Performance of English Language Learners as a Subgroup in Large-Scale Assessment: Interaction of Research and Policy

Jamal Abedi and Patricia Gándara, University of California, Davis

Factors that contribute to the performance gap between subgroups and mainstream students deserve special attention. Different subgroups are faced with different sets of challenges. To understand and control for factors leading to the performance gap between subgroups and mainstream students one must clearly understand the issues specific to each subgroup. This paper focuses on assessment and performance issues for English language learner (ELL) students as a subgroup. Identifying factors affecting the performance gap between ELL and non-ELL students may help gain insight into assessment issues for other subgroups of students as well as strengthen assessment of this group.

Keywords: assessment, language, English language learners, accountability, accommodation

Perspective

Current legislative mandates in the No Child Left Behind (NCLB Act of 2002) require the states to report performance of all students, including students in major ethnic groups, economically disadvantaged students, students with disabilities, and English language learners (ELL). While this requirement to report by major subgroups brings highly needed attention to achievement issues for these students, it also introduces a whole new set of technical and ethical issues in assessment.

Different subgroups are faced with different sets of challenges. Therefore, to understand and control for factors leading to the performance gap between subgroups and mainstream students, one must clearly understand the issues specific to each subgroup. Identifying factors affecting the performance gap between ELL and non-ELL students may help gain insight into assessment issues for other subgroups of students as well. Thus, the focus of this paper is on the specific issues associated with ELL academic performance.

English language learners now number over 4.5 million students; they represent about 8% of all K-12 students nationally (Zehler et al., 2003) and are the fastest growing subgroup in the nation (Kindler, 2002). Consequently, fairness and validity issues relating to their assessment must be among the top priorities of the national education agenda. ELL students have historically lagged behind their English proficient peers in all content areas, particularly academic subjects that are high in English language demand. The literature suggests that this performance gap is explained by many different factors including parent education level and poverty (Abedi, Leon, & Mirocha, 2003), the challenge of second language acquisition (Hakuta, Butler, & Witt, 2000; Moore & Redd, 2002), and a host of inequitable schooling conditions (Abedi, Herman, Courtney, Leon, & Kao, 2004; Gándara, Rumberger, Maxwell-Jolly, & Callahan, 2003) in addition to measurement tools that are ill-equipped to assess their skills and abilities.

English language learners are more likely to be taught by teachers without appropriate teaching credentials and with little classroom experience than other students (Rumberger & Gándara, 2004). A recent study of 4800 teachers of ELL students in California responding to a survey about the challenges they faced in teaching ELLs in their classrooms found large percentages of these teachers expressing the concern that they were not prepared to teach these students (Gándara et al., 2003).

While many ELL students are native born, many also have at least one parent who is an immigrant, and overwhelmingly they are reared in homes in which English is not the primary language spoken. Moreover, a disproportionately
high percentage of these students are raised in poverty and attend poor schools where they receive weak instruction (Gándara & Rumberger, in press). ELL students begin school significantly behind their English speaking peers and so require extra time and instruction from the very beginning of school to catch up (See Figure 1).

Within this context, ELL students must take on the challenges of learning English and U.S. culture, in addition to learning the academic content of the curriculum. The process of developing new language skills is difficult and requires time and effort that otherwise could be spent on acquiring academic content knowledge. For example, a number of studies have shown that it takes from five to seven years and even more for most ELLs to gain sufficient mastery of academic English to join English speaking peers in taking full advantage of instruction in English (Hakuta, Butler, & Witt, 2000). During this time, learning cannot occur at the same rate as for a native speaker of English when that instruction is offered only in English. Limited English proficiency may also make it difficult to benefit fully from the teacher’s instructions and to understand assessment questions.

Because ELL students, by definition, do not have a strong command of the English language, both learning and assessment are affected by their limited English proficiency, and both learning and assessment conditions must be addressed to help close the performance gap. Standardized achievement tests that have been constructed for mainstream students, but which do not take into account the special needs of English Learners, can provide a further challenge for these students, and may not provide a good indication of what these students know and can accomplish. Findings from studies suggest that a major factor contributing to the performance gap between ELL and non-ELL students is the impact of linguistic complexity of assessments (Abedi, 2006a; Abedi, Hofstetter, & Lord, 2004). Sometimes ELL students are provided with accommodations to help offset the linguistic challenge of the test, but these accommodations may not be relevant or helpful, and almost always fall short of ideal.

In this paper, we will discuss cognitive and non-cognitive (conative and affective) factors that contribute to a performance-gap between ELL and mainstream students, how such factors interact with academic achievement of ELL students and how research findings can help resolve some of these issues.


