

# Validity of Assessments for English Language Learning Students in a National/International Context

This paper provides a summary of research that shows a substantial performance gap between native and non-native speakers of assessment language. Such a performance gap increases with the level of language demand of the test items. Results of studies indicate that unnecessary linguistic complexity may impact reliability and validity of the assessment outcomes. To control for language factors as a source of bias, this paper introduces the concept of linguistic modification of assessment through which unnecessary linguistic complexity of test items is reduced or eliminated. The paper recommends that test developers be aware of the impact of linguistic complexity in assessment and should plan to control for such possible source of bias.

**Keywords:** non-native language learning, content based assessment, linguistic complexity, bias in language assessment.

## Nc016

Jamal Abedi

University of California, Davis  
jabedi@ucdavis.edu

## Validez de las evaluaciones en los estudiantes de la lengua inglesa en un contexto nacional e internacional

Este trabajo sintetiza algunos estudios que muestran importantes diferencias de rendimiento entre los alumnos nativos frente a los no nativos al ser evaluados en la lengua inglesa. Este déficit se incrementa a medida que la complejidad lingüística de las pruebas lo hace. Los resultados de numerosos estudios revelan que hay una complejidad lingüística innecesaria en las pruebas de evaluación, lo que produce un efecto negativo en la fiabilidad y en la validez de los resultados de las evaluaciones. Este trabajo propone un enfoque de modificaciones lingüísticas de la evaluación a partir del cual se puedan modificar o eliminar los

ítems que presenten una complejidad innecesaria evitando así el sesgo que ello produce.

**Palabras clave:** Aprendizaje de segundas lenguas, complejidad lingüística en la evaluación, evaluación basada en contenidos, sesgos en evaluación del lenguaje.

## 1. RATIONALE

Learning a new language is a challenge for all students throughout the world. Such learning becomes even more challenging if the language is used for instruction and assessment purposes. As elaborated by Meskill (in press), “[...] there are considerable differences between learning a language in formal settings for limited use outside of the target culture (“foreign language learning”) and mastering the language of the culture in which one lives and studies” (p. 2).

The focus of this paper is on students who are learning English as their academic language (i.e., language of instruction and assessment). The studies that are summarized in this paper are all conducted on students in the United States where a large majority of English language learners are instructed and assessed in English. The outcome of these studies may be generalized to other students all over the world who learn English as their academic language. The results may also be applicable to learners of other languages.

Students learning English as a second language (foreign language learning) may acquire English proficiency at the general level but they may not become proficient enough at the level that is required for learning academic content. A distinction exists between basic interpersonal communication skills (BICS) and cognitive academic language proficiency (CALP) (Bailey & Butler, 2003; see also Cummins, 2000). In the context of assessments, language proficiency tests could vary in the extent they gauge CALP. Bailey and Butler (2003) defined academic language as “language that stands in contrast to the everyday informal speech that students use outside the classroom environment” (p. 9). In other words a student could score high in basic interpersonal communication skills (BICS) but low in cognitive academic language proficiency (CALP).

Research literature on language acquisition has shown that it takes a much longer time for English language learners to gain sufficient mastery of academic English to join English speaking peers in taking full advantage of instruction and assessment in English (Hakuta, Butler, & Witt, 2000). During this period, learning cannot occur at

## NOTAS

VALIDITY OF ASSESSMENTS  
FOR ENGLISH LANGUAGE  
LEARNING STUDENTS IN A  
NATIONAL/INTERNATIONAL  
CONTEXT

the same rate it would for a native speaker of English when that instruction is offered only in English. Limited English proficiency may make it difficult to benefit fully from the teacher's instructions and to understand assessment questions. Students learning English as a foreign language may learn in terms of basic interpersonal communication skills. "Conversational fluency is often acquired to a functional level within about two years of initial exposure to the second language whereas at least five to seven years is usually required to catch up to native speakers in academic aspects of second language" (Cummins, 1981, p. 1; see also Collier, 1987; Hakuta, Butler, & Witt, 2000; Klesmer, 1994). Cummins indicates that failure to distinguish between BICS and CALP may lead to many problems including a premature exit from a language support program and students in mainstream classes encountering difficulty assimilating.

