English Language Learners with Disabilities: Classification, Assessment, and
Accommodation Issues

Jamal Abedi
CRESST/University of California, Davis
National Center for Research on Evaluation, Standards, and Student Testing

Author's Note

Correspondence concerning this article should be addressed to: Jamal Abedi, School of Education, University of California, Davis, One Shields Avenue, Davis, CA, 95616.

English Language Learners with Disabilities: Classification, Assessment, and Accommodation Issues

Abstract

English language learners with disabilities (ELLWD) face many challenges in their academic career. Learning a new language and coping with their disabilities create obstacles in their academic progress. Variables relegating accessibility of assessments for students with disabilities and ELL students may seriously hinder the academic performance of ELLWD students. Furthermore, classification and accommodation for these students requires a more complex design than those for either ELLs or students with disabilities. Proper identification of these students is a challenge if their disability is masked by their limited English proficiency, or vice versa. Improper identification may lead to inappropriate instruction, assessment and accommodation for these students. Linguistic and cultural biases may affect the validity of assessment for ELLWD students. In this paper, issues concerning accessibility of assessment, classification, and accommodations for ELLWD students are discussed and recommendations for more accessible assessments for these students are provided.

Perspective

The Office of English Language Acquisition estimated the number of English language learners with disabilities (ELLWD) in K-12 to be over 350,000 students, which represents 9% of all ELL students, and 8% of all children in special education (National Symposium on Learning Disabilities in English Language Learners, 2004; Zehler, Fleischman, Hopstock, Pendzick, & Stephenson, 2003). Aldus & Thurlow (2005) found little reported data on this population of students. These students need more attention in their academic career due to dual challenges they are facing--the combination of limited English proficiency and a disability or disabilities. Thus, it is imperative to closely examine issues concerning classification, assessment and accommodation for ELLWD students and provide research based recommendations for enhancing the academic life of these students.

Several major components set the foundation for education of ELLWD students. They include: classification, instruction, assessment, and accommodation for these students. There are major issues in these areas for ELLWD students. Literature clearly shows validity concerns in the classification system for both ELLs (Abedi, 2008) and students with disabilities (Reschly, 1996; Ysseldyke, & Bielinski, 2002). ELLWD students suffer from a combination of issues concerning validity of classification; those related to classification of ELLs and those specific to classification of students with disabilities (SWD). Classification issues may cause problems in curriculum, assessment and accommodation for these students. Furthermore, there are issues concerning opportunity to learn for these students (Abedi & Herman, in press). The focus of this paper is on classification, assessment and accommodation issues for ELLWD students.

Classification

Researchers expressed concerns over the validity of classification for ELLs and students with disabilities. Abedi (2008) found that the criteria used for classification of ELL students suffer from major validity shortcomings. For example, the study found that students' level of proficiency in English, which must be the foundation of any ELL classification system, is not a major determinant of ELL classification. The results indicated that less that 10% of the variance of ELL classification was explained by students' levels of English proficiency. The study also revealed that many factors other

than students' levels of English proficiency determine the outcome of classification systems for ELL students. These variables include ethnicity, socio-economic status, and teacher and parent opinion.

Research also points to issues in classification of ELL students with disabilities (Reschly, 1996; Ysseldyke, & Bielinski, 2002). Classification of ELLWD students requires special attention as misclassification directly impacts the instruction, assessment, and accommodation for these students. ELLWD students may not be properly identified as having a learning disability, if their disability is masked by their limited English proficiency. On the other hand, ELL students at the lower level of English proficiency spectrum with no disabilities may be misclassified as students with learning disability due to their limited proficiency in English.

Literature shows ELL students with lower levels of proficiency in their native language (L1) and English (L2) have the highest rate of identification in the special education categories. In addition, more ELL students tend to be placed in the "learning disability" category than in the "language and speech impairment" category (Artiles, Rueda, Salazar, and Higareda, 2005; Ortiz, 2002). Artiles and Ortiz (2002) found a differential rate of overrepresentation of ELL students in special education programs in some states. For example, based on their data, 26.5% of ELLs in Massachusetts, 25.3% in South Dakota, and 20.1% in New Mexico were placed in a special education program. This was compared to less than 1% of ELLs in Colorado, Maryland, and North Carolina that were placed in similar programs. Rueda, Artiles, Salazar, and Higareda (2002) reported that over a 5-year period (1993-1994 to 1998-1999), the placement rate of Latino English language learners increased by 345% while their overall population in the district increased by only 12% during this period of time. It must be noted at this point that some of the variability between states discussed above could be explained by the characteristics of specific ELL populations. Such variability might also be explained by the way resources are allocated to ELL versus students with disabilities.

Minnema, Thurlow, Anderson, & Stone (2005) indicated that "the classification of an English language learner with disabilities remains elusive in part due to the lack of consensus in the field on who an English language learner actually is and how to refer to these students" (p. 10). The authors indicated that "…a major concern rests on English

language proficiency measures and definitions of language proficiency, which is confounded even more when paired with a disability" (Minnema et al., 2005, p. 10).

A study by the National Accessible Reading Assessment Projects (Abedi et al., 2008) provided evidence on misclassification of ELL students at the lower level of English proficiency as students with disabilities. In this study, to make reading tests more accessible for students with disabilities, long reading comprehension passages were parsed into three to four segments and questions for each of those segments were placed immediately after each segment. Therefore, students did not have to read the entire passage to answer many of the reading comprehension questions related to the passage. This study was conducted on a group of 738 students, 117 of them were classified as students with disabilities.

Of the total 117 students with disabilities in this study, 79 or 68.1% were ELL and 38 or 31.9% were non-ELL. In contrast, only 22.5% of students without disabilities were English language learners. This disproportional rate of ELL students represented in the learning disability category (68.1% of students with disabilities versus 22.5% of students without disabilities) illustrates major issues in classification of ELL students as students with learning disabilities. This trend, which is consistent with the literature (see for example, Abedi, 2006a; Artiles, Rueda, Salazar, & Higareda, 2005), indicates that ELL students at the lower level of English proficiency have a much higher chance of being misclassified as having a learning disability than non-ELL students.

The summary of literature presented above suggests that classification of ELLWD students needs serious attention. Improper classification of ELLWD students may render assessment results unfair, invalid, and ineffective. Invalid assessment may then lead to inappropriate and inadequate instruction for these students. The first step in improving the quality of education for ELLWD students is to develop and validate a classification system that provides accurate results for these students.