**Cognitive Factors**

1. **The Development of Reading Skills among English Language Learners**

   Because reading is implicated in most assessments of cognitive skill, and because it is the fundamental window through which we have access to cognitive processes, we begin with the literature on learning to read. The newly released report from the National Literacy Panel (August & Shanahan, 2006) calls attention to the problems of assessment for English language learners. They note that most assessments do a poor job of gauging individual strengths and weaknesses. The research on the development of English literacy strongly suggests that adequate assessments are essential for gauging the individual strengths and weaknesses of language minority students. . . . Unfortunately, existing assessments are inadequate to the need . . .
Theories about how second language learning takes place often identify factors that may impact ELL students’ assessment outcomes. For example, several cognitive processes are assumed to be imperative to the development of reading skills in English. These processes include phonological processing, syntactic awareness, and working memory (for a review of this literature, see Siegel, 1993). Although there is a dearth of research that examines whether these same processes characterize the reading development of English language learners (Lesaux & Siegel, 2003), we cite a few studies that are relevant to the topic of assessment of second language learners.

One study by Chiappe, Siegel, and Wade-Woolley (2002) investigated the processes involved in learning to read among a sample of 727 native English speakers and 131 English language learners. While the English language learners performed more poorly than native speakers of English on phonological and linguistic processing tasks, the development of literacy skills followed a similar trajectory for both groups. This suggests that while the process of learning to read may follow the same pattern, more time may be necessary for ELLs to actually acquire the necessary skills. The implications for test makers are that they must be cognizant of these factors when developing, administering, and scoring test items for ELL students.

Some research has focused specifically on the transfer of phonological awareness skills from the first language to the second language. Durgunoglu, Nagy, & Hancin-Bhatt (1993) examined 27 Spanish-speaking first graders to find whether phonological awareness in a child’s native language is related to their word recognition in the second language. In addition, the researchers investigated whether oral language competence in the second language impacts acquisition of reading in the second language. The results of this study indicated that Spanish phonological awareness predicted reading acquisition in the second language, suggesting that transfer of pre-reading skills can occur across languages. Moreover, as August and Shanahan (2006) conclude, “the research indicates that instructional programs work when they provide opportunities for students to develop proficiency in their first language” (Executive Summary, p. 5). These findings have particular implications for use of primary language assessments in order to better predict which students will need extra help in English reading. Unfortunately, policy in many states serving large numbers of English learners has worked against the development of strong primary language assessments.

Most recently, a study by Gottardo, Pannucci, Kuske, & Brettin (2003) showed evidence for simultaneous phonological processing in English and the primary language for English language learners. A sample of 65 children whose first language was Cantonese and second language was English was examined to determine the relationship between phonological skill and first and second languages. The study found that phonological skill was correlated between the first and second language. This is an interesting finding considering that Cantonese, unlike English, is not written in alphabetic orthography. This provides further evidence of the transfer of basic pre-reading skills across languages and the utility of assessing at least some skills in students’ primary languages.

The areas of syntactic awareness and working memory have been studied very little in connection to English language learners’ acquisition of reading. However, a study by Geva, Yaghoub-Zadeh, and Schuster (2000) briefly addressed syntactic awareness. The findings of this study indicated that when compared to native English speaking students, average English language learners in reading had a deficit in syntactic awareness skills. The lesson from this research appears to be that test makers would do well to reduce syntactic complexity in test items that will be used for ELL students, eschewing atypical sentence formats.

Working memory as well as oral English proficiency are two competencies that bode well for English reading development. In a study of working memory and second language acquisition conducted by Geva and Siegel (2000) among a sample of English speakers learning Hebrew, the researchers found that verbal memory was a significant predictor of basic reading skills. Likewise, August and Shanahan (2006) summarize one of the primary findings of the National Reading Panel as the critical importance of oral proficiency in English and in the primary language as precursors of strong reading skills in English. Again, this research suggests that the testing of verbal memory and oral proficiency in either the first or second language will yield important information about students’ readiness to undertake reading instruction in English.

While the summary presented above clearly shows the challenges that English language learners are faced with in learning a new language, it also suggests that students are transferring their pre-reading skills across languages and so the time they spend learning in any language is going to be roughly predictive of learning to read. As such, knowledge of what students know in both primary and secondary languages can yield important information about the cognitive processing skills they have for learning to read.

2. Language Factors and Assessment of ELL Students

Title III of No Child Left Behind Act (NCLB 2002) requires states to measure ELL students’ level of English proficiency (in addition to the Title I assessment requirements for these students) in four domains: reading, writing, listening, and speaking (and comprehension). Therefore, students’ level of reading proficiency obviously plays a major role in their assessment outcomes since without proficiency in reading, students will have difficulty understanding test questions. Proficiency in other domains is also essential for a valid assessment. In this section we use the term “English proficiency” in a broader sense to include various aspects of students’ level of proficiency.

