

World-Class Instructional Design and Assessment



**The Bridge Study between Tests of English Language Proficiency
and *ACCESS for ELLs*®**

Part I: Background and Overview

Margo Gottlieb, Ph.D.
WIDA Lead Developer

Dorry Kenyon, Ph.D.
Center for Applied Linguistics

Technical Report #2

FINAL

August 31, 2006

Table of Contents for Part I: Background and Overview

I. Introduction to the Bridge Study	1
A. Purposes of the Bridge Study.....	1
B. Descriptions of the English Language Proficiency Tests Used in the Bridge Study	2
C. Major Differences between the Old Generation of English Language Proficiency Tests and ACCESS for ELLs	3
D. Uses of Results from the Bridge Study.....	4
E. Caveats on Interpreting the Data in the Bridge Study	4
II. General Procedures.....	5
III. Analyses.....	6
A. General Issues	6
B. Comparing the Four Tests with ACCESS.....	6
IPT	6
LAS.....	9
LPTS	10
MAC II	11
IV. Results	12
A. Participants.....	12
B. Analytical Procedures	12
C. How to Read the Results (Sample).....	13
V. General Discussion.....	24
A. Analytical Caveats.....	24
B. Future Research	24

I. INTRODUCTION TO THE BRIDGE STUDY

As part of the development of a comprehensive assessment system for English language learners in kindergarten through Grade 12, states in the World-Class Instructional Design and Assessment (WIDA) Consortium have developed **Assessing Comprehension and Communication in English State to State for English Language Learners (ACCESS for ELLs)**, a new English language proficiency test, as their annual, large-scale measure. By February 2004, the Consortium's English language proficiency standards with their large-scale and classroom frameworks had been crafted and introduced to member states. The large-scale assessment framework of the English language proficiency standards, aligned with state academic content standards, served as the basis for designing the specifications and anchoring ACCESS for ELLs.

Prior to the inauguration of ACCESS for ELLs in spring 2005, states within the Consortium were using a variety of marketed English language proficiency tests for initial identification, placement, and annual assessment of their English language learners, most notably the Idea Proficiency Test (IPT), the Language Assessment Scales (LAS), the Language Proficiency Test Series (LPTS), and the Revised Maculaitis II (MAC II). These measures existed before passage of the No Child Left Behind Act in 2001, and consequently many of their features were not in concert with the requirements set forth in the law.

This report summarizes the Consortium's efforts to establish estimates of comparability between the four English language proficiency measures and ACCESS for ELLs. The relative equivalence between the tests was determined through a bridge study conducted in spring 2005 in two member states, Illinois and Rhode Island. School districts in Illinois, the state with the largest ELL population among WIDA Consortium states, had discretion in selecting three of the four named measures: IPT, LAS, and LPTS. School districts in Rhode Island administered the MAC II. As all the member states used one or more of these English language proficiency tests, we felt that the results from the study would be generalizable across the Consortium.

Part I of this report presents background to the bridge study and an overview of procedures. It also contains information on how to read the results of the study. Part II contains the results of the bridge study and is divided into four subsections, one for each of the four English language proficiency measures in the study:

Part II A: IPT Results

Part II B: LAS Results

Part II C: LPTS Results

Part II D: MAC II Results

A. Purposes of the Bridge Study

The overall purpose of the complete bridge study is to determine the strength of the relationship between the results from each of the four English language proficiency tests being used in Consortium states and those from ACCESS for ELLs. In other words, the intent is to establish links between older measures of English language proficiency and the Consortium's new test when a representative sample of English language learners and proficient English students are double tested within a specified time frame. The estimates of comparability ascertained through the bridge study are helpful for various stakeholders in 2005–06, the year when states are transitioning from their old to the new generation of English language proficiency testing. Table 1 outlines the various levels of implementation of language proficiency assessment, with a corresponding purpose for conducting the study.

Table 1
The Purposes for the Bridge Study by Level of Implementation

Level of Implementation	Purposes for the Bridge Study
State	<ul style="list-style-type: none"> • Meet compliance with Title III requirements under No Child Left Behind • Provide continuity of data flow for cohorts of English language learners identified in 2002–03, the baseline year • Determine Annual Measurable Achievement Objectives (AMAOs) for the established cohorts in the transitional year
School District	<ul style="list-style-type: none"> • Explain the differential performance of English language learners' English language proficiency in the transitional year
School	<ul style="list-style-type: none"> • Enable continuity of support services for English language learners

B. Descriptions of the English Language Proficiency Tests Used in the Bridge Study

There is substantial disparity between tests of English language proficiency in terms of their theoretical orientation, design, materials, and reporting of results. The four tests are briefly described here.

Idea Proficiency Test (IPT): The Early Literacy Test, first published in 2001, joined the IPT compendium of oral language measures introduced in the late 1970s and literacy measures introduced in the late 1980s. The tests have undergone some revision and renorming throughout their history. The tests are components of a management program, and the results are referenced to accompanying instructional materials. Unlike other language proficiency tests, they include a subsection on reading for life skills. Listening/speaking is directly assessed with answers coded as correct or incorrect; writing is another direct subsection that utilizes a four-point rubric. Results are reported as three designations (six levels of oral language proficiency) and three levels of literacy.

Language Assessment Scales (LAS): This test was developed in response to a Supreme Court decision of 1974, *Lau v. Nichols*, and was initially published in 1976 as an oral measure. The K–12 oral language forms have undergone minor modifications through the years. The most current literacy sections date back to 1988, the oral language sections were last revised in 1990, and pre-LAS was introduced in 1998. Speaking and writing are direct subsections that utilize rubrics for teacher scoring, while listening and reading are selected response. Tapes are available for administration of listening and speaking; procedures to establish interrater reliability of the direct subsections are listed in the administration manual. The results yield five levels of oral language proficiency and three levels of literacy.