In this paper, we discuss issues concerning assessment of ELL (English language learners) students; therefore, our focus will be on English as the language of assessment. We distinguish between academic language that facilitates learning content knowledge and unnecessary linguistic complexity that impacts the validity and authenticity of assessment for ELL students in content-based areas such as mathematics and science. We will also demonstrate how unnecessary linguistic complexity as a source of construct irrelevant variance may negatively impact student learning and consequently affecting the validity and reliability of assessment outcomes for ELL students.

Our focus in this paper will be on two aspects of language used for instruction and assessment purposes: 1) reaching the level of proficiency in English that is needed to successfully participate in academic discourse both in instruction and assessment, and 2) having access to English language that is free of unnecessary linguistic complexity that prevents students from understanding the intended construct. We will elaborate on each of these two important aspects that affect academic performance of ELL students.

## **2. LEVEL OF PROFICIENCY IN ENGLISH NEEDED TO SUCCESSFULLY PARTICIPATE IN INSTRUCTION AND ASSESSMENTS IN ENGLISH**

In the United States and other countries with English as their official language, state and national assessments are mainly constructed and normed for students who are fluent in English. Therefore, there might be linguistic factors that could seriously undermine the validity of content-based assessment for English

language learners. Hence, it is imperative to determine at what level of proficiency in English ELL students can meaningfully participate in the state and national content-based assessments in English. Literature has clearly demonstrated that participation of ELL students in content-based assessment in English when they are not fully prepared to understand the assessment questions may not be productive. For example, research suggests that ELL students do poorly in content-based assessments in English compared with non-ELL students and the performance gap between ELL and non-ELL students increases as the level of language demand of the assessment increases (Abedi, 2006b; Abedi & Gándara, 2006; Maihoff, 2002; Solano-Flores & Trumbull, 2003). Unfortunately, however, issues concerning the impact of language factors are not considered in the development process of many standardized achievement tests. Therefore, many of these assessments at the state and national levels may not present a comprehensive picture of what ELL students know and are able to do. It is therefore imperative for ELL students to participate in the state content-based assessments only when their English language proficiency level matches that of the content-based assessment language demand.

### **3. THE IMPACT OF LINGUISTIC COMPLEXITY ON THE ASSESSMENT OF ELL STUDENTS IN CONTENT-BASED AREAS**

Complex linguistic structure of instructional and assessment materials may have negative consequences on ELL student learning. The impact of such linguistic complexity could seriously jeopardize ELL students' academic progress. A number of studies have examined the impact of language factors on the assessment and instruction of ELL students. The results of these studies clearly and consistently suggest that language factors play a major role in the academic performance of ELL students. When instructional materials contain complex linguistic structures, ELL students may be faced with serious difficulty in following such instruction and understanding the content of instruction. Similarly, it is extremely difficult for ELL students to understand test items that are complex in their linguistic structure. In such cases, ELL students with a fair level of knowledge of the content may not perform well not because of lack of content knowledge but because of difficulty understanding the assessment questions.

Research has provided ample evidence on the impact of linguistic complexity in the assessment of ELL students. For example, findings

## NOTAS

VALIDITY OF ASSESSMENTS  
FOR ENGLISH LANGUAGE  
LEARNING STUDENTS IN A  
NATIONAL/INTERNATIONAL  
CONTEXT

from the analyses of extant data from the National Assessment of Educational Progress (NAEP) in the United States demonstrated that ELL students had difficulty with the test items that were linguistically complex regardless of content difficulty. The study also found that ELL students exhibited a substantially higher number of omitted/not-reached test items since it took them a much longer time to read and understand assessment questions (Abedi, Lord, & Plummer, 1997). Other studies also suggested that mathematics test performance of some students has been affected by differences in the syntactic complexity of the language of word problems (Larsen, Parker, & Trenholme, 1978; Wheeler & McNutt, 1983).