There are many factors affecting academic progress of ELL students that are influenced by students’ level of English proficiency in general. For example, linguistic complexity of assessment may impact the performance outcome of these students and increase the performance gap between ELL students and their native English speaker peers. Invalid assessment results may affect decisions regarding classification and inclusion of ELL students and these all may impact the validity and fairness of the accountability system for these students. We first discuss the impact of language factors on the assessment outcomes of ELL students and then provide some research-based recommendations on how to improve assessment for these students by controlling for language factors that could impact the
validity of assessment for these students. We will also discuss language-based accommodations that would help ELL students with their English language needs.

Impact of language factors on the validity of assessment for ELL students. The National Research Council has warned that the use of achievement tests developed for English-speaking students will not likely yield valid results for students who are not proficient in English. As the NRC noted, “if a student is not proficient in the language of the test, her performance is likely to be affected by construct-irrelevant variance—that is, her test score is likely to underestimate her knowledge of the subject matter being tested (Huebert & Hauser, 1999, p. 225). We know that consistently low test scores often lead to placement in remedial and low-level instruction that further disadvantages already disadvantaged learners (Heubert, 2000) and exacerbates existing achievement gaps. Hence the consequences of such testing can be serious for these students.

Findings from several studies conducted recently show the impact of language factors on the assessment of ELL students (Abedi, 2006a; Abedi, Hofstetter, & Lord 2004; Maihoff, 2002; Kiplinger, Haug, & Abedi, 2000). These findings suggest that unnecessary linguistic complexity may hinder ELL students’ ability to express their knowledge of the construct being measured. Results from these studies question the appropriateness of mainstream assessments for English language learners, an issue that has long been debated. Educational psychologists, measurement experts, and educational practitioners have raised concerns over testing ELL students with assessment tools that are developed for native speakers of English, and which have been field tested mainly for these students:

For all test takers, any test that employs language is, in part, a measure of their language skills. This is of particular concern for test takers whose first language is not the language of the test. Test use with individuals who have not sufficiently acquired the language of the test may introduce construct-irrelevant components to the testing process. In such instances, test results may not reflect accurately the qualities and competencies intended to be measured … Therefore it is important to consider language background in developing, selecting, and administering tests and in interpreting test performance. (American Educational Research Association, American Psychological Association, & National Council on Measurement in Education, 1999, p. 91)

In the assessment of ELL students, research suggests that unnecessary linguistic complexity introduces a source of measurement error and is considered as a construct-irrelevant factor in the assessment (Abedi, 2006a; Abedi, Leon & Mirocha, 2003). Thus, language factors may seriously confound the outcomes of instruction and assessment in content-based areas (Abedi, Lord, & Hofstetter, 1998; Cocking & Chipman, 1988; De Corte, Verschaffel, & DeWin, 1985; Kintsch & Greeno, 1985; Larsen, Parker, & Trenholm, 1978; Lepik, 1990; Mestre, 1988; Munro, 1979; Noonan, 1990; Orr, 1987; Rothman & Cohen, 1989; Spanos, Rhodes, Dale, & Crandall, 1988).

The results of a series of studies on the impact of language on assessment of ELLs resulted in two major conclusions: (1) reducing the linguistic complexity of assessment tools helps ELL students perform significantly better because it reduced the performance gap between ELL and non-ELL, and (2) the process of reducing linguistic complexity of test items did not alter the construct under measurement. Therefore by reducing the linguistic complexity of assessment tools, the validity of assessment was not compromised (see Abedi, Lord, Hofstetter, & Baker, 2000; Abedi & Lord, 2001; Abedi, Courtney, Mirocha, Leon & Goldberg, 2005; Kiplinger, Haug, & Abedi, 2000; Abedi, Lord, and Hofstetter, 1998; Abedi, Lord, and Plummer, 1997).

In other studies on the impact of language factors on the assessment of ELL students, extant data from the National Assessment of Educational Progress (NAEP) were analyzed. Results (see, for example, Abedi, Lord, & Plummer, 1997) suggested that ELL students had difficulty with test items that were linguistically complex. The results also indicated that ELL students exhibited substantially higher numbers of omitted/not-reached test items.

Linguistic features that researchers identified in test items slow down the reader, make misinterpretation more likely, and add to the reader’s cognitive load, thus interfering with concurrent tasks. Indexes of language difficulty include word frequency/familiarity, word length, and sentence length. Other linguistic features that may cause difficulty for readers include long noun phrases, long question phrases, passive voice constructions, comparative structures, prepositional phrases, sentence and discourse structure, subordinate clauses, conditional clauses, relative clauses, concrete versus abstract or impersonal presentations, and negation (Abedi, Lord, & Plummer, 1997).