Language Proficiency Test Series (LPTS): Launched in 1999, this test is the newest of the marketed English language proficiency measures. Its conceptual base and design were derived from the Illinois Measure of Annual Growth in English (IMAGE), which at that time served as the state’s English language proficiency test. Unlike the other language proficiency measures, LPTS draws from themes with sets of interrelated items, and it is vertically scaled. Oral language and writing rubrics accompany the direct subsections of listening/speaking and writing. There are no supplementary materials outside of the paper-and-pencil test. The results are reported as two levels of oral language proficiency and four levels of literacy.

Revised Maculaitis II (MAC II): In its second edition, this test has incorporated the Degrees of Reading Power as a subtest starting in Grade 4. It has the most grade level cluster forms (five) of the four English language proficiency tests that measure oral language and literacy. Listening is assessed and reported independently from speaking. Five-point rubrics are used to interpret oral language and writing samples. Results are converted into five levels of oral language competency and five levels of literacy.

C. Major Differences between the Old Generation of English Language Proficiency Tests and ACCESS for ELLs

No Child Left Behind marks a major shift in the use of English language proficiency tests for accountability purposes. The Consortium constructed ACCESS for ELLs in compliance with the mandates of the act. In addition, ACCESS for ELLs reflects the growing body of theory and research on academic language and exemplifies current instructional practice in the field, centering on the integration of language and content. Table 2 compares the features of the generation of English language proficiency tests developed prior to 2001 and those of ACCESS for ELLs.

Table 2

Comparison of Features of the Old Generation of English Language Proficiency Tests and ACCESS for ELLs

Old Generation of English Language Proficiency Tests	ACCESS for ELLs
<ul style="list-style-type: none"> • Not standards-based 	<ul style="list-style-type: none"> • Anchored in WIDA’s English language proficiency standards
<ul style="list-style-type: none"> • Nonsecure 	<ul style="list-style-type: none"> • Secure
<ul style="list-style-type: none"> • Emphasize social language 	<ul style="list-style-type: none"> • Emphasizes academic language
<ul style="list-style-type: none"> • Generally integrated oral language domains 	<ul style="list-style-type: none"> • Independent language domains (i.e., listening and speaking)
<ul style="list-style-type: none"> • Different tests for each grade level cluster (no comparability) 	<ul style="list-style-type: none"> • Vertically scaled across tiers and grade level clusters
<ul style="list-style-type: none"> • Low-stakes 	<ul style="list-style-type: none"> • High-stakes
<ul style="list-style-type: none"> • Not compliant with the requirements of No Child Left Behind 	<ul style="list-style-type: none"> • Compliant with the requirements of No Child Left Behind

D. Uses of Results from the Bridge Study

The results from the bridge study have applicability at every level of implementation. States may use the data to define the criteria to identify English language learners and set revised AMAOs for established cohorts. School districts will be better able to inform stakeholders of the relationship between their previously used English language proficiency test and ACCESS for ELLs. Schools will have additional information on the range of performance of their English language learners to use to align curriculum, instruction, and classroom assessment. Table 3 presents these varied uses of the bridge study data.

Table 3

Uses of the Results from the Bridge Study by Level of Implementation

Level of Implementation	Uses of the Results
State	<ul style="list-style-type: none">• Examine data in the aggregate to formulate policy for school districts in regard to identification, placement, and reclassification decisions for English language learners• Assist in determining initial tier placement for ACCESS for ELLs
School District	<ul style="list-style-type: none">• Implement state policy for identification, placement, and reclassification of English language learners• Use multiple data sources for decision-making in the transitional year• Offer justification for reexamining support services for English language learners based on the results
School	<ul style="list-style-type: none">• Communicate to teachers working with English language learners• Stimulate examination of curriculum, instruction, and classroom assessment

E. Caveats on Interpreting the Data in the Bridge Study

The results from the bridge study must be tempered with much caution. The goal of the study is to predict scores on ACCESS for ELLs based on scores from other English language proficiency tests; however, the tests under examination do not have a matched set of characteristics. The differences among the basic structures of the tests alone are reason to interpret the results cautiously. Differences among all five tests include the following:

- Degree of alignment with English language proficiency and academic content standards
- Number and types of items in each subsection or language domain
- Depth of knowledge of the items
- Inclusion of the language of math, science, and social studies
- Ceiling levels of the measures
- Rubrics used for interpreting speaking and writing
- Treatment of grade level clusters in multiple forms
- Types of scoring and reporting (metrics used)

For these reasons, and those listed in Table 2, high levels of correspondence between the measures in the bridge study cannot be established. The recalibration of previously used English language proficiency tests with ACCESS for ELLs must be viewed as an *estimate* of comparability at best. Results from the empirical analyses on the equivalences for each test form are presented in the four subsections (Parts II A-D) of this report.

II. GENERAL PROCEDURES

During the first operational administration of the ACCESS for ELLs in spring 2005 in Alabama, Maine, and Vermont, students in selected districts in Illinois and Rhode Island were also administered the test. Within three months of taking the ACCESS for ELLs between April 15 and June 15, these students were also administered one of four other tests used by their district: the IPT, LAS, LPTS, or MAC II.

The ACCESS for ELLs was administered and scored under operational conditions by MetriTech, the vendor that operationalizes the test for the WIDA Consortium. MetriTech also collected scores on the other tests from the districts participating in the bridge study. The data were matched at MetriTech; MetriTech confirmed that performances on ACCESS for ELLs matched the scores data for the appropriate students. Among all participating students, about 3% of scores could not be matched. The data were then sent to the Center for Applied Linguistics (CAL) for analyses.

While the ACCESS for ELLs data came from the operational administration of the test, results from the other tests were reported by the schools. This meant that several kinds of errors were detected in the data. The following is an illustrative but not complete list of the types of data errors that CAL needed to correct before proceeding with the analyses:

- Test form information was missing.
- Test form information was inaccurate or incomplete.
- Range of test results did not match range for reported test form.
- Test results across districts for same reported test form were not comparable (some of the results could be made comparable, but many could not).