<div align="center">Assessment</div>

Literature shows a substantial performance-gap between ELL and non-ELL students (see, for example, Abedi, 2008; Abedi, Leon & Mirocha, 2003; Abedi, 2004). Many different factors contribute to such a performance-gap, among the most influential of which are linguistic and cultural factors. Linguistic complexity of test items makes them more difficult for ELL students to comprehend and contributes to a performance-gap between ELL and non-ELL students.

The literature also documents a major performance-gap between students with disabilities and regular students (Le Floch, Martinez, O'Day, Stecher, Taylor, & Cook, 2007; Perie, Grigg, & Donahue, 2005; Thurlow & Liu, 2001). As a result of facing dual challenges, the academic performance of ELLWD students is expected to fall far behind both groups: students with disabilities and ELL students.

The higher the level of language demands of assessments, the larger the performance gap between ELL and non-ELL students. For example, Abedi, Lord and Plummer (1997) showed that ELL students had more difficulty with linguistically complex test items regardless of the items' content difficulty. This same study found that ELL students also had a substantially higher number of omitted/not reached items due to difficulty understanding the test items (see also, Abedi, 2004; Abedi, Hofstetter & Lord, 2004; Abedi, 2006b; Maihoff, 2002; Solano-Flores & Trumbull, 2003). This situation becomes even more dramatic when ELL students have disabilities--particularly those with learning disabilities.

To illustrate the disparity between performances of ELLWD students compared to their peers (non-ELL/non-SWD), we present summaries of analyses of data from two different states with large numbers of ELL students. The large number of ELL students in each state allows for detailed analyses of the data.

Due to the confidentiality agreement with data providers, we refer to these two states as Site 1 and Site 2. Site 1 provided pre-NCLB data for Grades 3 and 8 in reading and math and Site 2 provided post-NCLB data in reading and math for students in Grades 5 and 8. It is critical to note that data obtained from the two sites are quite different in many ways. They are from different assessments with different contents and test items and different scale scores. Therefore, a direct comparison between the data from pre-NCLB and post-NCLB may not be possible. We report performance gaps between subgroups of students in terms of percent of difference which we refer to as the Disparity Index (DI).

We were also interested in examining possible changes in the performance of ELLWD students over time and possible changes due to the impact of the No Child Left Behind legislation (NCLB, 2002) on ELLWD students since one of the major goals of the

NCLB legislation was to improve the quality of education for all students particularly those with a challenging academic career such as ELLWD students.

Pre-NCLB Data

Table 1 presents summaries of analyses of data from Site 1 (pre-NCLB) in reading and math for students in Grade 3. This table reports the Stanford Achievement Test, version 9 (SAT9) by the three subgroups (ELL, SWD, ELLWD) and the reference group (non-ELL/non-SWD) for Grade 3 students in reading and math. Test scores are reported in the form of a NCE (Normal Curve Equivalent) scale score with a mean of 50 and standard deviation of 21.06 (Linn & Gronlund, 1995). Descriptive statistics including mean and standard deviation of reading and math tests are reported for ELL-only, SWD-only, ELLWD and the non-ELL/non-SWD (the reference group). Table 1 also reports reliability coefficients, standard error of measurement (SEM) and the correlation between reading and math test scores for each of the three subgroups (ELL, SWD, and ELLWD) as compared with the reference group.

Table 1.
Descriptive statistics by subgroups of students for Grade 3 Students in Reading and Math (pre-NCLB)

Site 1, Grade 3 Reading					
ELL/ SWD Status	Mean	Std. Deviation	DI	Alpha Coefficient (SEM)	Correlation of Reading & Math
ELL-ONLY	27.99	16.37	-53.4	.82 (6.95)	.65
SWD-ONLY	27.46	22.38	-56.4	.79 (10.26)	.64
ELLWD	13.93	13.37	-208.0	.70 (7.32)	.52
Non-ELL/non-SWD	42.94	19.59		.89 (6.49)	.71
Total	40.69	20.29		.85(7.86)	.66
Site 1, Grade 3 Math					
LEP/ SWD Status	Mean	Std. Deviation	DI	Alpha Coefficient (SEM)	Correlation of Reading & Math
ELL-ONLY	41.11	18.55	-25.7	.84 (7.42)	.66
SWD-ONLY	38.57	23.09	-34.0	.83 (9.52)	.65
ELLWD	25.71	14.39	-101.0	.80 (6.44)	.50

Non-ELL/non-SWD	51.68	20.76		.91 (6.23)	.75
Total	49.93	21.18		.87	.70

As indicated earlier, comparisons were made between performance of the subgroups (ELL, SWD, and ELLWD) and the reference group (non-ELL/non-SWD) using a DI index. A DI is simply the percent of disparity between the performance of a subgroup with the reference group which is defined by subtracting the mean of the reference group from the mean of the subgroup, divided by the mean of the subgroup multiplied by 100. A negative DI suggests lower performance of the subgroup when compared with the reference group. A positive DI on the other hand indicates higher performance for the subgroup when compared with the reference group.

As data in Table 1 show, the performance-gap between ELLWDs and the reference group in both reading and math is substantially higher than the performance-gap between either ELL and the reference group or SWD and the reference group. For example, the performance-gap between ELLs and the reference group in reading for Grade 3 students is -14.95 (27.99-42.94), and for the SWDs subgroup, the performance gap is -13.23 as compared with a performance-gap of -26.76 for ELLWD students. Converting these performance gaps into a DI provides data in the form of percent of difference which is easier to read. As Table 1 shows, the DI for ELL students (comparing ELL with the reference group, non-ELL/non-SWD), is -53.4 which suggests that ELL students underperform the reference group by 53.4% ([27.99 – 42.94/27.99] * 100). Similarly, the DI for SWD is -56.4 which suggests that SWD students underperform the reference group by 56.4% as compared with the DI of -208 for ELLWD students suggesting that they underperformed the reference group by 208%.

Similarly, in math the performance gap between ELLWD students and the reference group is larger as compared with the performance gap for either ELL or SWD. However, these performance gaps are generally lower than those reported above for reading. The DI for ELL students is -25.7, for SWD, -34.0 and for ELLWD students it is -101.0. These results present two important points: (1) ELLWD students perform substantially lower than the other subgroups in all content areas and (2) such a performance gap is larger in content areas with more language demands. Thus, the gap between the three subgroups and the reference group is substantially lower in math than

in reading. This shows the serious impact of language on the academic performance of these students (ELLs, SWDs, and ELLWD).

