August
2022

ISASP
IOWA STATEWIDE ASSESSMENT of STUDENT PROGRESS

Validity of English Language Proficiency and Statewide Standards-Based Assessments: Convergent Evidence or Divergent Purposes?

Prepared By

Stephen Dunbar, Catherine Welch, Yen Vo, Annette Vernon, Jun Hee Park, Jacob Herring and Heather Rickels

Validity of English Language Proficiency and Statewide Standards-Based Assessments: Convergent Evidence or Divergent Purposes?

For the past 15 years there has been a significant change in assessing the English learner (EL). Standards for large-scale English Language Proficiency (ELP) assessments have been created to address federal mandates. These mandates, including those under both the No Child Left Behind Act (NCLB) and Every Student Succeeds Act (ESSA), suggest that ELP assessments provide a critical path for identifying EL students in need of instructional supports and for monitoring their progress in ELP programs.

Since the accountability reform efforts of NCLB (2001), an emphasis has been placed on measuring the kind of language proficiency students need to be successful in an academic setting. NCLB established the initial connection between state academic content standards and each state's ELP assessment and associated standards (Wolf, Farnsworth & Herman, 2008; Bailey & Wolf, 2020). The 2015 authorization of ESSA further stipulated that "each state plan shall demonstrate that the state has adopted English language proficiency standards that are (iii) aligned with the challenging state academic standards" (Every Student Succeeds Act, 2015). In response to this mandate two multistate consortia, World-Class Instructional Design and Assessment (WIDA) and English Language Proficiency Assessment for the 21st Century (ELPA21), modified existing ELP assessments aligned to the new ELP standards, and most states have been using one or the other of these new ELP assessments in federally mandated accountability programs.

A challenge facing the newly developed ELP assessments has been the articulation of evidence to support their use for the multiple purposes they attempt to serve. As described above and by Huang & Flores (2018), these assessments are intended to serve as accountability

measures, provide classification evidence for placement in English language instruction, estimate standards-based achievement, and monitor progress toward exiting EL programs. One goal of this report is to establish a context for evaluating the validity evidence that is available for the English Language Proficiency Assessment for the 21st Century (ELPA21) using its relationship with a state's ESSA accountability measure, the Iowa Statewide Assessment of Student Progress (ISASP), as one source. A second goal is to evaluate the practical and policy issues associated with using the state's accountability measure in lieu of a separate ELP assessment, particularly for students who have received EL instruction for multiple years.

Validity Perspectives

To ensure that ELP and state accountability assessments support their intended purposes and inferences, systematic and critical reviews should be conducted, and empirical validity evidence should be provided. The 2014 *Standards for Educational and Psychological Testing* (hereafter *Standards*) refer to validity as "the degree to which evidence and theory support the interpretations of test scores for proposed uses of tests" (American Educational Research Association, American Psychological Association, & National Council on Measurement in Education, 2014, p. 11). Validity requires first an understanding of what the intended purposes and uses of a test are, followed by the gathering of appropriate evidence based on the *Standards*. Each score and intended interpretation must be validated to justify decisions based on these scores (AERA, APA & NCME, 2014).

Interpretation of test scores requires an understanding of what constructs are being measured by the instrument. Any influences outside the intended scope result in what the *Standards* call construct-irrelevant variance; likewise, failing to measure all aspects of the

intended construct results in construct underrepresentation according to the *Standards*. The

collection of appropriate evidence allows users of tests to judge the quality of the instrument

with respect to validity. This evidence also provides a basis for informing users that the test can

confidently be used for the intended purpose(s).