CAL made every effort to contact the school districts to clean the data from the four tests. While ACCESS for ELLs data were clean and complete, the scores for students for whom all individual responses were missing for any domain (listening, speaking, reading, or writing) were treated as missing. In a very few cases the responses on a teacher-administered section (the kindergarten test in all domains or the speaking test in all grade level clusters) indicated that the test was incorrectly administered (e.g., stopped before a ceiling was reached). In such cases, the student's score for that domain was also treated as missing.

III. ANALYSES

A. General Issues

As mentioned earlier, there are many fundamental differences between the older English language tests and ACCESS for ELLs. For the purpose of the bridge study analyses, two major differences became important.

First, three of the four tests (IPT, LAS, and LPTS) do not provide a separate score for listening and speaking. For purposes of the bridge study, listening and speaking scores on ACCESS for ELLs were each separately related to these tests' single oral language (or listening/speaking) score.

Second, the way most of the tests (with the exception of the LPTS) divided the grade level clusters differed from that of the ACCESS for ELLs. Because the ACCESS for ELLs scores are vertically scaled, results on the ACCESS for ELLs could cross grade level cluster boundaries. This was not true of the English language tests (again, with the exception of the LPTS). In addition, apart from the LPTS, scores on adjacent-level test forms could not be compared. The score scale for each form was unique, although, according to the technical manuals, scores on parallel forms (e.g., versions A and B) were comparable. Thus, it was decided that each unique English language test would form the basis for score comparison.

To compare proficiency levels was a little harder. Some tests have different cut scores on the same form by grade levels. Most notably, the LPTS has different cut scores for its proficiency levels for practically each grade. In addition, since each grade level cluster on the ACCESS for ELLs has a separate set of cut scores, care must be taken to interpret the proficiency levels across the two forms in light of the grade of the students being compared.

The following tables show the number of unique comparisons made between performances on the ACCESS for ELLs and on the other tests, by test and ACCESS for ELLs domain (listening, speaking, reading, and writing). In the case of the MAC II, ACCESS for ELLs composite scores were compared with MAC II total scores. At the top of each chart are the grade levels. A space is left between the five different ACCESS for ELLs test forms (K, 1–2, 3–5, 6–8, and 9–12). Note, however, that although the K form differs from the 1–2 form in its mode of administration, many of the items are the same and there is only one set of cut scores for the K–2 grade level cluster. The rows in the table show the domains and test forms. The first domain is Listening. The "ACCESS" row fills in all blanks by grade level cluster. Below that is the listening score from the English language test (often an "oral language" score). Following this are the unique forms the English language test uses to test the domain across the grade level clusters. Each number represents a unique comparison presented in the results section of this report, although not necessarily in the order of this table.

The following is a discussion highlighting each set of comparisons in turn.

B. Comparing the Four Tests with ACCESS

IPT

The IPT provides a combined listening and speaking score (oral language). While an oral pre-IPT test exists, it is for 3- to 5-year-olds, and no students in the final cleaned database were identified as having taken it. Forms 1 and 2 span the K to 12 range. In this study, enough students were administered the older Form 1 version (C and D) and the newer Form 1 version (E

and F) to merit separate analyses. According to the publishers, performances on C and D could not be compared to performances on E and F.

Spanning from K to Grade 6, the Form 1 Oral IPT crosses four sets of test forms and three grade level clusters on the ACCESS for ELLs. Form 2 Oral IPT crosses two grade level clusters.

On the IPT, for reading and writing, a low-level form (EL, or Early Literacy) exists for kindergarteners and first-graders. This form and the other elementary forms divide across grade level cluster boundaries on the ACCESS for ELLs. The highest level reading and writing form (Form 3) spans Grades 7 to 12 and also crosses two grade level cluster boundaries on the ACCESS for ELLs. Because of these boundary crossings, when comparing outcomes in terms of proficiency levels, it is imperative to interpret the results with caution, paying strict attention to the grade level of the students who are being compared.

Table 4

Comparing the IPT with the ACCESS for ELLs

1. ACCESS/IPT Overlap		K	1	2	3	4	5	6	7	8	9	10	11	12
LISTENING														
	ACCESS	■	■	■	■	■	■	■	■	■	■	■	■	■
Oral Language														
1	IPT Oral 1C, 1D	■	■	■	■	■	■	■	■					
2	IPT Oral 1E, 1F	■	■	■	■	■	■	■	■					
3	IPT Oral 2C, 2D									■	■	■	■	■
SPEAKING														
	ACCESS	■	■	■	■	■	■	■	■	■	■	■	■	■
Oral Language														
4	IPT Oral 1C, 1D	■	■	■	■	■	■	■	■					
5	IPT Oral 1E, 1F	■	■	■	■	■	■	■	■					
6	IPT Oral 2C, 2D									■	■	■	■	■
READING														
	ACCESS	■	■	■	■	■	■	■	■	■	■	■	■	■
Reading														
7	IPT EL Reading	■	■	■	■	■	■	■	■					
8	IPT R 1A, 1B			■	■	■	■	■	■					
9	IPT R 2A, 2B					■	■	■	■					
10	IPT R 3A, 3B									■	■	■	■	■
WRITING														
	ACCESS	■	■	■	■	■	■	■	■	■	■	■	■	■
Writing														
11	IPT EL Writing	■	■	■	■	■	■	■	■					
12	IPT W 1A, 1B			■	■	■	■	■	■					
13	IPT W 2A, 2B					■	■	■	■					
14	IPT W 3A, 3B									■	■	■	■	■

LAS

The clusters on the LAS were very similar to those on the IPT. One difference, however, was in the pre-LAS. All kindergarten and first-grade students in the final clean dataset were administered the pre-LAS for preliteracy skills. All kindergarteners also took the pre-LAS for oral language skills, while some first-graders took the pre-LAS and others were administered Form 1 of the oral test. It should also be noted that the pre-LAS preliteracy test for kindergarteners and first-graders provided only a combined score for preliteracy skills, not a separate score for reading and writing.