3. Language Factors and Accountability System for ELL Students

The impact of language factors on the assessment outcomes for ELL students has direct consequences on the accountability system for these students. Under the NCLB Title I accountability requirements, ELL students are assessed with the same tests that are developed and field tested mainly for native speakers of English. Earlier in this paper we discussed challenges that ELL students are faced with when taking these tests. Due to the complex linguistic structure of these tests, ELL students’ performance outcomes are very likely to be underestimated. Unfortunately, however, there is no simple approach to determine the level of underestimation that may occur due to the linguistic complexity of these tests.

The large performance gap between ELL and non-ELL students—which is caused by many factors, including parent education, poverty, the challenge of second language acquisition, and inequity in opportunity to learn—establish a substantially lower baseline score for ELL students. That is, from the very beginning, these students are clearly at risk of failing. It is quite clear that with starting low and progressing lower than their English proficient peers, ELL students cannot make the race for reaching the intended goal of 100% proficiency by the 2014 target date or even any time soon after that date.

Data from the California Department of Education for 2005 reveal how key subgroups are performing on NCLB accountability measures. Of the English learners tested in California, 79% were designated as not meeting proficiency standards. In addition, of the students with disabilities tested, 83% were classified as non-proficient (California Department of Education, 2006). NCLB also requires states to annually test English language learners on their progress toward gaining proficiency in
English, and some states have developed their own tests of English proficiency for this purpose. In California, the state developed test is the California English Language Development Test (CELDT). As concern over the lack of progress in closing the achievement gap between English language learners and English speakers has intensified, debates have ensued about the appropriate level of proficiency as measured by the CELDT at which to place students into the regular English-only curriculum (“basic”, “early advanced”, or “advanced”) (Parrish et al., 2006). Pressure to mainstream students at lower levels of proficiency has increased. Katz, Low, Stack, & Tsang (2004), in their study of San Francisco ELL students, also found little relationship between proficiency on the CELDT and students’ ability to navigate the content of English-only standardized tests. They argued that the CELDT and standardized tests of academic competencies are either measuring very different constructs or measuring the same constructs at very different levels of difficulty, but in either case, English proficiency levels established by English language proficiency tests such as CELDT should not be used to determine students’ readiness to take on English-only coursework that requires academic as opposed to social language. Gándara and Rumberger (in press) likewise found that 60% of tenth grade English language learners in 2005 were able to pass the CELDT at a level of early advanced or advanced English proficiency, but only 3% of these students could pass the state test of English Language Arts at a level of proficient—again suggesting that the tests measure different types of language proficiency, or have different standards of language proficiency.

Linn (in press) indicated that for large schools with diverse student populations, the disaggregated requirement can create major obstacles in reaching their AYP goal. “A school with a large enough number in, say, three racial/ethnic groups, students with limited English proficiency, economically disadvantaged students, and students with disabilities, would have a total of 29 hurdles to clear, 4 for each of the 6 subgroups plus the 5 that all schools have for the total student body” (p. 21). Sunderman and Kim (2004) expressed concerns over the racial and ethnic equity of NCLB in states with a large number of minority enrollments. Nichols, Glass, and Berliner (2005) found that states with a larger number and proportion of minority students are under more pressure. They concluded that problems associated with high-stakes assessment disproportionately affect minority students.

4. Accommodation Issues in the Assessment of ELL Students

To assist ELL students with their English language limitations, several accommodation strategies have been proposed. However, results of studies on the assessment and accommodations for ELL students have raised concerns over appropriateness of some of these accommodations. For example, researchers have found that many of the commonly used accommodations for ELL students are neither effective in helping ELL students with their language barriers nor are the results of these accommodated assessments valid (see, for example, Abedi, Lord, Hofstetter, & Baker, 2000; Abedi, Hofstetter, & Lord, 2004; Sireci, Li, & Scarpati, 2003).

To investigate the relevance of accommodations used for ELL students, Abedi (2006b) analyzed a list of 73 accommodations that have been used nationally in the assessment of ELL students (Rivera, 2003). The results indicated that of the 73 accommodations used by different states only 11 (or 15%) of them were highly relevant for these students. The main questions here are: How do states decide what accommodation to use for English language learners and students with disabilities? What type of criteria they use for such decisions? Do states use research findings as criteria in their decisions?

To answer these questions and to understand the states’ policies and practices regarding inclusion and accommodations for students with disabilities (SDs) and ELL students, eight states with large numbers of ELLs and/or students with disabilities were surveyed (Pitoniak, Lutkus, Cahalan-Laitusis, Cook, & Abedi, 2006.) State testing directors were asked to identify criteria used in their decision about the type of accommodations offered to students and about their policies regarding inclusion of both SDs and ELLs in NAEP and state assessments. Responses of the participating states across the two assessments (state and NAEP) were similar. Therefore, we discuss the results of these interviews regarding state assessments both in terms of decisions regarding accommodations and the inclusion of students. In this paper, we limit our discussion to the results for ELL students.