There are several sources of validity evidence outlined in the *Standards*: evidence based on

test content, on response processes, on internal structure, on relationships to other variables, and

on the consequences of testing. These sources often are prioritized to varying degrees depending

on the purpose of the test. The next section of this report is organized around these sources of

validity evidence as they pertain to English learners (ELs) who have taken both a state

accountability measure (ISASP English Language Arts) and a test specifically designed for

English Learners (ELPA21). As presented in Table 1 and consistent with the *Standards,* we

organize the presentation and analysis of validity evidence standard by standard before

considering how the aggregated evidence contributes to a unified validity argument (Kane,

2006). We also explore the potential consequences of using the ELPA21 for accountability for

ELs while at the same time requiring ELs take the statewide accountability assessment

developed to measure the performance of all students. Of particular interest is evidence that

suggests that these two assessments provide complementary information, raising questions about

duplication in testing, a potential unintended consequence of multiple testing requirements in

federal and state accountability programs.

Table 1

Sources of Validity Evidence for Review

Proposition	Standards of Interest	Evidence Needed
Tests are measuring the same or similar constructs	*Content (Standard 1.11)* ELPA21 and ISASP measure similar constructs	Review of standards and alignment-related information
	Cognitive Processes (Standard 1.12) ELPA21 and ISASP measure similar levels of cognitive processing	Review of test specifications to determine the processes to be assessed by the two assessments
Decision Concerning Proposition is Made		
Tests are predictive of a given criterion	*Relationships with related constructs and criteria (Standards 1.16 and 1.17)* ELPA21 and ISASP results classify test taker performance in a similar way.	Patterns of association between the two assessments will be examined. Decision consistency, correlations and proficiency indicators will be used.
	Evidence regarding relationship with criteria (Standards 1.18 and 1.20) ELPA21 and ISASP both accurately predict readiness to exit ELPs	Correlations between test scores used to draw inferences about student readiness
Test results are not affected by construct-irrelevance	*Internal Structure (Standard 1.13)* ELPA21 and ISASP are not unduly influenced by construct-irrelevant variance	Review of administrative conditions for both assessments; testing time, universal features and accommodated features compared

Participants

For this report, performance of Iowa students in grades 6 through 8 who took both the ELP assessment (ELPA21) and the state accountability measure (ISASP) during spring 2021 was considered to assemble validity evidence. This sample included 4,922 students and is described in Table 2. We focused on students in the middle grades in this report because of the correspondence between the ELP standards of ELPA21 and the ELA standards of ISASP as well as because by the time EL students reach middle school, they are more likely to have received EL instruction for one or more years. The sample is predominantly Hispanic/Latino (66 percent)

in ethnicity, White (58 percent) in terms of race, and eligible for free- or reduced-price lunch (81

percent). About one-fourth of the sample received special education services via an

Individualized Education Program (IEP) or 504 Plan. Appendix A includes results for students in

ELPA21 grade bands 4-5 and 9-11 parallel to those presented for students in grades 6 through 8.

Table 2

Demographic Characteristics of Study Sample (N = 4922)

Characteristic	Percent of Sample
Female	43.6
Male	56.4
Hispanic/Latino	65.5
Not Hispanic/Latino	34.5
American Indian/Alaskan Native	12.7
Hawaiian/Other Pacific Islander	4.9
Asian	8.8
Black/African American	18.0
White	58.3
Students eligible for free- or reduced-lunch	80.8
Students with an IEP or 504 Plan	24.5

Content-Oriented Evidence (Standard 1.11)

Critical validity information can be obtained from comparing the content of a given test

with the construct it is intended to measure. The test specifications carefully describe the content

of a test in detail, including the domains to be measured, the proportion of each domain and the

types of items. Review of information in Table 3 suggests a strong similarity between the two

assessments with respect to structure and item types. The primary purpose, to determine

proficiency in English language, is consistent across the two assessments, although the intended

audience of test information is obviously different.

ELPA21 measures four ELP domains (Reading, Writing, Listening and Speaking) using

three different item formats (selected response, constructed-response and technology-enhanced).

The ELPA21 Reading domain aligns with the ELP Standards developed by the Council of Chief

State School Officers. ISASP ELA domains align with the Common Core State Standards

(CCSS, National Governors Association Center for Best Practices & Council of Chief State

School Officers, 2010) in English Language Arts and uses similar item formats, with the addition

of an extended constructed-response item in writing.