Table 5
Comparing the LAS with the ACCESS for ELLs

ACCESS/LAS Overlap		K	1	2	3	4	5	6	7	8	9	10	11	12
LISTENING														
	ACCESS	■	■	■	■	■	■	■	■	■	■	■	■	■
Oral Language														
1	Pre-LAS Oral	■	■	■										
2	LAS Oral 1C, 1D		■	■	■	■	■	■	■					
3	LAS Oral 2C, 2D								■	■	■	■	■	■
SPEAKING														
	ACCESS	■	■	■	■	■	■	■	■	■	■	■	■	■
Oral Language														
4	preLAS Oral	■	■	■										
5	LAS Oral 1C, 1D		■	■	■	■	■	■	■					
6	LAS Oral 2C, 2D								■	■	■	■	■	■
READING														
	ACCESS	■	■	■	■	■	■	■	■	■	■	■	■	■
Preliteracy														
7	Pre-LAS preliteracy	■	■	■										
Reading														
8	LAS R/W 1A, 1B			■	■	■	■							
9	LAS R/W 2A, 2B					■	■	■	■					
10	LAS R/W 3A, 3B								■	■	■	■	■	■
WRITING														
	ACCESS	■	■	■	■	■	■	■	■	■	■	■	■	■
Preliteracy														
11	Pre-LAS preliteracy	■	■	■										
Writing														
12	LAS R/W 1A, 1B			■	■	■	■							
13	LAS R/W 2A, 2B					■	■	■	■					
14	LAS R/W 3A, 3B								■	■	■	■	■	■

LPTS

The grade level clusters of the LPTS parallel those of the ACCESS for ELLs. There are two main differences to note between the two tests. First, the LPTS does not give separate scores for listening and speaking as does the ACCESS for ELLs. Second, proficiency level cut scores on the LPTS, with the exception of Grades 11 and 12, are at the grade level; that is, there are separate cut scores to determine proficiency classifications for each grade. On the ACCESS for ELLs, the proficiency level cut scores are based on the grade level cluster.

Table 6
Comparing the LPTS with the ACCESS for ELLs

ACCESS/LPTS Overlap		K	1	2	3	4	5	6	7	8	9	10	11	12
LISTENING														
	ACCESS	■	■	■	■	■	■	■	■	■	■	■	■	■
Listening/Speaking														
1	LPTS 1, 5	■	■	■										
2	LPTS 2, 6				■	■	■							
3	LPTS 3, 7							■	■	■				
4	LPTS 4, 8										■	■	■	■
SPEAKING														
	ACCESS	■	■	■	■	■	■	■	■	■	■	■	■	■
Listening/Speaking														
5	LPTS 1, 5	■	■	■										
6	LPTS 2, 6				■	■	■							
7	LPTS 3, 7							■	■	■				
8	LPTS 4, 8										■	■	■	■
READING														
	ACCESS	■	■	■	■	■	■	■	■	■	■	■	■	■
Reading														
9	LPTS 1, 5	■	■	■										
10	LPTS 2, 6				■	■	■							
11	LPTS 3, 7							■	■	■				
12	LPTS 4, 8										■	■	■	■
WRITING														
	ACCESS	■	■	■	■	■	■	■	■	■	■	■	■	■
Writing														
13	LPTS 1, 5	■	■	■										
14	LPTS 2, 6				■	■	■							
15	LPTS 3, 7							■	■	■				
16	LPTS 4, 8										■	■	■	■

MAC II

The MAC II divides the K–12 range into five clusters, although the top two clusters (Grades 6–8 and 9–12) align completely with the ACCESS for ELLs. MAC II provides separate scores for listening and speaking. In addition, the centralized dataset sent to the Consortium from Rhode Island contained total scores on the MAC II, which were compared to the ACCESS for ELLs composite scores.

While two forms of the MAC II (A and B) exist at each level, all students in this study were administered form B. Since A and B may be compared only via standard scores, standard scores on the MAC II were used in the study.

Table 7

Comparing the MAC II with the ACCESS for ELLs

ACCESS/MACII Overlap		K	1	2	3	4	5	6	7	8	9	10	11	12
LISTENING														
	ACCESS	■	■	■	■	■	■	■	■	■	■	■	■	■
Listening														
1	A1, B1	■	■											
2	A2, B2			■	■									
3	A3, B3					■	■							
4	A4, B4							■	■	■				
5	A5, B5										■	■	■	■
SPEAKING														
	ACCESS	■	■	■	■	■	■	■	■	■	■	■	■	■
Speaking														
6	A1, B1	■	■											
7	A2, B2			■	■									
8	A3, B3					■	■							
9	A4, B4							■	■	■				
10	A5, B5										■	■	■	■
READING														
	ACCESS	■	■	■	■	■	■	■	■	■	■	■	■	■
Reading														
11	A1, B1	■	■											
12	A2, B2			■	■									
13	A3, B3					■	■							
14	A4, B4							■	■	■				
15	A5, B5										■	■	■	■
WRITING														
	ACCESS	■	■	■	■	■	■	■	■	■	■	■	■	■
Writing														
16	A1, B1	■	■											
17	A2, B2			■	■									
18	A3, B3					■	■							
19	A4, B4							■	■	■				
20	A5, B5										■	■	■	■
COMPOSITE														
	ACCESS	■	■	■	■	■	■	■	■	■	■	■	■	■
Total														

(7) Statistics from the regression analysis to predict ACCESS for ELLs scale scores from scores on the English language test

In addition, tables were prepared in Excel to show the relationship between scores on the English language test (with their corresponding proficiency levels) and predicted scores on the ACCESS for ELLs (with their corresponding proficiency levels). The tables were prepared showing every possible raw or scale score (as appropriate) on the English language test. The technical manuals for each test were consulted to determine the conversion into proficiency levels. For every score on an English language test, there is at least one proficiency level. Some tests, notably the LPTS, have different cut scores for the same test form, depending on the grade level of the student. In such cases, multiple columns were made for each appropriate grade level. Likewise, when the English language test spanned multiple grade level clusters on the ACCESS for ELLs, multiple proficiency levels were also possible and included in the table.