Even minor changes in the language of mathematics word problems to make them more accessible to ELL students can affect student performance (Cummins, Kintsch, Reusser, & Weimer, 1988; De Corte, Verschaffel, & De Win, 1985; Hudson, 1983; Riley, Greeno, & Heller, 1983). One way to examine the impact of language factors on the assessment of ELL students is to experimentally control the level of unnecessary linguistic complexity of assessments and observe the changes that such reduction in the level of linguistic complexity may have on ELL students' assessment outcomes. Results of the studies presented above encouraged experimental studies to examine the impact of language factors on the assessment of ELL students by randomly assigning test items with different level of linguistic complexity to groups of students and comparing their performance.

Several studies were conducted to identify sources of linguistic complexity that may slow down the reader and make misinterpretation more likely. In one study, researchers found 48 linguistic features that may have serious impact on students' understanding of the test items and grouped them into 14 general categories (Abedi et al., 1997). These features included subordinate clauses, unfamiliar vocabulary (low-frequency words, typically long words), passive voice constructions, modal verbs, and participial modifiers. For example, ELL students have more difficulty with mathematic items that are expressed in passive voice rather than active voice. They also have difficulty with unfamiliar vocabulary that is not related to the mathematics concepts being tested. The impact of these linguistic features on the performance of ELL students in content-based areas (math and science) was then examined. A short description of each of these 14 features along with research evidence of the impact of these features on assessment of ELL students is presented later in this paper.

#### **4. FINDINGS OF EXPERIMENTALLY CONTROLLED STUDIES ON THE IMPACT OF LINGUISTIC COMPLEXITIES ON THE ASSESSMENTS OF ELLs**

A study by Abedi and Lord (2001) compared the performance of eighth grade students on NAEP mathematics items with parallel items that were modified to reduce the complexity of sentence structures and to replace potentially unfamiliar vocabulary with words likely to be more familiar to the students. The mathematics tasks were not changed, nor were mathematical terms in the items. In this study two forms of the mathematics test were created: 1) the original form with some items that were linguistically complex and 2) the same items in the original version that were revised in order to reduce the level of unnecessary linguistic complexity. A sample of 1,031 students in intact classrooms was randomly assigned to original and modified versions of the items. Test results showed ELL students scored significantly higher on the modified items where the complexity of the linguistic structure of the items was reduced.

This study which controlled for many sources of threats to internal validity of the design clearly demonstrated that ELLs and low-performing students benefited the most from language modification. The linguistic features that appeared to contribute to item difficulty included low-frequency vocabulary and passive voice verb constructions (see Abedi et al., 1997, for discussion of linguistic features).

In another study with 1,394 eighth grade students in schools with a high enrollment of Spanish speakers, students generally scored higher on the linguistically modified version of the test. Modification of the item language contributed to improved performance on 49% of the items (Abedi, Hofstetter, & Lord, 1998).

#### **5. HOW THE EFFECTS OF LINGUISTIC COMPLEXITIES OF ITEMS CAN BE REDUCED IN THE ASSESSMENT OF ELL STUDENTS?**

To minimize the impact of language factors on ELL students' performance outcomes, one may reduce the level of unnecessary (non-essential) linguistic features of the assessment—that is, complexity unrelated to the construct being assessed—and then examine the impact of such reduction of linguistic complexity on student performance. If the hypothesis that reducing linguistic complexity on assessment provides clearer interpretations of student performance is supported, then improvements on the outcome of assessments can be observed by using less linguistically complex test items. This process is referred to as *linguistic modification*.

## 6. LINGUISTIC MODIFICATION OF ASSESSMENT

The concept of linguistic modification applies to areas in which content other than language is being assessed (e.g., mathematics, science and social sciences) since the language construct may be unrelated to the purpose of assessment. However, the judgment of whether language is related or unrelated to the target of assessment is arguable. Some researchers determine whether language is related or unrelated based on the judgment of content experts. For example, Abedi et al. (1997) presented both the original and the linguistically modified version of the math test used in their study to two math content experts independently. The math content experts were asked to compare the original test item with its modified version and make a judgment as whether or not the content was altered in the process of linguistic modification. They both provided some minor suggestions but generally agreed that the math content in the items had not been altered.