Table 3

Comparison of ELPA21 and ISASP ELA Structure and Features

	ELPA21	ISASP ELA
Primary Purposes	Measure ELP Standards Fulfill accountability requirements for ESSA Monitor proficiency for ELP	Measure Iowa Core Standards Fulfill accountability requirements for ESSA Monitor proficiency in ELA Measure student growth

Content Domains	Reading
	Writing
	Speaking
	Listening
		English Language Arts
		  Reading
		  Language and Writing

Grades (Bands)	Grade bands for
	  Kindergarten to Grade 1
	  Grades 2 to 3
	  Grades 4 to 5
	  Grades 6 to 8
	  Grades 9 to 11
		Grade by grade assessments for grades 3 to 11

Item Formats	Multiple Choice Technology Enhanced Constructed-Response Extended Constructed-Response	Multiple Choice Technology Enhanced Constructed-Response Extended Constructed-Response
Scores Provided	Scale scores for Reading, Writing, Speaking and Listening	Scale scores for Reading, Language/Writing and ELA composite
Performance Levels	Five	Three

Performance Descriptors	Based on multi-state standard setting process
	  Beginning
	  Early Intermediate
	  Intermediate
	  Early Advanced
	  Advanced
		Based on State of Iowa standard setting process
		  Not Yet Proficient
		  Proficient
		  Advanced

The ELPA21 test specifications are described in Office of Teaching, Learning and Assessment (2021). The Listening test in the 6 to 8 grade band measures a student's ability to listen to and comprehend spoken English at a level sufficient to fully participate in and learn from grade-level instruction. The ELPA21 Reading test measures a student's ability to read and comprehend written English, also at grade level. The Speaking test measures a student's ability to produce speech in a variety of academic contexts that vary in length and complexity, and the Writing test measures the ability to produce written texts ranging, in grades 6 to 8, from brief communications to an extended essay.

The ISASP test specifications in ELA are provided in Welch & Dunbar (2022) and are aligned by grade level to the CCSS. The Reading test measures how well students can comprehend a variety of written materials. The assessment contains authentic reading passages of varying length and complexity consistent with common core grade-level standards. Both literary passages (e.g., fiction and folktales) and informational passages (e.g., expository science and social studies materials, procedural texts, and general nonfiction) are included. A significant number of questions may require students to draw inferences or to generalize about what they have read, and the questions reflect a variety of cognitive levels with respect to the depth of knowledge (DOK) required to comprehend the texts that students read. The Reading test provides teachers with information about the kinds of comprehension skills students are expected to develop as they read across the curriculum, engage with literature, and process information in materials encountered outside of school.

Questions associated with each reading passage require students to demonstrate understanding in a variety of content domains and cognitive levels associated with reading comprehension. For the tests in grades 6 to 8 of interest in this study, approximately 50 percent

of the items are in the CCSS domain of Key Ideas and Details, approximately 40 percent in Craft

and Structure, and the remaining 10 percent in Integration of Knowledge and Ideas. In addition,

the DOK levels of the items in the ISASP Reading test have been aligned with the specifications

of the Iowa Core Standards (Dickinson, Michaels & Thacker, 2018), the implementation of the

CCSS operative in the state of Iowa.

The ISASP Language/Writing test measures a student's ability to write in one of several

modes of expository discourse, for example, to argue in support of claims or to explain complex

ideas and information. Students are asked to produce coherent writing in which development,

organization and style are appropriate to task, purpose, and audience. Students' writing must be

supported by evidence and information from print sources provided as part of the writing task.

Students demonstrate command of the conventions of standard English with respect to usage,

mechanics, and spelling when writing and demonstrate understanding of how language functions

in different contexts to make effective choices for meaning and style. The Language/Writing test

also provides passage-based editing tasks in which students identify text that needs revision or

choose appropriate opening, closing and transitional sentences that are consistent in tone and

style of the passage, its discourse type and purpose.