Table 2 presents data for grade 8 students. Similar to the data presented in Table 1 for Grade 3 students, the performance-gap between ELLWD and the reference group is much larger than the performance-gap for the other two subgroups (ELL and SWD). For ELL students the DI for reading was -123.4 (i.e., ELLs performed 123.4% less that non-ELL/non-SWD). The DI for SWDs was -135.5 and for ELLWD students the average DI was -372.9. Once again while all the three subgroups performed substantially lower than the reference group in reading, the gap in performance of ELLWD students was substantially larger than the gap of the other subgroups.

Table 2.
Descriptive statistics by subgroups of students for Grade 8 Students in Reading and Math (pre-NCLB)

Site 1, Grade 8 Reading					
ELL/ SWD Status	Mean	Std. Deviation	DI	Alpha Coefficient (SEM)	Correlation of Reading & Math
ELL-ONLY	20.53	16.37	-123.4	.82 (6.95)	.66
SWD-ONLY	19.47	19.86	-135.5	.80 (8.88)	.68
ELLWD	9.70	11.56	-372.9	.73 (6.01)	.49
Non-ELL/non-SWD	45.87	21.04		.89 (6.98)	.70
Total	41.91	22.70		.85(8.79)	.68
Site 1, Grade 8 Math					
LEP/ SWD Status	Mean	Std. Deviation	DI	Alpha Coefficient (SEM)	Correlation of Reading & Math
ELL-ONLY	36.13	18.56	-36.9	.85 (7.19)	.69
SWD-ONLY	28.01	14.18	-76.6	.82 (6.02)	.70
ELLWD	21.38	10.87	-131.3	.75 (5.44)	.52
Non-ELL/non-SWD	49.46	20.48		.88 (7.09)	.78
Total	46.71	21.00			.75

The average DI over grades (Grade 3 and 8) and over content areas (reading and math) shows that ELLWD students performed 203.3% less than their non-ELL/non-SWD peers. This is a great sign for concern and requires urgent attention. Many factors could

contribute to such gaps--among them, classification, instruction, and assessment issues can be mentioned.

The data presented above show another interesting trend. In many of the content areas, the performance gap between ELL and the non-ELL/non-SWD is similar with the performance gap between SWD and non-ELL/non-SWD. This trend underscores the importance of language on the assessment of ELL students and shows how the complex linguistic structure of the assessment may create a substantial gap between ELLs and native speakers of English.

The passage of No Child Left Behind heightened accountability by mandating the inclusion of these students in high-stakes state and national assessments and by requiring states to provide equal education opportunity for these students. In the next section of this paper we present analyses based on the post-NCLB education system. We will examine any systematic changes across time that could possibly be explained by the implementation of NCLB.

Post-NCLB Data Analyses Results

As indicated earlier, an extension to our research question/hypothesis on the performance gap between ELLWD and non-ELL/non-SWD students was whether or not making schools more accountable for academic progress of ELLWD students helps reduce such performance gaps. To test this hypothesis, we obtained post-NCLB data from another state (Site 2) and performed analyses similar with those conducted on the pre-NCLB data that were reported in Tables 1 and 2. Before presenting the results of post-NCLB analyses, it must be noted that the results from pre-NCLB and post-NCLB may not be comparable due to many factors. The two assessments may not be the same in content, format, and purpose. For example, the pre-NCLB assessment systems were mainly based on the notion of normative interpretation while the post-NCLB assessments are based more on the concept of criterion-referenced interpretation. Therefore, the post-NCLB assessments are mostly aligned with the state content standards while the pre-NCLB assessments were not necessarily aligned with the state content standards. Furthermore, different test formats and the use of assessments with different psychometric characteristics are among the differences in the two assessment systems**.**

The post-NCLB data came from a state with a large number of ELL students. Item-level data were available for students in several grade levels. We chose Grades 5 and 8 to be more comparable with the pre-NCLB data (Grade 3 data were not available). Due to the confidentiality agreement with the state, we do not report any identifiable information and refer to this data site as Site 2. We report the same descriptive statistics (mean, standard deviation, DI, alpha, standard error of measurement (SEM) and correlation between reading and math) for this site as were reported for Site 1. It must also be indicated that the study did not find major systematic differences between the population characteristics of the pre- and post-NCLB sites.

Table 3 presents a summary of analyses for students in Grade 5 in reading and math content areas. As data in Table 3 shows, the results of analyses are very consistent with those reported in Table 1 for Site 1, pre-NCLB. The data suggest a substantial performance gap between ELLs, SWDs, and ELLWDs on one hand and the reference group (non-ELL/non-SWD) on the other hand. For example, the Disparity Index (DI) for ELL students in reading for Grade 5 students is -33.4, suggesting that ELL students performed 33.4% lower than non-ELL/non-SWD students. The DI for SWD is -35.7 and for ELLWD the DI is -65.9. Once again, consistent with what was reported in Tables 1 and 2 (analyses on data from Site 1), ELLWD students showed a much larger performance gap than either ELLs or SWDs. For math, the DI for ELL is -22.6, for SWD it is -29.7 and for ELLWD the DI is -47.5.

Table 3.
Descriptive statistics by subgroups of students for Grade 5 Students in Reading and Math (post-NCLB)

Reading					
ELL/ SWD Status	Mean	Std. Deviation	DI	Alpha Coefficient (SEM)	Math & Reading Correlation
ELL-ONLY	222.62	60.12	-33.4	0.85 (23.28)	0.65**
SPEC.ED-ONLY	218.88	71.14	-35.7	0.91 (21.34)	0.69**
ELL&SP.ED	179.02	54.92	-65.9	0.80 (24.56)	0.49**
ENG-ONLY	296.96	61.00		0.89 (20.23)	0.71**
Total	277.49	70.96		0.92 (20.07)	0.76**
Math					
ELL/ SWD Status	Mean	Std. Deviation	DI	Alpha Coefficient (SEM)	Math & Reading Correlation

ELL-ONLY	259.78	59.65	-22.6	0.88 (20.66)	0.65**
SPEC.ED-ONLY	245.56	72.39	-29.7	0.93 (19.15)	0.69**
ELL&SP.ED	216.02	58.12	-47.5	0.89 (19.28)	0.49**
ENG-ONLY	318.53	64.15		0.90 (20.29)	0.71**
Total	301.87	70.97		0.92 (20.08)	0.76**

Reading test has 42 items and math test has 46 items.
** Correlation is significant at the 0.01 level (2-tailed).