In general, the interview results suggest that decisions regarding both inclusion of and accommodations for ELL students were mainly based on the guidelines and recommendations from the state department of education. All of the eight participating states in this interview indicated that accommodations for ELL students are “determined by following guidelines provided by the state department of education” and only two of the eight indicated that “accommodations must be language related to assist LEP students with their language needs.” This raises serious concerns over the appropriateness of accommodations for these students.

It is important to examine the pattern of responses in these interviews in terms of recent research findings in the area of accommodations and inclusion of English language learners. It does not seem that research findings greatly influenced the decision making process. For example, when states were asked whether they determine accommodations based on opinion or research evidence, only two of the eight states indicated that they determine NAEP accommodations based on research evidence.

When states were asked about specific research evidence affecting their decisions regarding accommodations and inclusion of ELL students, a majority of them did not find research evidence to influence their decisions. There are, however, some examples of research findings that have a direct bearing on issues of accommodation and inclusion. Results from experimentally controlled studies suggest accommodations that are language-based or consistent with students’ language needs are more effective and valid for ELL students than those originally created and proposed specifically for students with disabilities (see, for example, Abedi, Lord, Hofstetter, & Baker, 2000; Abedi, Hofstetter & Lord, 2004). In contrast, the interviews show only two of the eight states indicated that they use accommodations that address ELL students’ language needs. Similarly, research evidence supports the use of the length of time students have been instructed in the United States as well as level of English proficiency.
as valid criteria for decisions regarding appropriate accommodations for ELL students (Abedi, Lord, Hofstetter, & Baker, 2000; Abedi, Hofstetter, & Lord, 2004). The results from the interviews suggested that only three of the eight states reported using such criteria for selecting appropriate accommodations for ELL students.

5. Research Findings that may Help Improve the Validity of Assessment and Accountability Systems for ELL Students

Results from studies on the assessment and accommodations of ELL students can help to identify assessment issues and provide research-based recommendations for improving assessment and accountability systems for ELL students. For example, research on the impact of language factors on the assessment of ELL students demonstrates that unnecessary linguistic complexity of assessment may jeopardize the validity of assessment and accountability system and provides suggestions for improving the assessment and accountability systems by reducing the unnecessary linguistic complexity of assessments.

Abedi and Lord (2001) examined the effects of the linguistic modification approach with 1,631 eighth-grade students in Southern California. In this study, NAEP mathematics items were modified to reduce the complexity of sentence structures and to replace potentially unfamiliar vocabulary with more familiar words. Content-related terminologies, i.e., mathematical terms, were not changed. The results showed significant improvements in the scores of ELL and non-ELL students in low and average-level mathematics classes, but no effect on scores among other non-ELL students. Among the linguistic features that appeared to contribute to the differences were low-frequency vocabulary and passive voice verb constructions.

In another study, Abedi, Lord, and Hofstetter (1998) examined the impact of language modification on the mathematics performance of English learners and non-English learners on a sample of 1,394 eighth graders in schools with a large number of Spanish speakers enrolled. Results showed that modifications to the language of items contributed to improved performance on 49% of the items. The ELL students generally scored higher on shorter/less linguistically complex problem statements. In another study on a sample of 946 eighth graders Abedi, Lord, Hofstetter, & Baker (2000) found that among four different accommodation strategies for ELL students, only the linguistically modified English form narrowed the score gap between English learners and other students.

Another study (Abedi, Courtney, & Leon, 2003) examined 1,594 eighth grade students using items from NAEP and TIMSS. Students were given one of the following accommodations: a customized English dictionary (words were selected directly from test items), a bilingual glossary, a linguistically modified test version, or the standard test items. Only the linguistically modified version improved the ELL students’ scores without affecting the non-ELL students’ scores.

Other researchers have employed a language modification approach based on CRESST studies. Maihoff (2002) found linguistic modification of content-based test items to be a valid and effective accommodation for ELL students. Kiplinger, Haug, and Abedi (2000) found linguistic modification of math items to help improve the performance of ELL students in math. Rivera and Stansfield (2001) compared English learner performance on regular and simplified fourth and sixth grade science items. Although the small sample size did not show significant differences in scores, the study did demonstrate that linguistic simplification had no effect on the scores of English-proficient students. This indicates that linguistic simplification is not a threat to score comparability.

Research findings can also help identify accommodations that make assessments more accessible to ELL students. The results from several studies have demonstrated that some accommodations may impact the validity of assessments. For example, some forms of accommodation strategies, such as glossary use plus extra time, raised performance of both ELL and non-ELL students. The level of increase due to such accommodation strategies was higher for non-ELL students. This raised concerns regarding the validity of these accommodations (Abedi, Lord, Hofstetter, & Baker, 2000; Abedi, Lord, & Hofstetter, 1998).