Cognitive Processes (Standard 1.12)

The depth-of-knowledge (DOK) should be consistent with what is required by the

appropriate standards. For the ISASP, this would suggest consistency to the level of cognitive

complexity found in the CCSS. In grades 6 to 8, the average percents of items in each DOK

category were, in order from category 1 to 3, 20, 45 and 35 for ISASP Reading and 10, 20 and 70

for ISASP Language/Writing. Although items on ELPA21 are not coded for DOK, nor are there

test or blueprint specifications for DOK, the rubrics used for scoring responses on the Writing

and Speaking tests are designed to award the highest scores to written and spoken English

characterized by variety in syntactic structure, coherence in organization and development, and

relevance to the task. These aspects of language are consistent with level 3 in the DOK

framework (Office of Teaching, Learning and Assessment, 2021).

Internal Structure (Standard 1.13) Construct Irrelevance

Construct irrelevant variance may also occur when other features influence a student's

outcome on the test. Students who have limited fluency in English may inadvertently be

prevented from demonstrating what they know or what they can do on a test (Abedi, et al., 2000;

Haladyna & Rodriguez, 2013; Koenig & Bachman, 2004; Wolf, et al., 2008). Test

accommodations are used by students to help demonstrate their knowledge and skills when they

are identified as appropriate in an EL student's IEP or 504 Plan.

According to Iowa Testing Programs (2021), there is a three-tiered approach to

accessibility in the ISASP approach: universal features, designated features, and

accommodations. Examples of universal features available to all students include a bookmark

tool, a highlighter tool, and a notepad. In contrast, designated features are available to students

only when the need has been indicated in advance by a district test coordinator. The purpose of

designated features is to provide features that closely align with the instructional supports used in

the classroom; this allows for consistency of classroom instruction and assessment. For example,

designated features include a bilingual dictionary for Writing, color contrast display tools for

online tests, large-print test booklets for visual support, and text-to-speech or read-aloud audio

support when appropriate given a test's construct definition. Students may use a designated

feature or disable it depending on individual preference during testing. The final tier of

accessibility features are accommodations, which are available based on a student's IEP or 504

Plan. Accommodations give students fair and equitable access to assessment content. Examples

of accommodations include Braille and sign language interpretation.

Similar accessibility and accommodations features guide the administration of ELPA21,

and the same three-tiered approach of universal, designated, and accommodated features was

adopted (CRESST, 2021) by the Iowa Department of Education.

Internal Structure (1.15) Performance Level Descriptors

Achievement assessments such as ISASP and ELP assessments are predicated on

adherence to education standards and descriptors of performance that represent the claims they

support. In assessments used for accountability, it can be argued that the performance level

descriptors (PLDs) established by states to characterize proficiency represent the most concrete

definitions of the constructs intended for the assessment. PLDs take many forms, but in English

language assessments they are typically narrative statements of what students along a

performance continuum know and can do.

Because ELPA21 uses a fixed set of standards across all grade levels, rather than

standards for each individual grade, it is difficult to judge how the standards between ELPA21

and the Common Core standards in reading, language and writing align. This is a likely reason

there is limited literature addressing the alignment of ELPA21. However, Lee (2018) and

Wheeler & Davis-Becker (2022) provide examples of other sources of alignment for ELPA21

with the examination of achievement level descriptors at different grades. By examining these

descriptors for specific grades and content areas, both validity evidence for ELPA21 can be

collected and explanations for different patterns of test performance among students taking both tests can be explored.

In comparing the achievement level descriptors for the 6-8 grade band on ELPA21 in reading to the corresponding descriptors for the ISASP assessment at these three grade levels, several differences can be identified. For example, ELPA21 includes descriptors for all five levels of proficiency while ISASP includes descriptors only for students achieving a level of proficient or advanced. For the sake of understanding how well aligned these two tests are, we chose to compare the label 'proficient' on ISASP to the level of 4 or 'Early Advanced' on ELPA21 as this is the level at which students are considered ready to complete academic work at their grade level. Following from this, the advanced level 5 on ELPA21 should more closely align with the ISASP label of 'Advanced'. Another significant difference between the two tests is the treatment of literary and informational texts. ISASP differentiates between these two types of reading while ELPA21 does not.