There is a difference between language that is an essential part of the content of the question and language that makes the question incomprehensible to many students, particularly to English language learners. While it is important to understand and value the richness of language in an assessment system; it is also important to make sure that English language learner students and other students with similar language needs not be penalized for their lack of English proficiency in areas where the target of assessment is *not* language. Though we understand the views of some language modification critics in not “dumbing down” assessment questions by simplifying the language, we also recognize the distinction between necessary and unnecessary linguistic complexity. Content assessment specialists should make these distinctions when creating test items. Therefore, it is imperative to clearly elaborate on these issues and provide recommendations on how to deal with the issues of language factors in the assessment of ELL students. Findings of the studies presented above clearly show the significant impact of language factors on the assessment of ELL students.

## 7. SOURCES OF LINGUISTIC COMPLEXITY AFFECTING COMPREHENSION

Research has identified several linguistic features that appear to contribute to the difficulty of comprehending text (Abedi et al., 1997; Abedi, 2006a; Abedi, 2006b; Shaftel, Belton-Kocher, Glasnapp, & Poggio, 2006). These features may slow down the

### NOTAS

VALIDITY OF ASSESSMENTS  
FOR ENGLISH LANGUAGE  
LEARNING STUDENTS IN A  
NATIONAL/INTERNATIONAL  
CONTEXT



reader, increase the likelihood of misinterpretation, or add to the reader's cognitive load thus interfering with concurrent tasks. Indexes of language difficulty include unfamiliar (or less commonly used) vocabulary, complex grammatical structures, and styles of discourse that include extra material, abstractions and passive voice (Abedi et al., 1997). In order to better understand the need for linguistic modification, some of these linguistic features are discussed in detail below.

## **8. LINGUISTIC MODIFICATION TO REDUCE LANGUAGE DEMANDS**

The process of identifying the potentially problematic linguistic features in test items must be based on the judgment of content and linguistic experts and the actual characteristics of test items. The process can also be informed by research literature (Abedi, 2006a; Abedi et al., 1997) and knowledge of the type of linguistic features likely to cause problems for English language learners.

To illustrate the process of identifying the potentially problematic linguistic features in assessment, a summary of linguistic modifications implemented in a previously mentioned study (Abedi et al., 1997) will be presented. A test with 69 NAEP math items for eighth grade students were used to demonstrate the linguistic modification approach. Each of the 69 items was read and the mathematical operations attempted. Items in which the language was considered potentially difficult for students to understand were flagged and analyzed; linguistic features likely to contribute to the difficulty were identified and categorized. Simplified forms of linguistically complex items were drafted in order to make these items easier for students to understand. From this set of features, only the most salient and frequent language problems were selected for investigation in the field study.

Changes were made to the language of the original NAEP items in the following categories: 1) familiarity/frequency of non-math vocabulary, 2) voice of the verb phrase, 3) length of nominals (noun phrases), 4) conditional clauses, 5) relative clauses, 6) question phrases, and abstract or impersonal presentations (for a more detailed description of these changes, see Abedi et al., 1997; Abedi & Lord, 2001).

### **8.1. Familiarity/Frequency of Non-Math Vocabulary**

Potentially unfamiliar, low-frequency lexical items were replaced with more familiar, higher frequency lexical items.



■ Original: *A certain reference file contains approximately six billion facts.*

■ Revision: *Mack's company sold six billion pencils.*

The concepts of “company” and “pencils” are more familiar to ELL students, and are encountered more frequently, than “certain reference file.” If a student does not understand all the words in a test item, s/he may not understand what the question is asking and may be unable to respond to it even if those terms are not related to the mathematics content. A task places greater demands on a student if his attention is divided between employing mathematics problem-solving strategies and coping with difficult vocabulary and unfamiliar content (Gathercole & Baddeley, 1993).

In revising the items, estimates of familiarity/frequency of vocabulary were made based on established word frequency sources as well as judgment from linguistic experts of the students' familiarity with the words and concepts. For example, *The American Heritage Word Frequency Book* (Carroll, Davies, & Richman, 1971), based upon 5 million words from textbooks and library materials for Grades 3 through 9, and the *Frequency Analysis of English Usage: Lexicon and Grammar* (Francis & Kucera, 1982), based on the one million-word Brown University Corpus, listed the word “company” as occurring more frequently than “reference” or “file”.