English and bilingual dictionaries were used as different forms of accommodation strategies. The results of these studies suggested that by gaining access to the definition of content-related terms, recipients of dictionaries may be advantaged over those who did not have access to dictionaries. This may jeopardize the validity of assessments. The dictionary as a form of accommodation suffers from yet another major limitation, the feasibility issue. For a variety of reasons, it was logistically difficult to provide this form of accommodation to students (Abedi, Courtney, Mirocha, Leon, & Goldberg, 2005).

There are however, some accommodations that help ELL students with their English language needs that do not affect the performance of non-ELL students. For example, studies by CRESST researchers along with other studies nationwide suggested that linguistically modified versions of tests items are an effective and valid accommodation for ELL students (Abedi, Lord, & Hofstetter, 1998; Abedi, Lord, Hofstetter, & Baker, 2000). Duncan et al. (2005), using NAEP mathematics test data for 402 Spanish-speaking eighth-grade students, found that the students appreciated and preferred dual-language test booklets, with no undue advantage using the booklet. These investigators found the method hopeful, but cautioned that it was important to do more study to determine for which categories of ELLs the dual-language test booklets appear to be most useful.

Non-Cognitive Factors

1. Communicative Competence and Socio-cultural Factors

In this article, we use the concept of linguistic complexity in a more comprehensive manner, as it interacts with a complex set of social, cultural, and psychological factors. We believe English language learners who inadvertently suffer the most from language complexity in instruction and assessments are not only different from their native English speaking peers in their language background, but in their cultural, family, and personal characteristics. We believe, as research results in the past 10 years have clearly suggested, that linguistic and cultural factors may hinder a student’s ability to fully understand the instructional materials and do well on assessments. Celce-Murcia’s Model of Communicative Competence (Celce-Murcia, 1995) provides a comprehensive view of linguistic and cultural issues that may affect students’ academic performance.
Celce-Murcia, Dornyei, & Thurrell (1995) proposed a model which is an elaboration of the earlier communicative competence model (Canale, 1983) modified by Celce-Murcia (1995). This model is based on the following six components: linguistic competence, strategic competence, discourse competence, socio-cultural competence, formulaic competence, and paralinguistic competence (see also Bachman, 1990; Bachman & Palmer, 1996). Celce-Murcia (1995) emphasizes the dynamic aspect of her model and indicates that the different components in the model interact with each other:

It should be emphasized that the interactions and overlaps schematically represented in Figure 3 [Figure 2 in this article] are integral aspects of the proposed model of communicative competence. It is not sufficient simply to list all the components as was done in Figure 2; it is important to show the potential overlaps, interrelations and interactions, and to realize that discourse is where all the competencies most obviously reveal themselves. Discourse thus is the component in which (or through which) all the other competencies must be studied—and ultimately assessed—if one is concerned with communicative competence, which is not a hierarchical system of discrete competencies or abilities but a dynamic, interactive construct. (p. 704)

Figure 2 shows Celce-Murcia’s model. In this model, socio-cultural competence represents the speaker/listener’s background knowledge of the target community (e.g., understanding communications, beliefs, values, conventions, and taboos of the target community), which makes informed comprehension and communication possible. The linguistic/grammatical competence and formulaic competence complement each other and allow the user to express original and creative messages or to draw upon conventionalized prepackaged and mundane formulaic messages. A comprehensive view of the communicative competence model (CCM) and its related issues is beyond the scope of this paper. However, the main reason for presenting this brief introduction to CCM is to indicate how complex the role of language factors is in the instruction and assessment of ELL students and how different factors interact within the framework.

2. Affective Factors in Learning and Assessment

There is a multitudinous literature on the role of affect and learning. Bandura’s (1993) very large body of work on social learning theory is based on the notion that learners who feel they have control over their own learning are more apt to excel and that the sense of control is bound up in perceptions of self-efficacy—how they feel about themselves and their abilities. Likewise, Nolen-Hoeksema, Gurgus, & Seligman (1986) showed how a lack of feeling of self-efficacy could lead to “learned helplessness,” resulting in a failure to even try to accomplish tasks that are perceived as too difficult—whether in fact they are or not. English Learners are fundamentally learners, like all other students, and so are vulnerable to the same social and psychological factors that affect learning. If they come to feel that they are not capable of learning either a new language or a new content area, they will falter. Similarly, lack of confidence in one’s skills has been shown to affect test performance as well. Steele (1997), through a series of novel experiments, demonstrated that stereotypes or beliefs about how others view their ability can affect the test performance of both women and minorities. Often subjects are told that others do not believe they are particularly strong in the material being tested, and this results unwaveringly in depressed test performance. It is difficult for the test maker to account for these affective influences on test performance in the building of a test, but the interpretation of test results can certainly take into account the error that arises from these affective factors in particular learners, such as those who are still learning English.