While the ISASP proficiency level descriptors are more detailed than the corresponding ELPA21 achievement level descriptors, there are significant parallels between the two. For example, both sets of descriptors reference determining the meanings of words and phrases, with ELPA21 including 'idiomatic expressions' where ISASP includes 'figurative language'. Both sets of descriptors identify key actions that students should be able to take such as identifying details and summarizing texts. Similarly, both sets refer to understanding arguments in detail, with the ISASP advanced description stating that students can distinguish claims supported by evidence and ELPA21 level 5 refers to determining if evidence is sufficient to support claims.

Given that there are limited ways in which standardized tests such as the ELPA21 and ISASP can test the constructs presented in the performance descriptors, it is reasonable to

conclude based on the documentation provided in support materials that some degree of similarity for the construct of reading in grades 6 through 8 exists between these two tests.

Relationship with Related Constructs and Criteria (Standards 1.18 and 1.20)

Correlation and Regression Analyses   To understand the evidence base related to the concurrent validity of these instruments we examined both the correlations among scores and the best linear combination of the ELPA21 tests that predicts ISASP scores using standard regression techniques. In addition, classification consistency was examined using the proficiency benchmarks established for the two assessments. With respect to the ISASP, both the ELA composite score and the Reading score were of interest, the former because it represents the broadest coverage of the ELA standards in the CCSS and the latter because of the importance of reading skills in the access ELs have to academic content across the entire school curriculum. Note that because both ELPA21 and ISASP were spring administrations the correlation and regression results as well as the classification results presented here are interpreted as evidence related to concurrent validity.

Correlations among the ELPA21 tests and between those tests and ISASP ELA and Reading for grades 6 through 8 are reported in Table 4 with internal-consistency reliability estimates for each test in parentheses. The ELPA21 intercorrelations are high in all grades; they ranged from .69 to .91, and the highest values were among the Reading, Listening and Writing tests. Considering the reliability estimates, the rank-orders of examinees on these tests would be identical in all grades as the reliabilities of difference scores are essentially zero. This observed collinearity is also important in interpreting the regression results below. Another result of

interest in Table 4 is the markedly weaker relationship in all grades between ELPA21 Speaking

and the two ISASP measures as compared to the other ELPA21 tests.

Table 4

Correlations Among ELPA21 and ISASP Tests

| | | | ELPA21 | | | | ISASP | |
			Reading	Writing	Listening	Speaking	Reading	ELA Total
Grade 6	ELPA21	Reading	(.80)					
		Writing	.86	(.88)				
		Listening	.87	.87	(.90)			
		Speaking	.69	.78	.76	(.84)		
	ISASP	Reading	.66	.58	.60	.46	(.88)	
		ELA Total	.71	.67	.65	.53	.86	(.92)
Grade 7	ELPA21	Reading	(.82)					
		Writing	.87	(.90)				
		Listening	.89	.89	(.92)			
		Speaking	.74	.81	.78	(.86)		
	ISASP	Reading	.64	.56	.58	.46	(.88)	
		ELA Total	.69	.66	.64	.54	.87	(.93)
Grade 8	ELPA21	Reading	(.86)					
		Writing	.89	(.90)				
		Listening	.91	.91	(.83)			
		Speaking	.76	.80	.81	(.87)		
	ISASP	Reading	.62	.56	.58	.46	(.87)	
		ELA Total	.70	.67	.66	.55	.89	(.93)

Regression analyses treated ISASP ELA and ISASP Reading as dependent variables and

the four ELPA21 tests as independent variables. Standard regression diagnostics were examined

to support tests of linearity in the regression models as well as inferential tests related to the

stability and magnitude of regression weights. Tables 5 and 6 provide essential results of the

regression analyses for modeling ISASP ELA and ISASP Reading, respectively. The results in

these tables show evidence of the predictive influence of ELPA21 scores on ISASP scores and

evidence of concurrent validity (given that both measures were administered at the same time of the school year) of two sorts.