## 8.2. Voice of Verb Phrase

Verbs in the passive voice were replaced with verbs in the active voice.

■ Original: *A sample of 25 was selected.*

■ Revision: *He selected a sample of 25.*

Passive constructions occur less frequently than active constructions in English and many other languages (Biber, 1988; Celce-Murcia & Larsen-Freeman, 1983). Children learning English as a second language have more difficulty understanding passive verb forms than active verb forms (Bever, 1970; deVilliers & deVilliers, 1973).

## 8.3. Length of Nominals

In processing longer and novel nominal compounds, people use lexical information as well as knowledge of the world and the context to rule out implausible readings. A student with a limited English vocabulary may encounter difficulty with the long nominals.

### NOTAS

VALIDITY OF ASSESSMENTS  
FOR ENGLISH LANGUAGE  
LEARNING STUDENTS IN A  
NATIONAL/INTERNATIONAL  
CONTEXT

The number of pre-nominal modifiers in a noun phrase was reduced, as in the example below:

■ Original: ...*last year's class vice president*...

■ Revision: ...*vice president*...

Post modifiers can also be ambiguous. Adding more modifiers multiplies the possibilities for ambiguity and adds to the complexity of assessments.

In a noun phrase followed by two prepositional phrase modifiers, such as:

*"the man in the car from Mexico,"* the man may be from Mexico, or the car may be from Mexico, or both.

#### 8.4. Conditional Clauses

Some conditional if clauses were replaced with separate sentences. In some instances the order of the if clause and the main clause was reversed.

■ Original: *If two batteries in the sample were found to be dead.*

■ Revision: *He found two broken pencils in the box.*

In this item, in addition to removing the conditional clause, unfamiliar vocabulary (dead batteries) was replaced with familiar vocabulary (broken pencils).

Separate sentences, rather than subordinate if clauses, may be easier for some students to understand (Spanos, Rhodes, Dale, & Crandall, 1988). Some languages do not allow sentences with the conditional clause in last position (Haiman, 1985). Consequently, sentences with the conditional clause last may cause difficulty for some non-native speakers (such as "I won't go *if it is raining*").

#### 8.5. Relative Clauses

While sometimes the number of sentences in the revised item is increased, the number of clauses per sentence is reduced. Shorter sentences with lower information density levels are more easily processed by students. Some relative clauses are removed or recast.

■ Original: *A report contains 64 sheets of paper.*

■ Revised: *He needs 64 sheets of paper for each report.*

In this example, the original version contains information in a relative clause, whereas the revised item contains the same information in a simple sentence.

#### 8.6. Complex Question Phrases

Some question structures were changed from complex question phrases to simple question words.

- Original: *At which of the following times...?*
- Revision: *When...?*

In the first example, the complex question phrase in the original version was replaced with a single question word in the revision. The single-word structure is simpler syntactically, and the placement of the question word at the beginning of the sentence gives it greater salience. The longer question phrases occur with lower frequency, and low-frequency expressions will in general be harder to read and understand (Adams, 1990).

### 8.7. Concrete versus Abstract or Impersonal Presentations

In some instances, an abstract presentation mode was made more concrete.

- Original: *The weights of three objects were compared using a pan balance. Two comparisons were made.*
- Revision: *Sandra weighed three objects using a pan balance. She made two comparisons.*

In this example, the problem statement was made more story-like by the introduction of “Sandra”. Abstract or non-situated items may employ the passive voice, but not all passive constructions are abstract or non-situated; abstract/impersonal presentations may also employ modals or generic nominal, for example. A problem expressed in concrete terms may be easier for students to understand than an abstract problem statement (Lemke, 1986).