High levels of pressure can also negatively impact learning. The level of pressure on all students, and particularly on ELL students due to language barriers, increases as the stakes associated with assessment increases. For
example, ELL students, their teachers and their parents may be blamed partly when their schools are not meeting the AYP requirements. Nichols, Glass, and Berliner (2005) indicated that the theory of action underlying the NCLB accountability system is that the pressure of high-stakes testing will increase student achievement. The authors created the Pressure Rating Index (PRI) and correlated the PRI outcome with the NAEP results from 1990 to 2003 in 25 states. They found that “States with greater proportions of minority students implement accountability systems that exert greater pressure” and “High-stakes testing pressure is negatively associated with the likelihood that eighth and tenth graders will move into twelfth grade.” They also found that “increases in testing pressure are related to larger numbers of students being held back or dropping out of schools” (p. ii).

3. Conative Factors: Motivational and Volitational Aspects of Learning

Non-cognitive factors—such as attitudes, beliefs, and motivation—have been shown to be related to student achievement (Schreiber, 2002). For instance, using an attitudes inventory, Tapia and Marsh (2000) found that students failing in mathematics scored lowest on self-confidence, motivation, value, and enjoyment.

Using nationally representative data of eighth-grade students from the National Education Longitudinal Study of 1988, Singh, Granville, & Dika (2002) found significant positive effects of motivation factors. Research has found that students who struggle with reading grow to have less desire and motivation to read as a direct result of their struggles (O’Brien & Dillon, 2002).

Snow (1994) advocated that cognitive processes alone could not account for students' learning, but that both affective and conative (motivational) processes had to be viewed as interacting with cognitive processes to produce learning. There is considerable literature to demonstrate this point in the research on second language learning. A number of studies (see Bialystok & Hakuta, 1994) have shown that individuals who are motivated to learn a language, whether by necessity or desire, are more effective learners of that language. Likewise, we know that the context of second language learning can have powerful effects on motivation, as when the learner perceives that she is being asked to reject her identity by displacing her primary language with a new language. In such cases, second language learning can be retarded (Schmidt, 2000). It was, in fact, this very principle—that students who were made to feel that their primary language and culture had value in the society would be more motivated to learn a second language than those who felt their native language threatened—that was behind the development of many bilingual programs during the 1970s. Adolescents, especially those who may be acutely concerned about how they appear as they struggle with a second language, can experience embarrassment that leads to lack of motivation to use the language, which in turn delays development of the second language. (Tse, 2001).

Just as language learning is affected by motivation, so too can test performance be affected. Students who do not “care” about learning are not likely to “care” about how they perform on tests of that learning. Gándara et al. (2003) found that many Latino high school students (especially males) who felt marginalized in their schools also professed not to care how they did in school and routinely failed their courses and made little or no effort to study for tests.

Given these findings, it is important to examine how policy issues regarding assessment measures consider possible differences in cognitive, conative, and affective factors for the ELL subgroup. Recent studies on the assessment and accommodations of English language learners and students with disabilities have raised concerns over a variety of issues related to academic assessment of these students.

Discussion

Fairness demands equal educational opportunities for all students including ethnic minorities, students with disabilities, low-income students, and English language learners. However, these subgroups of students have historically lagged behind their mainstream peers on test scores due, in part, to factors that may not be closely related to their academic achievement but do influence their performance outcomes. In order to begin reducing the performance gaps between subgroups and mainstream students, improvements must be made in both instruction and assessment by identifying and controlling for these factors (often referred to as nuisance variables) that affect a subgroup's academic achievement.

The four subgroups identified in the NCLB legislation are different in their background and are faced with different issues in their academic careers. While such differences are relevant to consider, in this paper we focused on assessment and accountability issues for ELL students, a key subgroup of students that is faced with serious challenges. Understanding issues that are involved in the instruction and assessment of ELL students will help us realize the limitations of the nation's current accountability measures for all the subgroups under consideration.

Many different factors such as parent education, poverty, and schooling conditions contribute to the existing performance gap between ELL and non-ELL students. However, language factors have a greater impact on ELL student performance than any other factors. While ELL students perform behind non-ELL students in almost all subject areas in school, the performance gap between ELL and non-ELL is largest in areas with higher levels of language demand. Because ELL students, by definition, do not have a strong command of the English language, both learning and assessment are affected by their limited English proficiency, and both learning and assessment conditions must be addressed to help close the performance gap.

The concept of language as it affects ELL students' performance is truly multidimensional and complex in nature. As elaborated in Celce-Murcia's model of communicative competence, language interacts with many different socio-cultural factors. By understanding the complex nature of language factors, one can then discuss the impact of language on the academic achievement of ELL students.