When the focus of the association between assessments is the broadly defined ELA domain of the Iowa Core Standards, the criterion variable of interest is the ISASP ELA composite score. The regression summary in Table 5 for this case indicates that ELAP21 Reading and ELPA21 Writing have significant regression weights, whereas ELAP21 Listening and Speaking do not contribute to the explained variance of the ISASP ELA composite. That is, both ELPA21 Reading and Writing scores contribute significantly to the explained variance of ISASP ELA when entered last in the regression equation, and the effect of ELPA21 Reading in standard units (beta weights equal to .51, .51 and .48 in grades 6, 7 and 8, respectively) is about twice the magnitude of the ELPA21 Writing effect (beta weights equal to .25, .24 and .20 in grades 6, 7 and 8, respectively). Correlations between observed ELA composite scores and those estimated from the regression equations in each of grades 6, 7 and 8 are uniform in magnitude at .72, .70 and .71, respectively. These findings are consistent with previous studies that found literacy skills measured by an ELP assessment were strongly predictive of EL student performance on a state content assessment (Cook, Linquanti, Chinen, & Jung, 2012; Parker, Louie, & O'Dwyer, 2009). The nonsignificant findings with respect to Listening also corroborate results related to the influence of a text-to-speech (TTS) feature available in ISASP discussed later in this report, namely that the use of TTS did not create any advantage to EL students on ISASP ELA.

When the focus of a concurrent validity argument is on the narrower ELA domain of reading, the criterion variable of interest is the ISASP Reading score and Table 6. The regression results for this case indicate that ELAP21 Reading and ELPA21 Listening, both tests that involve

content standards related to comprehension of written or spoken language, have significant

regression weights. There is little evidence from the results in Table 6 that the productive skills

in English measured by ELPA21 Writing and ELAP21 Speaking contribute to the explained

variance of the ISASP Reading score considered in isolation. Moreover, the regression weights

in Table 6 suggest the influence of ELPA21 Reading on ISASP Reading is three to five times

greater than the influence of ELPA21 Listening, again suggesting no influence of TTS.

Table 5

Regression Summary Table for Predicting ISASP ELA Score from ELPA21

Grade	N	Multiple R	Intercept	ELPA21 Regression Weights			
				Reading	Writing	Listening	Speaking
6	1663	.72	250.18	$.25^*$	$.11^*$	-.00	-.00
				.51*	*.25**	*-.01*	*-.01*
7	1649	.70	281.04	$.22^*$	$.09^*$	.00	-.01
				.51*	*.24**	*.00*	*-.03*
8	1610	.71	274.34	$.23^*$	$.09^*$	.04	-.02
				.48*	*.20**	*.08*	*-.03*

Note. Standardized (beta) weights in italics.
$^*p < .05$ as last variable entered

Table 6

Regression Summary Table for Predicting ISASP Reading Score from ELPA21

Grade	N	Multiple R	Intercept	ELPA21 Regression Weights			
				Reading	Writing	Listening	Speaking
6	1663	.67	263.99	$.28^*$	.00	$.07^*$	-.02
				.56*	*.00*	*.14**	*-.04*
7	1649	.64	290.39	$.27^*$	.00	.04	-.03
				.60*	*.01*	*.09*	*-.06*
8	1610	.63	285.73	$.27^*$	.01	$.07^*$	$-.04^*$
				.53*	*.03*	*.14**	*-.07**

Note. Standardized (beta) weights in italics.
$^*p < .05$ as last variable entered