Table 4 shows a summary of analyses for Grade 8 students in reading and math. Data presented in Table 4 are also quite consistent with the data presented in Tables 1 through 3. There is a substantial performance gap between ELLs, SWDs, and ELLWDs as compared with the performance of non-ELL/non-SWDs. For example, the DI for ELL students in reading is -28.5, for SWDs, the DI is -36.7 and for ELLSWDs the DI is -60.0. Once again, while there is a huge performance gap between the three subgroups (ELLs, SWDs, and ELLWDs) and the reference group, the performance gap between ELLWDs and the reference group is substantially higher. In math, the DI for ELLs was -37.4, for SWDs it was -56.3 and for ELLWDs it was -89.1.

Table 4.
Descriptive statistics by subgroups of students for Grade 8 Students in Reading and Math (post-NCLB)

Reading					
ELL/ SWD Status	Mean	Std. Deviation	DI	Alpha Coefficient (SEM)	Reading & Math Correlation
ELL-ONLY	244.12	64.68	-28.5	0.88 (22.14)	0.65(**)
SWD-ONLY	229.45	68.79	-36.7	0.91 (20.64)	0.60(**)
ELLWD	196.01	57.81	-60.0	0.81 (25.20)	0.38(**)
Non-ELL/non-SWD	313.63	63.98		0.92 (18.10)	0.73(**)
Total	297.43	72.21		0.93(19.10)	0.77(**)
Math					
ELL/ SWD Status	Mean	Std. Deviation	N	Alpha Coefficient (SEM)	Reading & Math Correlation
ELL-ONLY	230.35	85.09	-37.4	0.87 (30.68)	0.65(**)
SWD-ONLY	202.58	79.34	-56.3	0.88 (27.48)	0.60(**)
ELLWD	167.42	61.99	-89.1	0.81 (27.02)	0.38(**)
Non-ELL/non-SWD	316.54	90.01		0.92 (25.46)	0.73(**)
Total	295.58	97.93		0.92(27.70)	0.77(**)

Reading test has 48 items and math test has 65 items.
** Correlation is significant at the 0.01 level (2-tailed).

In sum, the data in Tables 3 and 4 show that the post-NCLB analyses results are consistent with the pre-NCLB trends that were presented in Tables 1 and 2 based on data from Site 1. These data suggest that: (1) all three subgroups underperform the reference group, (2) ELLWDs demonstrate a substantially higher performance gap than either ELLs or SWDs, and (3) the performance gap between all three subgroups with the main group is lower in math than in reading suggesting that language factors play a major role in this performance gap.

The post-NCLB data show that ELLWD students with an average DI of -65.7 over the two content areas and two grades underperformed the reference group by 65.7%. A comparison between the overall percent of post-NCLB underperformance (65.7%) with the respective pre-NCLB percent (203.3%) may shed light on the impact of NCLB in reducing the performance gap between the subgroups and the reference group. However, such comparisons should be interpreted with extreme caution because the pre- and the post-NCLB assessments vary in many different aspects. They are based on different tests that are aligned with content standards presumably being taught in the classrooms. Additionally, they are based on different measurement concepts (norm-referenced versus criterion-referenced). The difference may also be due to more attention to student test performance during the post-NCLB era since such performance was a major requirement in the post-NCLB accountability system.

Issues concerning technical/psychometric characteristics of assessments used for ELLWD students

Linguistic and cultural factors have a profound impact on the assessment of ELLWDs since these students are often at the lower level of the English language proficiency spectrum (Abedi, 2007). Secondly, they are disproportionately classified as students with learning disabilities (Abedi, 2006a; Artiles et al., 2005). Due to the impact of linguistic and cultural factors and also because of their disabilities, assessments that are developed and field-tested for the mainstream student population may not provide valid outcomes for these students. Results of analyses of existing data from different locations nationwide show a substantial gap in reliability and validity between these and

other students particularly on test items with a high level of language demand. In this section of the paper, we present data on the reliability and validity of tests for ELLWD students compared to the data from other subgroups of students.

Reliability

To compare reliability of assessments for ELLWD students with the reference group, we used data from the two states referred to as Site 1 and Site 2. Site 1 provides data from the pre-NCLB assessments and Site 2 provides data for post-NCLB assessments. In both sites we had access to the item-level data, and therefore, we were able to compute internal consistency (alpha coefficient) as an estimate of the reliability of the tests.

Pre-NCLB

The results of analyses on existing state assessment data (Site 1) showed a large gap in the reliability coefficient between the three subgroups on one hand (ELLs, SWDs, and ELLWDs) and the reference group (non-ELL/non-SWD) on the other hand. Internal consistency (alpha) coefficients for pre-NCLB data are reported in Table 1 (reading and math for Grade 3, Site 1) and Table 2 (reading and math for Grade 8, Site 1). Averaging over Grades 3 and 8, the reliability coefficient for reading tests was .820 for ELL students, .795 for SWDs, and .715 for ELLWD students as compared with the reliability coefficient of .890 for the reference group. Thus, according to the data in Tables 1 and 2, the reliability coefficients for the subgroups are lower than the coefficient for the reference group. However, the reliability gap is much higher for ELLWD than any other subgroups.

The reliability coefficients for math (averaging over Grades 3 and 8) are generally higher than those for reading. However the gap between subgroups and the reference group is still high. The reliability coefficients for math are .845 for ELLs, .825 for SWDs, and .775 for ELLWDs as compared with a reliability coefficient of .895 for the reference group (non-ELL/non-SWD).

Post-NCLB

Reliability coefficients are also reported for the post-NCLB data from Site 2 in Tables 3 and 4. Averaging over Grade 5 (Table 3) and Grade 8 (Table 4), the average

reliability coefficient for reading was .865 for ELLs, .910 for SWDs, and .805 for
ELLWDs as compared with a reliability coefficient of .925 for the reference group. For
the math test, the average reliability coefficient was .875 for ELLs, .905 for SWDs, and
.850 for ELLWDs as compared with the coefficient of .910 for the reference group.