Following the SPSS analysis, the regression coefficients from the SPSS output were used to build an equation that was applied to the Excel tables to produce a predicted ACCESS for ELLs scale score for every possible score on the English language test. Once these were entered into the table, a look-up table was used to find the corresponding ACCESS for ELLs proficiency level score (e.g., 4.8) that corresponded to the predicted ACCESS for ELLs scale score. As mentioned above, some tables needed multiple ACCESS for ELLs proficiency level score columns.

Once the table was complete in Excel, it was cut and pasted into the appropriate place in the output document.

An initial draft of the format in which bridge study results would be reported was circulated to several members of the WIDA Steering Committee prior to their meeting at the end of October 2005. Following comments from one member, a revised version was presented to the Steering Committee on October 25, 2005. The contents of the output were approved with minor revisions, which have been incorporated into the final format. The first draft of this report was made available to WIDA Consortium states at the end of November, 2005.

C. How to Read the Results (Sample)

The following output from the comparison between performances on the MAC II and ACCESS for ELLs is annotated to help the reader completely understand the bridge study output.

Annotations appear in a dotted text box in *cursive script*.

graders. Most of the students were administered Tier C of the ACCESS.

Results for MAC II B2 Read with ACCESS Read

Results for MAC II B2 Read with ACCESS Read
1. Participants (All)

grade * tier Crosstabulation

Count

		tier			Total
		A	B	C	
grade	2	20	3	52	75
	3	3	8	52	63
Total		23	11	104	138

The title tells us these results are reading scores on the MAC II ACCESS reading scores. The scores are listed before each section.

Results for MAC II B2 Read with ACCESS Read

2. Descriptives (Listwise Deletion)

Descriptive Statistics

	N	Range	Minimum	Maximum	Mean	Std. Deviation
ACCESS Read Scale Score	136	203	166	369	307.50	29.169
MAC II Read Standard Score	136	154.00	154.00	308.00	230.1324	30.98191
Valid N (listwise)	136					

This table under section 2, Descriptives, shows us in the first column the two scores used in the comparison: the ACCESS Reading Scale Score and the MAC II Reading Standard Score. The number of students (N) is the number of those who had both scores and were used in the analysis. In this case, it was 136 of the 138 students who were administered the test. The next column (range) shows the number of points between the minimum observed score in the Minimum column and the maximum observed score in the Maximum column. The next column shows the average (mean) score, and the final column shows the standard deviation (Std. Deviation). Locating the mean scores in the final table can give an indication of the average ability level of the students in this comparison.

Results for MAC II B2 Read with ACCESS Read
 3. Crosstabulation of Results (Numbers)

MAC II Read Prof Level * ACCESS Read Prof Level Crosstabulation

Count

				ACCESS Read Prof Level						
				Entering	Beginning	Developing	Expanding	Bridging	Attained	Total
Cluster	Gr 1-2	MAC II Read Prof Level	Beginner	4	0	0	0	0	0	4
			Low Intermediate	3	7	3	2	0	2	17
			High Intermediate	0	4	5	6	6	7	28
			Advanced	0	1	5	3	5	10	24
		Total	7	12	13	11	11	19	73	
	Gr 3-5	MAC II Read Prof Level	Beginner	1	0	0	0	0		1
			Low Intermediate	3	2	0	0	0		5
			High Intermediate	4	8	2	0	0		14
			Advanced	5	26	8	2	2		43
		Total	13	36	10	2	2		63	

This table under section 3, Crosstabulation of Results (Numbers), shows how students classified into a proficiency level based on their performance on the MAC II B2 reading test were classified into an ACCESS proficiency level based on their performance on the ACCESS. Since on the ACCESS there are separate cut scores for each grade level cluster, and MAC II B2, administered to grades 2 and 3, spans two ACCESS grade level clusters, there are two parts to this table, as indicated in the leftmost columns. The top part presents results for the grade level cluster 1-2 (i.e., for the second-graders in this study). The bottom part presents results for the grade level cluster 3-5 (i.e., for the third-graders in this study). In each part, the rows show the proficiency level outcome on the MAC II; the columns show the proficiency level outcomes on the ACCESS. In terms of the raw numbers of students, we see that of the four students in second grade who were classified as "Beginner" based on their performance on MAC II were classified as "Entering" based on their performance on ACCESS. Checking the total column on the far right, we see that these were the only students in this group. In the next row, of the 17 students classified as "Low Intermediate" by their MAC II performance, 3 were classified as "Entering" based on ACCESS, 7 as "Beginning," 3 as "Developing," 2 as "Expanding," none as "Bridging," and 2 as "Attained."

Results for MAC II B2 Read with ACCESS Read
 4. Crosstabulation of Results (Percentages)

MAC II Read Prof Level * ACCESS Read Prof Level Crosstabulation

% within MAC II Read Prof Level

				ACCESS Read Prof Level						
				Entering	Beginning	Developing	Expanding	Bridging	Attained	Total
Cluster	Gr 1-2	MAC II Read Prof Level	Beginner	100.0%						100.0%
			Low Intermediate	17.6%	41.2%	17.6%	11.8%		11.8%	100.0%
			High Intermediate		14.3%	17.9%	21.4%	21.4%	25.0%	100.0%
			Advanced		4.2%	20.8%	12.5%	20.8%	41.7%	100.0%
		Total	9.6%	16.4%	17.8%	15.1%	15.1%	26.0%	100.0%	
	Gr 3-5	MAC II Read Prof Level	Beginner	100.0%						100.0%
			Low Intermediate	60.0%	40.0%					100.0%
			High Intermediate	28.6%	57.1%	14.3%				100.0%
			Advanced	11.6%	60.5%	18.6%	4.7%	4.7%		100.0%
		Total	20.6%	57.1%	15.9%	3.2%	3.2%		100.0%	

This table under section 4, Crosstabulation of Results (Percentages), shows the exact same information as the table under section 3, except that each cell now shows the percentage of students classified in a certain level by the MAC II B2 (that is, the row) who are classified in a certain level by the ACCESS (that is, in the column). In other words, since all four students in second grade who were classified as "Beginner" based on their performance on the MAC II were also classified as "Entering" based on their performance on the ACCESS, the result in that cell of the table is 100%. In the next row, of all the students classified as "Low Intermediate" by their MAC II performance, 17.6% were classified as "Entering" based on the ACCESS, 41.2% as "Beginning," 17.6% as "Developing," 11.8% as "Expanding," none as "Bridging," and 11.8% as "Attained." This table should be understood only in light of the table under section 3; the fewer students in a row, the less weight should be placed on the results. Nevertheless, the table can be useful for quickly seeing what percentage of students was at or above a certain level or what percentage of students was below a certain level. For example, we can quickly see that of the second-graders classified as "Advanced" by their performance on the MAC II, 62.5% (= 20.8% + 41.7%) were classified in the top two ACCESS proficiency levels.