## 9. LINGUISTIC MODIFICATION AS A FORM OF ACCOMMODATION FOR ELL STUDENTS

English language learners often perform lower than native speakers of English in academic content areas such as mathematics and science mainly due to their language barriers. To provide a fair assessment for ELL students, some forms of accommodations are offered to these students. Accordingly, many states are using accommodations for English language learners (Abedi, Kim-Boscardin, & Larson, 2000; Rivera, Stansfield, Scialdone, & Sharkey, 2000; Thurlow & Bolt, 2001). The most commonly used accommodations for ELL students are: extended time (42 of the 48 states), use of a glossary (26 states); use of an English dictionary (33 states); use of a bilingual dictionary (22 states); and linguistically-simplified test items (12 states). Rivera (2003) presents a list of commonly used accommodations for ELL students, a list that includes 73 accommodations. Based on an evaluation of the effects of these accommo-

#### NOTAS

VALIDITY OF ASSESSMENTS  
FOR ENGLISH LANGUAGE  
LEARNING STUDENTS IN A  
NATIONAL/INTERNATIONAL  
CONTEXT

dations, only 11 or 15% of those accommodations found to be relevant for ELL students. Since ELL students need language assistance the most, only those accommodations that provide such assistance would be relevant. Therefore, the linguistic modification approach was introduced as a relevant accommodation for ELL students.

Studies showed that linguistic modification as a form of accommodation helps ELL students to provide a more valid picture of what they know and are able to do without jeopardizing the validity of assessment (Abedi, Hofstetter, & Lord, 2004). For example, in a study (Abedi, Lord, Hofstetter, & Baker, 2000) comparing various accommodations with eighth grade mathematics items, the modified English version was the only accommodation that narrowed the score gap between ELLs and other students.

## 10. SUMMARY AND DISCUSSION

Learning English in the context of academic discourse may require different language content, different levels of proficiency and a different process than learning English as a foreign language. This is true in learning any second language. There are several major issues regarding learning English as the academic language. First, a distinction must be made between basic interpersonal communication skills (BICS) and cognitive academic language proficiency (CALP). Learning English as an academic language require proficiency of language content that facilitates functioning in an academic environment. High scores in basic interpersonal communication skills (BICS) do not necessarily satisfy academic language requirements. For ELL students to efficiently participate in English only academic environments, they must be equipped with the language content and vocabulary that are essential in content-based instruction and assessment. If students are not at the level of English proficiency to meaningfully participate in instruction and assessments in English, then their academic career might be at risk.

Second, the language of instruction and assessment for speakers of other languages must be free from unnecessary linguistic complexity. Instructional materials that have complex linguistic structure may not be relevant for ELL students and may not provide the same level of intended content knowledge as they provide for native speakers of English. Similarly, assessments with complex linguistic structure may not provide reliable and valid outcomes for ELL students. That is, such linguistically complex assessments may

not provide a true picture of what ELL students know and are able to do.

In this paper, we have provided both conceptual and research evidence on the impact that linguistic complexity has on assessments in order to help teachers and test developers better understand the impact of language factors on the assessment of ELL students. While language is an essential aspect of any assessment, one must distinguish between language that is related to the assessment content and language that may be irrelevant to the content. One can then reduce the impact of unrelated language factors to make assessments more reliable and valid for all students particularly those with limited language proficiency.

Based on the years of research on the impact of language factors on the assessment of ELL students, we introduced the concept of linguistic modification and provided a methodology to perform linguistic modification of test items. Research presented in this paper clearly suggests that linguistic modification of test items that does not alter the construct of assessment provides a more reliable and valid tool for assessing the content knowledge of ELL students.

The principle underlying linguistic modification of assessment is the *relevance* of language to the content being measured. If some of the language used in the assessment is judged to be irrelevant to the assessment then language may interfere with the content being measured. For example, if a content-based test such as a mathematic test has a complex linguistic structure, then the interpretation of the outcome of that test is difficult. A low score of an ELL student may be due to lack of content (mathematics) or lack of understanding of the assessment questions due to their linguistic complexity of the questions or a combination of both. Therefore, content knowledge and language factors are said to be confounded in such tests. However, judging the relevance of language to the content is a difficult undertaking. This judgment must be made by content and linguistic experts.