The post-NCLB reliability data that are consistent with the pre-NCLB results
show that: (1) the reliability coefficients for the subgroups (ELLs, SWDs, ELLWDs) are
mainly lower than the coefficient for the reference group (non-ELL/non-SWD); (2) the
average reliability coefficient for ELLWDs is generally lower than reliability coefficients
for the other two subgroups (ELLs and SWDs); and (3) the math reliability coefficients
are higher for all three subgroups, and (4) the post-NCLB reliability coefficients are
generally higher than the pre-NCLB reliabilities.

It is important to note at this point that reliability coefficients in this study were
estimated based on the concept of internal consistency. It is clear that in such
computations, the score distributions would have substantial impact on the size of alpha
coefficients. Data presented in Tables 1 through 4 clearly shows the trend of lower
standard deviation for the subgroups (ELL, SWD and ELLWD) as compared with the
reference group (non-ELL/non-SWD). To account for differences in the score
distributions across the subgroups and the reference group we computed and reported
standard error of measurement by different groups. The results indicate that while the
trend of discrepancies between the standard error of measurement between the subgroups
and the reference group remains (with the reference group having lower SEM in general)
but such discrepancies are substantially lower than those reported for the reliability
coefficients.

Validity

Literature has also provided evidence on the validity of standardized tests for
ELLs and students with disabilities. Researchers focusing on the validity of assessments
for ELL students are concerned about the confounding of the focal construct with the
construct that may be irrelevant to the assessment. For example, if a test measuring math
content knowledge (the focal construct) has a complex linguistic structure or it is
influenced by cultural biases, then the linguistic complexity and cultural factors which
are irrelevant to the focal construct may affect the validity of interpretation of the test

results (Abedi, 2006b). It such cases, the focal construct is said to be confounded with the sources irrelevant to the construct being measured. These sources are often referred to as construct-irrelevant variance (Messick, 1994).

For ELL students, assessments suffer from the impact of linguistic and cultural biases as sources of construct irrelevant variance. For students with disabilities, a variety of nuisance variables including linguistic complexity of assessments (for students with learning disabilities), test format, fatigue and frustration factors due to test length, and overcrowding of pages, influence their assessment outcomes. Assessments for ELLWDs suffer from the combination of both sources of construct irrelevant variance and nuisance variables.

Confounding the focal construct with the construct-irrelevant sources affects the structural relationship between test items within a test and across different tests measuring the same construct. To shed light on the construct validity issues for the assessments of subgroups (ELLs, SWDs, and ELLWDs) we carried out a study that compared the structural relationship underlying the performance of subgroups with the reference group (non-ELL/non-SWD) on different tests. A simple-structure confirmatory factor analysis model was used for this comparison. These analyses were conducted using the post-NCLB data as they are derived from more recent assessments for these students.

Additionally, the data from Site 2 provided an opportunity for examining the validity of assessments for ELL students. For this site, there was access to a longitudinal cohort for who statewide mathematics test scores were available for two consecutive years for each student. For example, for students in Grade 4, in addition to the state mathematics assessment scores for the current year, the same data were available from their previous year.

In a confirmatory factor analytic model, item-level scores from the state mathematic tests in two consecutive years were used. Rather than factoring individual test items, factor analyses were conducted on a group of items as a "testlet" often referred to as "parcels" of items. The main reason for conducting factor analyses on item parcels was to avoid issues on factoring phi and tetrachoric correlations (Bejar, 1980; Carroll, 1983; Cook, Dorans and Eignor, 1988).

Several item parcels were created for each of the test items. For Grade 4 students, the math test had 60 items and for the previous year their math test had 60 items as well. The 60 item Grade 4 math test was divided into 6 item parcels with 10 items in each parcel. Similarly, the Grade 3 math test was divided into 6 item parcels with 10 items in each parcel. A math latent variable was created from the 6 item parcels for the Grade 4 math test and another math latent variable was created from the 6 item parcels for Grade 3 math. Factor loadings of the parcels on the math latent variables as well as the correlation between the two latent variables (Grade 4 and Grade 3 math) were obtained. Analyses were performed separately for each of the subgroups (ELL, SWD, and ELLWD) as well as for the reference group (non-ELL/non-SWD).

A similar model was used for Grade 8 students. As indicated earlier, for these students, in addition to their current state test scores in math, their state test scores in math from the year before were also obtained. Table 5 summarizes the results of the confirmatory factor analyses for Grade 4 students. Correlations between the item parcels and latent variables (for both Grade 4 and Grade 3 math) are reported separately for the subgroups (ELL, SWD, ELLWD) and the non-ELL/non-SWD. Correlations between the two math latent variables (Grade 4 and Grade 3) are also reported.

Table 5. Comparing Factorial Validity Data Across the Subgroups and the Reference Group for Students in Grade 4 (Post NCLB)

Item Parcels/Subgroups	ELL	SWD	ELLWD	Non-ELL Non/SWD
	Factor Loadings			
Math (Grade 3)				
Parcel 1	.721	.765	.740	.790
Parcel 2	.724	.789	.633	.772
Parcel 3	.734	.781	.757	.791
Parcel 4	.765	.781	.749	.770
Parcel 5	.436	.551	.269	.644
Parcel 6	.707	.763	.626	.772
Math (Grade 4)				
Parcel 1	.806	.851	.795	.880
Parcel 2	.780	.819	.776	.876
Parcel 3	.828	.860	.745	.826
Parcel 4	.816	.855	.792	.810
Parcel 5	.830	.864	.788	.841
Parcel 6	.720	.751	.693	.880

Average Loadings	.739	.786	.697	.804
	Factor Correlation			
Grade 3 & 4	.780(**)	.781(**)	.734(**)	.782(**)

* Factor Loadings from the second component, the rest have only one component extracted.
** Correlation is significant at the 0.01 level (2-tailed).