Results for MAC II B2 Read with ACCESS Read

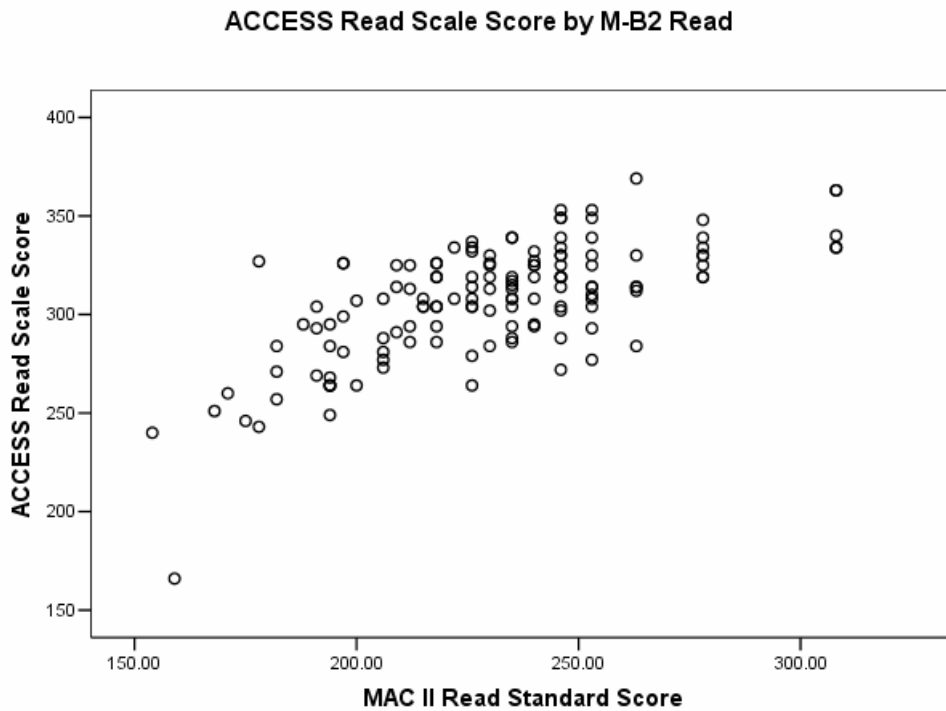
5. Correlation

Correlation

		MAC II Read Standard Score
ACCESS Read Scale Score	Pearson Correlation	.675
	Sig. (2-tailed)	.000
	N	136

This table under section 5, Correlation, shows the correlation between students' scores on the two tests. The higher the correlation, the stronger the relationship between performances on the two tests. The second row (Sig.) shows the probability that there is no correlation between the two tests. The result of .000 shows that there is very little chance that there is no correlation between the tests. The final row again shows the number of students involved in determining the correlation.

Results for MAC II B2 Read with ACCESS Read
6. Scatterplot



The scatterplot under section 6, Scatterplot, shows every observed data point in the comparison. Scores on the MAC II are on the horizontal axis, while scores on the ACCESS are on the vertical axis. Each circle represents the intersection of at least one student's results on the two tests. The scatterplot can indicate the range of scores that were observed in the study. Results will be stronger if there is a wide distribution of scores on both tests; if scores in one area of the test range are missing, results might be skewed. This plot shows that there was a fairly good range of scores for this comparison, though there were few at the very lowest and few at the highest end of the range.

Results for MAC II B2 Read with ACCESS Read
7. Regression Coefficients

Model Summary

	R	R Square	Adjusted R Square	Std. Error of the Estimate
Model 1	.675	.456	.452	21.589

ANOVA

		Sum of Squares	df	Mean Square	F	Sig.
Model 1	Regression	52404.378	1	52404.378	112.431	.000
	Residual	62457.622	134	466.102		
	Total	114862.000	135			

Coefficients

		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
Model 1	(Constant)	161.152	13.926		11.572	.000
	MAC II Read Standard Score	.636	.060	.675	10.603	.000

The three tables under section 7, Regression Coefficients, show basic results from the analysis to predict ACCESS scores from MAC II scores. The "R" in the Model Summary table is always equal to the correlation. The results of the ANOVA show the variance explained by the regression (Regression) and the variance not explained by the regression (Residual). The comparative size of these amounts can give a feel for how much the variance of the scores can be accounted for by the regression. The significance test (Sig.) shows the probability that the variation explained by the regression model is due to chance. The result of .000 shows that this is not likely. The results presented in the Coefficients table are the most important

While the preceding tables are based on output from SPSS, the final table (on the following pages) was developed from Excel. It shows predicted ACCESS scores based on scores on the older test as per the SPSS regression results. At the very top of the table is the regression equation. (Note that depending on the Excel table column width, some of the coefficients may be shown as rounded. The exact coefficients from the SPSS outputs were used in all calculations. Predicted ACCESS scores are rounded to the nearest whole number.) The title of the table indicates the scores involved. To save space, the results are presented in two parallel columns that for many tests cross several pages. The first column shows the score on the older test. The second column shows the proficiency level associated with that score by grade as per published materials or information direct from the test publisher. (Note that in some cases there will be multiple columns here when there are several different cut scores by grade level for the same test form. For the MAC II there is only one.) The next column shows the predicted ACCESS score given the MAC II score. The final column(s) shows the proficiency level associated with the ACCESS score by grade(s) involved in the comparison. In the table below, there are two columns. The first column is for the second-graders and is based on the cut scores for ACCESS grade level cluster K-2. The second column is for the third-graders and is based on the cut scores for ACCESS grade level cluster 3-5.