Based on the research and concepts presented in this paper, it can be concluded that language factors play an important role in assessment for everyone, particularly for non-native speakers of the language (English in this case). Test publishers and assessment experts do a commendable job of examining content and psychometric properties of the tests that are developed for the national and state assessments by pilot testing and field testing these tests on large and representative samples of students. However, in the process of

**NOTAS**

VALIDITY OF ASSESSMENTS  
FOR ENGLISH LANGUAGE  
LEARNING STUDENTS IN A  
NATIONAL/INTERNATIONAL  
CONTEXT

test development not enough attention is paid to the impact of cultural and linguistic factors that can be a source of bias for non-native speakers of the language of the test. We hope this paper helps to bring this issue to the attention of test publishers and policy makers to seriously consider these factors in the process of test development.■

Manuscript received: September 16<sup>th</sup>, 2008

Revised manuscript received: November 21<sup>st</sup>, 2008

## REFERENCIAS

- Abedi, J. (2006a). Psychometric issues in the ELL assessment and special education eligibility. *Teacher's College Record*, 108(11), 2282-2303.
- Abedi, J. (2006b). Language issues in item-development. In S. M. Downing & T. M. Haladyna (Eds.), *Handbook of Test Development* (pp. 377-398). Mahwah, NJ: Erlbaum.
- Abedi, J., & Gándara, P. (2006). Performance of English language learners as a subgroup in large-scale assessment: Interaction of research and policy. *Educational Measurement: Issues and Practice*, 25(40), 36-46.
- Abedi, J., & Lord, C. (2001). The language factor in mathematics tests. *Applied Measurement in Education*, 14(3), 219-234.
- Abedi, J., Hofstetter, C., & Lord, C. (1998). *Impact of selected background variables on students' NAEP math performance* (CSE Rep. No. 478). Los Angeles: University of California, National Center for Research on Evaluation, Standards, and Student Testing.
- Abedi, J., Hofstetter, C., & 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., Kim-Boscardin, C., & Larson, H. (2000). *Summaries of research on the inclusion of students with disabilities and limited English proficient students in large-scale assessment*. Los Angeles: University of California, National Center for Research on Evaluation, Standards, and Student Testing.
- Abedi, J., Lord, C., Hofstetter, C., & Baker, E. (2000). Impact of accommodation strategies on English language learners' test performance. *Educational Measurement: Issues and Practice*, 19(3), 16-26.
- Abedi, J., Lord, C., & Plummer, J. (1997). *Language background as a variable in NAEP mathematics performance* (CSE Technical Report 429). Los Angeles: University of California, Center for the Study of Evaluation, National Center for Research on Evaluation, Standards, and Student Testing.
- Adams, M. J. (1990). *Beginning to read: Thinking and learning about print*. Cambridge, MA: MIT Press.
- Bailey, A. L., & Butler, F. A. (2003). *An evidentiary framework for operationalizing academic language for broad application to K-12 education: A design document* (CSE Tech. Rep. No. 611). Los Angeles: University of California, National Center for Research on Evaluation, Standards, and Student Testing.
- Bever, T. (1970). The cognitive basis for linguistic structure. In J. R. Hayes (Ed.), *Cognition and the development of language* (pp. 279-353). New York: John Wiley.
- Biber, D. (1988). *Variation across speech and writing*. New York: Cambridge University Press.
- Carroll, J. B., Davies, P., & Richman, B. (1971). *The American heritage word frequency book*. Boston: Houghton Mifflin.
- Celce-Murcia, M., & Larsen-Freeman, D. (1983). *The grammar book: An ESL/EFL teacher's book*. Rowley, MA: Newbury House.
- Collier, V. (1987). Age and rate of acquisition of second language for academic purposes. *TESOL Quarterly*, 21(4), 617-41.
- Cummins, D. D., Kintsch, W., Reusser, K., & Weimer, R. (1988). The role of understanding in solving word problems. *Cognitive Psychology*, 20, 405-438.