Comparing factor loading and factor correlations across the three subgroups (ELL, SWD, ELLWD), and non-ELL/non-SWD reveals an interesting trend. The trend of factor loadings for the subgroups is generally lower than the trend for the reference group (non-ELL/non-SWD). These findings are consistent with the literature (see, for example, Abedi, Leon & Mirocha, 2003). The trend of factor loadings for the ELLWD students is even more interesting. Factor loading for the ELLWD students are substantially lower than the loadings for the reference group and even lower than those for either ELLs or SWDs. To discuss the trend of factor loadings across the subgroups and the reference group, due to differences in the size of factor loadings across the item parcels, reporting median factor loadings (rather than mean factor loadings) deemed more relevant. However, we preferred mean factor loading since it is more reflective of the factor loadings of individual parcels. But whether using mean or median, the trend of difference between the subgroups and the reference group remained the same. For example, both mean and median factor loadings show that all three subgroups (ELL, SWD and ELLWD) had lower factor loadings compared with the reference group. Furthermore, both mean and median factor loadings showed that ELL students with disabilities showed substantially lower factor loadings with the reference group as compared with the other two subgroups.

Based on the data in Table 5, the mean factor loadings (for current and previous math) is .739 for ELLs, .786 for SWDs, and .697 for ELLWD students as compared to the mean factor loadings of .804 for the reference group. These data suggest that the structural relationship between test items and the total test is lower for the three subgroups in general and for the ELLWD students in particular when compared with the reference group. The data also show that the correlation between the two math scores (Grade 3 and Grade 4) is lower for all three subgroups (.739 for ELL, .781 for SWD, and .734 for ELLWD) when they are compared with the reference group (.782). Once again,

the ELLWD group has the lowest factor correlation between the two math scores as compared with other groups.

The results of analyses for Grade 8 students are quite consistent with those for Grade 4 students. Recall that for Grade 8 students, state math test scores for the current year and the previous year were analyzed. Two math latent variables were created, one for Grade 8 (current) and one for Grade 7 (previous year) for the same students. Table 6 summarizes the results of confirmatory factor analyses for students in Grade 8. As the data in Table 6 show, factor loading (correlation between parcel scores and the latent variables) are generally lower for subgroups when compared with the reference group. The results also suggest that factor loadings for ELLWD students are substantially lower than those for either ELLs or SWDs. The mean factor loadings over parcels for Grades 7 and 8 is .640 for ELLs, .607 for SWDs, and .528 for ELLWD as compared with a mean factor loading of .780 for the reference group. Once again, the data suggest the structural relationship between test items and total test is lower for the three subgroups when compared with the reference group, and lowest for ELLWD students.

Table 6 also shows factor correlations between the two math tests (Grade 8 and Grade 7). These correlations are .640 for ELLs, .697 for SWDs, .528 for ELLWDs, as compared with .780 for the non-ELL/non disability group. It is important to note that for both Grade 4 and 8, the item parcels had substantial loadings on the first factor and only one factor with a given value greater than 1 was extracted for ELLs, SWDs, and the reference group, but for ELLWD, two factors emerged. These findings suggest that assessments for ELLWD students suffer more from construct irrelevant variance which leads to the multidimensionality issue. That is, in addition to the single dimension of math, which is the focal construct, other dimensions such as linguistic factors enter in the assessment system for these students.

Table 6. Comparing Factorial Validity Data Across the Subgroups and the Reference Groups for Students in Grade 8 (Post NCLB)

Item Parcels/Subgroups	ELL	SWD	ELLWD	Non-ELL Non/SWD
	Factor Loadings			
Math (Grade 7)				
Parcel 1	.616	.686	.501	.758

Parcel 2	.583	.657	.394	.771
Parcel 3	.704	.751	.671	.802
Parcel 4	.730	.770	.650	.772
Parcel 5	.655	.724	.604	.816
Parcel 6	.430	.582	.433	.754
Math (Grade 8)				
Parcel 1	.646	.695	.312	.814
Parcel 2	.691	.748	.668	.780
Parcel 3	.650	.692	.640	.751
Parcel 4	.646	.655	.515	.805
Parcel 5	.671	.704	.497	.788
Parcel 6	.663	.701	.457	.744
Mean Loadings	**.640**	**.697**	**.528**	**.780**
	Factor Correlation			
Grade 7 & 8	.845	.845	.708	.845

Parcels/Subgroups = 8 items each, except for Parcel 6 which contains 5 items;

However, as indicated in the reliability results section earlier, the lower factor loadings and lower correlations between assessments for subgroups and the reference group may be mainly due to differences in the distribution of scores across the subgroups. As elaborated earlier, the distributions of scores for subgroups (ELL, SWD and ELLWD) show less variability than those for the reference group.

The results of analyses examining psychometric characteristics of the assessments (current and previous year math) clearly show the serious impact of nuisance variables on the reliability and validity of assessments for ELL, SWD and ELLWD students. These findings suggest that academic achievement instruments that are developed for the mainstream students may not be quite relevant for special needs student populations, particularly for ELLWD students who suffer from a combination of challenges. To provide fair assessment and accountability for all, these nuisance variables impacting the reliability and validity of assessments for ELL, SWD, and ELLWD students must be carefully identified and their effects on the assessment for these students must be controlled.

Accommodations

To support ELLWD students in overcoming challenges they face in their academic careers, accommodations that help them with their disabilities and linguistic

needs must be provided. Unfortunately, many of the accommodations used for SWD and ELL students may not be effective in making assessments more accessible for these students. The strategy of test accommodations for ELL students originates in testing students with disabilities. Because of this, many forms of accommodations used for ELL students were created for SWDs and may not be appropriate for ELL students. For example, in analyses of 73 accommodations used for ELL students across the nation (Rivera, 2003), only 11 or 15% of them were deemed relevant for ELL students. In making this determination, accommodations that directly address language needs of ELL students, in addition to their disabilities, were identified as relevant accommodations for ELLWD students.

As states have aimed toward greater inclusion of these students in their content-based assessments, multiple testing accommodations have been implemented. As indicated above, these accommodations are provided to help ELLWD overcome challenges due to both language acquisition and their disabilities. Unfortunately, these accommodations have not been sufficiently evaluated (Rivera, Collum, Shafer, & Sia, 2004). Thus, research examining the effectiveness and validity of accommodations used for ELLWD students is scarce. Minnema, Thurlow, Anderson, and Stone (2005) reviewed literature regarding English language learners with disabilities and large-scale assessments. They found only 10 articles related to all three of these criteria (ELLs, disabilities, and assessments) from the mid-1960s to 2004 (see also, Thurlow & Liu, 2001).