For example, our best prediction of what a student receiving a MAC II Reading Standard Score of 226 would get on the ACCESS is 305. This is regardless of whether the child is in second or third grade. According to the MAC II, the score of 226 is at the "HI" or "High Intermediate" proficiency level. On the ACCESS, if that child is in second grade, the score of 305 on reading would place the child at 4.8, "Expanding," but fairly close to level 5, "Bridging." However, a third-grade child with a score of 305 would be placed 2 (=2.0), which is the ACCESS level "Beginning." This example is one of the most extreme in all the bridge study results. Yet it shows the importance of interpreting proficiency level designations in light of appropriate grade levels.

This table can also shed light on the raw data and some of the results from SPSS. For example, under number 2, Descriptives, we saw that the lowest observed standard score on MAC II reading was 154 and the highest 308. From this table, which shows all possible standard scores on the MAC II according to test publisher documentation, we see that a score of 154 is proficiency level B ("Beginner"), above the lowest MAC II proficiency level of BB ("Basic Beginner"). There were no very low scores in this group on the MAC II. The highest observed MAC II score of 308 was quite high relative to the maximum score of 325 and well above the cut score for the highest MAC II proficiency level of A ("Advanced") which is 231. The mean MAC II score of 230 occurs right at the cut score between HI and A, indicating that this group of test takers was, in terms of the MAC II, on the relatively more able side. Because regression procedures are generally most accurate near the means of the score distribution, this information is useful in interpreting results from the bridge study.

$$\text{Predicted ACCESS} = 161.152 + 0.636 * \text{MACII}$$

MAC II B2 Reading Standard Score to WIDA ACCESS Reading Scale Score									
MAC II B2 Standard Score	MAC II B2 Proficiency Level (by grade)	Predicted ACCESS Score	ACCESS Proficiency Level (by grade)		MAC II B2 Standard Score	MAC II B2 Proficiency Level (by grade)	Predicted ACCESS Score	ACCESS Proficiency Level (by grade)	
			Reading	2				3	Reading
105	BB	228	1.8	1.6	218	HI	300	4.3	1.9
106	BB	229	1.8	1.6	219	HI	300	4.3	1.9
107	BB	229	1.8	1.6	220	HI	301	4.4	1.9
108	BB	230	1.8	1.6	221	HI	302	4.5	1.9
109	BB	230	1.8	1.6	222	HI	302	4.5	1.9
110	BB	231	1.8	1.6	223	HI	303	4.6	1.9
111	BB	232	1.8	1.6	224	HI	304	4.7	1.9
112	BB	232	1.8	1.6	225	HI	304	4.7	1.9
113	BB	233	1.8	1.7	226	HI	305	4.8	2
114	BB	234	1.8	1.7	227	HI	306	4.9	2
115	BB	234	1.8	1.7	228	HI	306	4.9	2
116	BB	235	1.8	1.7	229	HI	307	4.9	2.1
117	BB	236	1.8	1.7	230	HI	307	4.9	2.1
118	BB	236	1.8	1.7	231	A	308	5	2.1
119	BB	237	1.8	1.7	232	A	309	5.1	2.2
120	BB	237	1.8	1.7	233	A	309	5.1	2.2
121	BB	238	1.9	1.7	234	A	310	5.1	2.2
122	BB	239	1.9	1.7	235	A	311	5.2	2.2
123	BB	239	1.9	1.7	236	A	311	5.2	2.2
124	BB	240	1.9	1.7	237	A	312	5.3	2.3
125	BB	241	1.9	1.7	238	A	313	5.4	2.3
126	BB	241	1.9	1.7	239	A	313	5.4	2.3
127	BB	242	1.9	1.7	240	A	314	5.4	2.4
128	BB	243	1.9	1.7	241	A	314	5.4	2.4
129	BB	243	1.9	1.7	242	A	315	5.5	2.4
130	BB	244	1.9	1.7	243	A	316	5.6	2.4
131	BB	244	1.9	1.7	244	A	316	5.6	2.4
132	BB	245	1.9	1.7	245	A	317	5.6	2.5
133	BB	246	1.9	1.7	246	A	318	5.7	2.5
134	BB	246	1.9	1.7	247	A	318	5.7	2.5
135	BB	247	1.9	1.7	248	A	319	5.8	2.6
136	BB	248	1.9	1.7	249	A	320	5.9	2.6
137	BB	248	1.9	1.7	250	A	320	5.9	2.6
138	BB	249	1.9	1.7	251	A	321	5.9	2.6
139	BB	250	1.9	1.7	252	A	321	5.9	2.6
140	BB	250	1.9	1.7	253	A	322	5.9	2.7
141	BB	251	1.9	1.7	254	A	323	6	2.7
142	BB	251	1.9	1.7	255	A	323	6	2.7
143	BB	252	1.9	1.7	256	A	324	6	2.8
144	BB	253	1.9	1.8	257	A	325	6	2.8
145	BB	253	1.9	1.8	258	A	325	6	2.8

