association (n.d.). *Business Education Standards, International Business.* Retrieved July 6, 2007, from http://www.nbea.org/curfbes.html

National Council for Accreditation of Teacher Education. (2006). *NCATE 101: A primer on accreditation*. Washington, DC: Author.

National Education Association. (2006). *A report on the status of Hispanics in education: Overcoming a history of neglect.* Retrieved January 5, 2007, from http://www.nea.org/mco/images/hispaniced.pdf

North Central Regional Educational Laboratory, Learning Point Associates (n.d.). *Critical issue: Ensuring equity with alternative assessments*. Retrieved April 4, 2007, from www.ncrel.org/sdrs/areas/issues/methods/assment/as800.htm

North Central Regional Educational Laboratory, Learning Point Associates (n.d.). *Purposes for assessing, pathways home*. Retrieved April 4, 2007, from www.ncrel.org/sdrs/areas/issues/methods/assment/as8lk9.htm

Pierce, R. (2007). Development of an identification procedure for a large urban school corporation: Identifying culturally diverse and academically gifted elementary students. *Roeper Review, 29*(2), 113–118.

Popham, W. J. (2005). *Classroom assessment: What teachers need to know.* MA: Pearson.

Rudman, H. (1989). Integrating testing with teaching. *Practical Assessment, Research & Evaluation*, *1*(6). Retrieved March 2, 2007, from http://PAREonline.net/getvn.asp?v=1&n=6

Secretary of Education's Commission on the Future of Higher Education. (2006). *A test of leadership: Charting the future of U.S. higher education*. Retrieved September 5, 2006, from http://www.ed.gov/about/bdscomm/list/hiedfuture/reports/pre-pub-report.pdf

Siegle, D. (n.d). *Reliability*. Retrieved May 10, 2007, from http://www.gifted.uconn.edu/siegle/research/Instrument%20Reliability%20and%20Validity/Reliability.htm

Strauss, V. (2007, March 26). Putting assessments to the test. *The Washington Post*, p. B2.

Tannenbaum, J. (1997). *Practical ideas on alternative assessment for ESL students*. ERIC Digest. Retrieved July 17, 2007, from http://www.ericdigests.org/1997-1/esl.html

Tanner, H., & Jones, S. (2006). *Assessment: A practical guide for secondary teachers (2nd ed).* NY: Continuum International Publishing.

U.S. Department of Education. (2006). *Teacher shortage areas nationwide listing*. Retrieved January 10, 2007, from http://studentaid.ed.gov/PORTALSWebApp/students/english/cancelperk.jsp?tab=repaying

# JARBI Publication Information

The ***Journal of Applied Research for Business Instruction*** is a refereed publication of Delta Pi Epsilon, Inc., applying research to the improvement of instruction in all business disciplines. The views expressed in this publication are those of the author(s) and not necessarily of Delta Pi Epsilon. Journal submissions and inquiries should be sent electronically to the 2008-2009 editor:

Dr. Lisa Gueldenzoph Snyder, Associate Professor
School of Business and Economics
North Carolina A&T State University
Greensboro, NC 27377
Email: gueldenzoph@att.net
Phone: 336.337.4983

Information concerning the Society's program of publications and other professional activities may be obtained from the National Office of the organization. Publication criteria may be found at http://www.dpe.org.

## 2008-2009 Editorial Review Board