Minnema et al.'s (2005) review of the literature did not find any research studies that specifically addressed the effectiveness of accommodations used in the assessment of English language learners with disabilities due to the challenge in finding appropriate accommodations for ELLWD students. State officials who were interviewed by Anderson, Minnema, Thurlow, and Hall-Lande (2005) reported having separate accommodation policies for ELLs and students with disabilities. In these cases, a collaborative effort is needed to ensure that students are receiving accommodations that address both their linguistic and disability needs.

For a better understanding of accommodations one must view them within the entire academic system. That is, major factors that influence selection of

accommodations and affect the outcome of accommodated assessments must be carefully examined. Validity of accommodations for ELLWD students is greatly influenced by factors affecting identification, classification and assessment of these students.

<center>Discussion</center>

Recent trends in federal legislation and national and state assessments call for a more inclusive policy for all students, particularly those with limited proficiency in English and those with disabilities. However, the more inclusive policies may not necessarily lead to more academic progress and to fairer assessment for subgroups of students who are at risk of education failure, unless more attention is focused toward the instruction and assessments of these students. Studies on the classification, assessment and accommodation for ELL with disabilities (ELLWD) students do not provide encouraging results. Inconsistencies in classification practices for both English language learners (ELLs) and students with disabilities (SWD) have a combined effect on the instruction and assessment of ELLWD students as these students suffer from dual challenges in their academic careers. They must deal with their limited English proficiency, learning a new language, and coping with their disabilities, all of which create obstacles in their academic progress.

In order to systematically address issues in the education of ELLWD students, all factors affecting their academic performance must be discussed in a comprehensive manner. Improper classification of ELLWD students may render assessment results unfair, invalid, and ineffective, which may lead to inappropriate and inadequate instruction for these students. Instructional and assessment materials for these students should be free of any systematic biases such as linguistic and cultural biases. Unnecessary linguistic complexity of assessment has a profound impact on this group of students since they are English learners and a majority of them are classified as students with learning disabilities. Literature clearly suggests that language factors affect performance of both groups (Abedi, 2006a, 2006b).

To help ELLWD students overcome problems due to their limited English proficiency and disabilities, different forms of accommodations have been proposed. However, due to an extremely complex situation in the assessment of ELLWD students, these accommodations are often ineffective and may even provide invalid assessment

outcomes. There are many factors affecting the selection of accommodations and the outcome of accommodated assessments. Among these factors issues concerning effectiveness, relevance, differential impact, and validity of accommodated assessments should be thoroughly examined. Inappropriate placement of students in the ELLWD category can seriously affect decisions concerning the selection of accommodations for these students. Accommodations which are not effective in making assessments more accessible for these students may have negative impact on their assessment outcomes. Accommodations that alter the construct being assessed may also impact the validity of assessments for these students. Furthermore, accommodations must be consistent with students' background variables and needs.

Discussion of the issues concerning assessment, classification and accommodation for ELL and ELLWD students helps us understand the complexity inherent in accommodation theory and practice for ELLWD students. Since these students are at a disadvantage from two distinct aspects (language and disability, they should be accommodated according to both their language needs and their disabilities. Therefore, in many cases a combination of accommodations should be provided.

For the language aspect, their language background variables must be reviewed and accommodations that are relevant to their language needs must be provided to them. For ELL students, many different forms of accommodations are used (Rivera, 2003) -- only a few of which could be appropriate for ELL students. Research-based recommendations can be provided for some accommodations that can help these students with their language needs. For example, depending on their level of English language proficiency, accommodations such as linguistically modified tests (Abedi, Hofstetter & Lord, 2004), customized English and bilingual dictionaries (Abedi, Courtney & Leon, 2003; Sireci, Li, & Scarpati, 2003), or native language or bilingual test booklets (Abedi, Lord, Hofstetter & Baker, 2000; Sireci et al., 2003) and computer-based testing may be helpful.

To assist them with their disabilities, recommendations should be made based on several factors. First, the accommodations must be consistent with their IEP. Second, there must be some research evidence on the validity and effectiveness of accommodations to justify their use. Many different accommodations for these students

have been used (Thurlow et al., 2000; Rivera, et. al. 2004). Unfortunately, however, there is not sufficient research evidence to support the effectiveness or validity of many of these accommodations (Thurlow et al. 2000).

Decisions for selecting accommodations for ELLWD and the outcome of accommodated assessments for these students can be influenced by issues concerning classification, instruction and assessment. If students are not correctly identified as ELLWD, then accommodations provided for these students may not work no matter how valid and effective they are as accommodations. This is particularly important in the case of ELLWD since misclassification can have grave consequences on their academic career. The selected accommodations can then impact both classification and assessment outcomes. Accommodations that are used in instructional settings can be more effective and more useful. All these factors can then determine the validity of assessment outcomes. A comprehensive view of any of these three components requires knowledge of the other two components.

Below are a few recommendations based on the literature cited above:

1. Make sure that students are properly classified as ELL, SWD, or ELLWD. Issues in the classification system would have a great influence on other aspects of education for these students.

2. Make assessments as accessible as they can be made to control for sources of biases (cultural and linguistic biases) and other nuisance variables. Assessments that are constructed and field tested for the mainstream students may not be accessible to the three subgroups of students discussed in this paper.

3. Use accommodations that help overcome some of the problems and challenges that students are faced with. For example, for ELLWD students, select accommodations that directly address either their linguistic needs or their disabilities.

4. Select accommodations that do not alter the construct being measured. If validity of the accommodated assessments is not well established then the outcome of such assessments may not be aggregated with the outcome from the mainstream students

5. Select accommodations that are practical and easy to implement.

References

Abedi, J. (2004). The No Child Left Behind Act and English language learners: Assessment and accountability issues. *Educational Researcher, 33*(1), 4-14.

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 Downing, S. M. and Haladyna, T. M. (Eds.). *Handbook of Test Development* (pp. 377-398). Mahwah, NJ: Lawrence Erlbaum Associates.

Abedi, J. (2007). *English Language Learners with Disabilities*. In Cahlan-Laitusis, C. & Cook, L. (Eds.). *Accommodating student with disabilities on state assessments: What works?* (pp. 23-35) Arlington, VA: Council for Exceptional Children.

Abedi, J. (2008). Classification system for English language learners: Issues and recommendations. *Educational Measurement: Issues and Practice*, *27* (3), 17-22.

Abedi, J., Courtney, M., & Leon, S. (2003). *Research-supported accommodation for English language learners in NAEP* (CSE Tech. Rep. No. 586). Los Angeles: University of California, National Center for Research on Evaluation, Standards, and Student Testing.