$$\text{Predicted ACCESS} = 161.152 + 0.636 * \text{MACII}$$

MAC II B2 Reading Standard Score to WIDA ACCESS Reading Scale Score										
MAC II B2 Standard Score	MAC II B2 Proficiency Level (by grade)	Predicted ACCESS Score	ACCESS Proficiency Level (by grade)		MAC II B2 Standard Score	MAC II B2 Proficiency Level (by grade)	Predicted ACCESS Score	ACCESS Proficiency Level (by grade)		
			Reading	2				3	Reading	2
146	BB	254	1.9	1.8	259	A	326	6	2.8	
147	BB	255	1.9	1.8	260	A	327	6	2.9	
148	BB	255	1.9	1.8	261	A	327	6	2.9	
149	BB	256	1.9	1.8	262	A	328	6	2.9	
150	B	257	1.9	1.8	263	A	328	6	2.9	
151	B	257	1.9	1.8	264	A	329	6	2.9	
152	B	258	1.9	1.8	265	A	330	6	2.9	
153	B	258	1.9	1.8	266	A	330	6	2.9	
154	B	259	1.9	1.8	267	A	331	6	3	
155	B	260	1.9	1.8	268	A	332	6	3.1	
156	B	260	1.9	1.8	269	A	332	6	3.1	
157	B	261	1.9	1.8	270	A	333	6	3.1	
158	B	262	1.9	1.8	271	A	334	6	3.2	
159	B	262	1.9	1.8	272	A	334	6	3.2	
160	B	263	2	1.8	273	A	335	6	3.2	
161	B	264	2.1	1.8	274	A	335	6	3.2	
162	B	264	2.1	1.8	275	A	336	6	3.3	
163	B	265	2.1	1.8	276	A	337	6	3.3	
164	B	265	2.1	1.8	277	A	337	6	3.3	
165	B	266	2.2	1.8	278	A	338	6	3.4	
166	B	267	2.2	1.8	279	A	339	6	3.4	
167	B	267	2.2	1.8	280	A	339	6	3.4	
168	B	268	2.3	1.8	281	A	340	6	3.5	
169	B	269	2.3	1.8	282	A	341	6	3.6	
170	B	269	2.3	1.8	283	A	341	6	3.6	
171	B	270	2.4	1.8	284	A	342	6	3.6	
172	B	271	2.4	1.8	285	A	342	6	3.6	
173	B	271	2.4	1.8	286	A	343	6	3.7	
174	B	272	2.5	1.8	287	A	344	6	3.7	
175	B	272	2.5	1.8	288	A	344	6	3.7	
176	LI	273	2.5	1.8	289	A	345	6	3.8	
177	LI	274	2.6	1.9	290	A	346	6	3.8	
178	LI	274	2.6	1.9	291	A	346	6	3.8	
179	LI	275	2.6	1.9	292	A	347	6	3.9	
180	LI	276	2.7	1.9	293	A	348	6	3.9	
181	LI	276	2.7	1.9	294	A	348	6	3.9	
182	LI	277	2.7	1.9	295	A	349	6	3.9	
183	LI	278	2.8	1.9	296	A	349	6	3.9	
184	LI	278	2.8	1.9	297	A	350	6	4	
185	LI	279	2.8	1.9	298	A	351	6	4.1	
186	LI	279	2.8	1.9	299	A	351	6	4.1	
187	LI	280	2.9	1.9	300	A	352	6	4.2	

$$\text{Predicted ACCESS} = 161.152 + 0.636 * \text{MACII}$$

MAC II B2 Reading Standard Score to WIDA ACCESS Reading Scale Score										
MAC II B2 Standard Score	<i>MAC II B2 Proficiency Level (by grade)</i>	Predicted ACCESS Score	ACCESS Proficiency Level (by grade)			MAC II B2 Standard Score	<i>MAC II B2 Proficiency Level (by grade)</i>	Predicted ACCESS Score	ACCESS Proficiency Level (by grade)	
			Reading	2					3	Reading
188	LI	281	2.9	1.9		301	A	353	6	4.3
189	LI	281	2.9	1.9		302	A	353	6	4.3
190	LI	282	2.9	1.9		303	A	354	6	4.4
191	LI	283	2.9	1.9		304	A	354	6	4.4
192	LI	283	2.9	1.9		305	A	355	6	4.5
193	LI	284	3	1.9		306	A	356	6	4.5
194	LI	285	3.1	1.9		307	A	356	6	4.5
195	LI	285	3.1	1.9		308	A	357	6	4.6
196	LI	286	3.2	1.9		309	A	358	6	4.7
197	LI	286	3.2	1.9		310	A	358	6	4.7
198	LI	287	3.3	1.9		311	A	359	6	4.8
199	LI	288	3.3	1.9		312	A	360	6	4.9
200	LI	288	3.3	1.9		313	A	360	6	4.9
201	LI	289	3.4	1.9		314	A	361	6	4.9
202	LI	290	3.5	1.9		315	A	361	6	4.9
203	LI	290	3.5	1.9		316	A	362	6	5
204	HI	291	3.6	1.9		317	A	363	6	5
205	HI	292	3.7	1.9		318	A	363	6	5
206	HI	292	3.7	1.9		319	A	364	6	5.1
207	HI	293	3.8	1.9		320	A	365	6	5.1
208	HI	293	3.8	1.9		321	A	365	6	5.1
209	HI	294	3.8	1.9		322	A	366	6	5.2
210	HI	295	3.9	1.9		323	A	367	6	5.2
215	HI	298	4.1	1.9		324	A	367	6	5.2
216	HI	299	4.2	1.9		325	A	368	6	5.3
217	HI	299	4.2	1.9						

V. GENERAL DISCUSSION

A. Analytical Caveats

In addition to the caveats mentioned in Section I, users of these results need to remember that the tests in these studies are not equated. The results only suggest how performances on the two tests may compare. These comparisons are made in two categories: scores and proficiency level designations. Proficiency level designations were also made using different approaches for the different tests, using different techniques and standards.

Users of these results should keep in mind the data that went in to determining them. It is important to consider the number of students in each separate study, the distribution of the scores on each of the two tests, and the extent of the correlation between performances on the two tests. The larger the number of students, the more scores were distributed throughout the range of the older generation proficiency test (i.e., low, mid, and high scorers); the stronger the correlation between the two tests, the more appropriate may be the comparisons between scores.

B. Future Research

The bridge study presents a wealth of important data on English language testing. Each of the comparisons, combined with a content review of the test forms, could shed light on various aspects of testing the developing English proficiency of English language learners. That research, however, is beyond the scope of the current study.