## NOTAS

VALIDITY OF ASSESSMENTS  
FOR ENGLISH LANGUAGE  
LEARNING STUDENTS IN A  
NATIONAL/INTERNATIONAL  
CONTEXT



- Cummins, J. (1981). Four misconceptions about language proficiency in bilingual education. *NABE Journal*, 5(3), 31-45.
- Cummins, J. (2000). *Language, power and pedagogy: Bilingual children in the crossfire*. Clevedon, England: Multilingual Matters, Ltd.
- De Corte, E., Verschaffel, L., & De Win, L. (1985). Influence of rewording verbal problems on children's problem representations and solutions. *Journal of Educational Psychology*, 77(4), 460-470.
- deVilliers, J. & deVilliers, P. (1973). Development of the use of word order in comprehension. *Journal of Psychological Research*, 2, 331-341.
- Francis, W. N., & Kucera, H. (1982). *Frequency analysis of English usage: Lexicon and grammar*. Boston: Houghton Mifflin.
- Gathercole, S. E., & Baddeley, A. D. (1993). *Working memory and language*. Hillsdale, NJ: Erlbaum.
- Haiman, J. (1985). *Natural syntax: Iconicity and erosion*. New York: Cambridge University Press.
- Hakuta, K., Butler, Y. G., & Witt, D. (2000). *How long does it take English learners to attain proficiency?* Santa Barbara: University of California Linguistic Minority Research Institute.
- Hudson, T. (1983). Correspondences and numerical differences between disjoint sets. *Child Development*, 54, 84-90.
- Klesmer, H. (1994). Assessment and teacher perceptions of ESL student achievement. *English Quarterly*, 26 (3), 8-11.
- Larsen, S. C., Parker, R. M., & Trenholme, B. (1978). The effects of syntactic complexity upon arithmetic performance. *Educational Studies in Mathematics*, 21, 83-90.
- Lemke, J. L. (1986). *Using language in classrooms*. Victoria, Australia: Deakin University Press.
- Maihoff, N. A. (2002, June). *Using Delaware data in making decisions regarding the education of LEP students*. Paper presented at the Council of Chief State School Officers 32nd Annual National Conference on Large-Scale Assessment, Palm Desert, CA.
- Meskill (in press). Moment-by-moment formative assessment of second language development: ESOL professionals at work. In H. L. Andrade & G., J. Cizek (Eds.), *Handbook of Formative Assessment*. London: Routledge.
- Riley, M. S., Greeno, J. G., & Heller, J. I. (1983). Development of children's problem-solving ability in arithmetic. In H. P. Ginsburg (Ed.), *The Development of Mathematical Thinking* (pp. 153-196). New York: Academic Press.
- Rivera, C. (2003). *State assessment policies for English language learners*. Presented at the 2003 Large-Scale Assessment Conference.
- Rivera, C., Stansfield, C. W., Scialdone, L., & Sharkey, M. (2000). *An analysis of state policies for the inclusion and accommodation of English language learners in state assessment programs during 1998-1999*. Arlington, VA: The George Washington University, Center for Equity and Excellence in Education.
- Solano-Flores, G. & Trumbull, E. (2003). Examining language in context: The need for new research and practice paradigms in the testing of English Language Learners. *Educational Researcher*, 32(2), 3-13.
- Shaftel, J., Belton-Kocher, E., Glasnapp, D., & Poggio, J., (2006). The impact of language characteristics in mathematics test items on the performance of English language learners and students with disabilities. *Educational Assessment*, 11 (2), 105-126.

- Spanos, G., Rhodes, N. C., Dale, T. C., & Crandall, J. (1988). Linguistic features of mathematical problem solving: Insights and applications. In R. R. Cocking & J. P. Mestre (Eds.), *Linguistic and cultural influences on learning mathematics* (pp. 221-240). Hillsdale, NJ: Erlbaum.
- Thurlow, M., & Bolt, S. (2001). *Empirical support for accommodations most often allowed in state policy* (Synthesis Report 41). Minneapolis: University of Minnesota, National Center on Educational Outcomes.
- Wheeler, L. J. & McNutt, G. (1983). The effect of syntax on low-achieving students' abilities to solve mathematical word problems. *Journal of Special Education*, 17(3), 309-315.

**NOTAS**

VALIDITY OF ASSESSMENTS  
FOR ENGLISH LANGUAGE  
LEARNING STUDENTS IN A  
NATIONAL/INTERNATIONAL  
CONTEXT