Abedi, J. & Herman, J. (in press). Assessing English language learners' opportunity to learn mathematics: Issues and limitations. *Teachers College Record*.

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., Kao, J.C., Leon, S., Sullivan, L., Herman, J., Pope, R., & Mastergeorge, A. (2008). *Exploring factors that affect the accessibility of reading comprehension assessments for students with disabilities: A study of segmented text.* Minneapolis, MN: University of Minnesota, Partnership for Accessible Reading Assessment.

Abedi, J., Leon, S., & Mirocha, J. (2003). *Impact of student language background on content-based performance: Analyses of extant data* (CSE Tech. Rep. No. 603). Los Angeles, CA: University of California, National Center for Research on Evaluation, Standards, and Student Testing.

Abedi, J., Lord, C., & Plummer, J. (1997). *Language background as a variable in NAEP mathematics performance* (CSE Tech. Rep. No. 429). Los Angeles, CA: 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.

Aldus, D., & Thurlow, M. (2005). *Beyond subgroup reporting: English language learners with disabilities in 2002-2003 online state assessment reports* (ELLs with Disabilities Report 10). Minneapolis, MN: University of Minnesota, National Center on Educational Outcomes. Retrieved August 28, 2008 from the World Wide Web: http://education.umn.edu/NCEO/OnlinePubs/ELLsDisReport10.html.

Anderson, M., Minnema, J., Thurlow, M., & Hall-Lande, J. (2005). *Confronting the unique challenges of including English language learners with disabilities in statewide assessments* (ELLs with Disabilities Report 9). Minneapolis, MN: University of Minnesota, National Center on Educational Outcomes.

Artiles, A. J., Rueda, R., Salazar, J., & Higareda, I. (2005). Within-group diversity in minority disproportionate representation: English Language Learners in urban school districts. *Exceptional Children*, *71*, 283-300.

Bejar, I. (1980). A Procedure for investigating the unidimensionality of achievement tests based on item parameter estimates. *Journal of Educational Measurement*, *7*, 283-296.

Carroll, J. B. (1983). The difficulty of a test and its factor composition revisited. In H. Wainer & S. Messick (Eds.). *Principals of modern psychological measurement: A festschrift for Frederic M. Lord* (pp. 257-282). Hillsdale, NJ: Lawrence Erlbaum Associates.

Cook, L. L., Dorans, N. J., & Eignor, D. R. (1988). An assessment of the dimensionality of three SAT – Verbal test editions. *Journal of Educational Statistics, 13,* 19-43.

Le Floch, K.C., Martinez, F.C., O'Day, J., Stecher, B., Taylor, J., Cook, A. (2007). State and Local Implementation of the "No Child Left Behind Act." Volume III--Accountability under "NCLB" Interim Report. US Department of Education.

Linn, R. L. & Gronlund, N. E. (1995). *Measurement and assessment in teaching* (7th ed.) Englewood Cliffs, NJ: Prentice-Hall.

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.

Messick, S. (1994). The interplay of evidence and consequences in the validation of performance assessments, *Educational Researcher, 23*(2): 13-23.

Minnema, J., Thurlow, M., Anderson, M., & Stone, K. (2005). *English language learners with disabilities and large-scale assessments: What the literature can tell us* (ELLs with Disabilities Report 6). Minneapolis, MN: University of Minnesota, National Center on Educational Outcomes.

No Child Left Behind Act of 2001, Pub. L. No. 107-110, 115 Stat. 1425 (2002).

Ortiz, A. A. (2002). Prevention of school failure and early intervention for English language learners. In Artiles A. J. & Ortiz, A. A. (Ed.). *English language learners with special education needs* (pp.31-63). Washington, DC: Center for Applied Linguistics.

Perie, M., Grigg, W. & Donahue, P. (2005). *The Nation's Report Card*: *Reading 2005* (NCES 2006-451). U.S. Department of Education, National Center for Eduction Statistics. Washington, DC: U.S. Government Printing Office.

Reschly, D.J. (1996). Identification and assessment of students with disabilities. *The Future of Children, 6*(1), 40-53.

Rivera. C. (2003, June) *State assessment policies for English language learners.* Presented at the 2003 Large-Scale Assessment Conference. **San Antonio, Texas.**

Rivera, C., Collum, E., Shafer, L., & Sia, J. K. (2004). *Analysis of state assessment policies regarding the accommodation of English language learners, SY 2000–2001.* Arlington, VA: George Washington University, Center for Equity and Excellence in Education.

Rueda, A., Artiles, A. J., Salazar, J., & Higareda, I. (2002). An analysis of special education as a response to the diminished academic achievement of Chicano/Latino students: An update. In R.R. Valenica (Ed.). *Chicano school*

failure and success: Past, present, and future (2nd ed., pp. 310-332). London: Routledge/Palmer.

Sireci, S. G., Li, S., & Scarpati, S. (2003). *The effects of test accommodation on test performance: A review of the literature* (Center for Educational Assessment Research Report No. 485). Amherst, MA: School of Education, University of Massachusetts, Amherst.

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.

Thurlow, M., & Liu, K. (2001). *State and district assessments as an avenue to equity and excellence for English language learners with disabilities* (LEP Projects Report 2). Minneapolis, MN: University of Minnesota, National Center on Educational Outcomes.

Thurlow, M. L., McGrew, S., Tindal, G., Thompson, S. J., Ysseldyke, J. E., & Elliott, J. L. (2000). *Assessment accommodations research: Considerations for design and analysis* (Technical Report 26). Minneapolis, MN: University of Minnesota, National Center on Educational Outcomes.

U.S. Department of Education, Office of Special Education and Rehabilitative Services, Office of the Assistant Secretary (2004). *National symposium on learning disabilities in English language learners, October 14-15, 2003: Symposium summary*. Washington, D.C.

Ysseldyke, J, & Bielinski, J. (2002). Effect of different methods of reporting and reclassification on trends in test scores for students with disabilities. *Exceptional Children, 68*(2), 189-200.

Zehler, A. M., Fleischman, H. L., Hopstock, P. J., Pendzick, M. L., & Stephenson, T. G. (2003). *Descriptive study of services to LEP students and LEP students with disabilities* (Special Topic Report No. 4: Findings on Special Education LEP Students). Arlington, VA: Development Associates, Inc.