The challenges of second language acquisition, such as learning in a language in which students are not quite proficient and working with measurement tools that are ill-equipped to assess their skills and abilities, have much more impact on ELL students' academic performance than is acknowledged by many researchers and educational practitioners. For example, as elaborated earlier in this paper, when compared to native English speaking students in reading, average English
language learners had a deficit in syntactic awareness skills.

Furthermore, the process of developing new language skills is difficult and requires time and effort. Thus, it takes a long time for ELL students to become proficient enough in English to understand teachers' instructions and test questions in a language with an unfamiliar structure and vocabulary. Based on some research evidence it may take from five to seven years or longer for most ELLs to gain sufficient mastery of academic English to join their English speaking peers in taking full advantage of instruction in English. During this period, ELL students may not be able to learn at the same rate as their native speaking peers when the instruction is offered only in English. Limited English proficiency may also make it difficult to benefit fully from the teacher’s instructions and to understand assessment questions. This has a direct impact on these students’ performance on assessments.

It is ironic that during the period that ELL students are struggling to learn English, they are expected to progress at the same rate as their native English speaking peers. In addition, their schools and teachers may receive punitive corrective action for a lack of adequate progress. Perhaps one of the most challenging tasks in ELL students’ academic life is dealing with the often complex linguistic structure of assessment tools. Standardized achievement tests that have been constructed for mainstream students do not take into account the special needs of English learners. This can be a major source of frustration for these students and can contribute substantially to the performance gap between ELL and non-ELL students as there is no evidence to suggest that the basic abilities of ELL students are different from non-ELL students.

Sometimes ELL students are provided with accommodations to help offset the linguistic challenges of assessment instruments, but many of these accommodations may not be relevant for these students since they were originally designed for other subgroups with completely different backgrounds and needs. That is, there is not enough evidence that these accommodations make the assessments more accessible for ELL students. Moreover, results of experimentally controlled studies show that some accommodations may alter the construct being measured which consequently may affect the validity of the assessment.

Problems in the assessment and instruction of ELL students, and the impact language background variables have on students’ learning and achievement outcomes undermine the authenticity of the nation’s accountability system for ELL students. The accountability system developed under NCLB has the expressed intent to pressure schools to be more responsive to the needs of at-risk students who have historically fared poorly in America’s public schools. By exerting pressure through sanctions, the belief has been that schools can be made to help these students improve academically. This is certainly a laudable and important goal, although there is far from universal agreement on the degree to which this strategy can achieve the goal. As Linn (2003) indicated, NCLB offers many contributions to the nation’s accountability system. For example, special attention given to groups of students who have had the lowest achievement in the past is encouraging. However “the goals that NCLB sets for student achievement would be wonderful if they could be reached, but, unfortunately, they are quite unrealistic” (Linn, 2003, p. 10).

It is important to consider that there are many other factors beyond the cognitive domain in general and language factors in particular that are affecting the academic achievement of ELL students. Conative factors (motivational and volitional aspects of learning) as well as affective factors can greatly impact students’ performance in school. Lack of progress due to serious challenges that ELL students are faced with may seriously impact their level of motivation. Undue pressure to perform well under the nation’s accountability system and other sources of pressure in the academic life of ELL students may cause frustration and impact their academic performance.

Research findings can help to resolve some of the major issues in the instruction and assessment of ELL students and consequently improve the accountability system for these students. While research in this area is very limited in content and scope, and some of the findings may be inconclusive, some of these findings can help improve the existing assessment and accountability systems. For example, by identifying the sources of linguistic complexity of assessments, and by providing ways to reduce the level of unnecessary linguistic complexity, research findings have contributed significantly to improving assessments for ELL students.

Similarly, by identifying factors that impact the effectiveness and validity of accommodated assessments and by providing research-based recommendations for selecting the most appropriate accommodations for ELL students, research findings can help create a more effective assessment system for ELL students. For example, as explained in this paper, research findings have identified some accommodations that help ELL students with their English language needs without compromising the validity of the assessment.

In this paper we discussed the challenges that English language learners are faced with in learning a new language. We also elaborated on the fact that ELL students transfer their pre-reading skills across languages. This implies that the time they spend learning in any language is going to be roughly predictive of learning to read. Therefore, knowledge of what students know in both primary and secondary languages can yield important information about the cognitive processing skills they can apply to learning to read. Including such information in the assessment and accountability systems for these students would be helpful.

Instruction and assessments for ELL students are affected by many different factors; therefore, it is not possible to present a comprehensive picture of all of the issues involved in the educational experiences of these students. In this article we tried to present a few tangible examples of factors influencing the academic achievement of these students. We hope this discussion might help policy makers, educational researchers, and educational practitioners to understand the complex nature of the academic life of ELL students in particular, and all four subgroups identified in NCLB in general, and plan their curriculum, instruction, and assessment accordingly.

References


Educational Studies in Mathematics, 21, 83–90.