Classification Consistency   To examine how consistently EL students were classified on ELPA21 and ISASP, a classification consistency analysis for each of the three grade levels was conducted using proficiency on ISASP ELA and overall ELPA21, presented in Table 7, as well as proficiency on ISASP Reading and ELPA21 Reading only, presented in Table 8. In the former case, results showed that students in grades 6, 7 and 8 were classified consistently 83, 88 and 83 percent of the time, respectively. With respect to inconsistent classification, those of greatest interest here are the so-called false negatives, i.e. those students not proficient overall on ELPA21 but proficient on the ISASP ELA composite score. These are students whose ELPA21 scores were not sufficient to exit EL instruction but whose ISASP scores met the standard of ELA proficiency for all students. The false negatives were 12, 8 and 15 percent of all EL students in grades 6, 7 and 8, respectively. Results in Table 8 show a similar pattern for ISASP Reading. Correct classification percents were 80 or above in all grades, and misclassifications were predominately false negatives.

Table 7

Classification Results for ELPA21 Overall and ISASP ELA

ISASP ELPA21		Not Proficient	Proficient
Grade 6	Not Proficient	1229 (74)	*200 (12)*
	Proficient	88 (5)	146 (9)
Grade 7	Not Proficient	1392 (84)	*138 (8)*
	Proficient	56 (3)	63 (4)
Grade 8	Not Proficient	1257 (78)	*247 (15)*
	Proficient	27 (2)	79 (5)

Table 8

Classification Results for ELPA21 and ISASP in Reading

ELPA21 \ ISASP		Not Proficient	Proficient
Grade 6	Not Proficient	1,117 (67)	*179 (11)*
	Proficient	146 (9)	221 (13)
Grade 7	Not Proficient	1,360 (82)	*94 (6)*
	Proficient	108 (7)	87 (5)
Grade 8	Not Proficient	1,303 (81)	*155 (10)*
	Proficient	58 (4)	94 (6)

To better understand the classification results, ISASP score distributions were examined by levels of performance on ELPA21. Figure 1 shows the distribution of ISASP ELA scores for EL students by five performance levels of the ELPA21 Reading assessment for grade 6. The box-and-whiskers plot for each performance level identifies the middle 50 percent of the distribution as a rectangular box and the remaining 25 percent as vertical line segments extending up and down (single points represent outlying observations). Recall that Level 4 (Early Advanced) on each ELPA21 domain is considered a criterion for obtaining the overall English proficiency standard. As illustrated in the figure, approximately 50% of EL students in grades 6 who obtained Level 4 on ELPA21 Reading did not exceed the academic standard for proficiency on ISASP Reading (indicated by the horizontal dashed line). Note that just over 75 percent of the EL students at Level 5 (Advanced) on ELPA21 Reading exceeded the ISASP ELA proficiency standard in grade 6. It becomes clear in grade 6 that many, though not all, ELs deemed proficient on ELPA21 Reading were also judged proficient on the ISASP ELA assessment designed for all students, particularly if they achieved Level 5 on ELPA21. Conversely, most EL students at not proficient Levels 1 through 3 on ELPA21 also were judged to be not proficient on ISASP ELA. The parts of the distributions below the ISASP proficiency line at Levels 1 through 3 and above

the ISASP proficiency line at Levels 4 and 5 represent consistent classifications in Table 7.

Figure 2 displays similar results for ISASP Reading.

     Also of interest in the boxplots are the observations at ELPA21 Levels 1 through 3 above

the horizontal line. These represent EL students not proficient on ELPA21 Reading although

they exceeded the proficiency cut score on the ISASP ELA or Reading assessment. Although

only a few such cases exist at Levels 1 and 2, about 20 percent of the students at Level 3 on

ELPA21 Reading were proficient on ISASP ELA or ISASP Reading. Such observations

constitute the bulk of the false negatives in grade 6 discussed in the classification consistency

results presented in Tables 7 and 8, that is, students perhaps just below the ELPA21 proficiency

standard in Reading and above the corresponding ISASP standard. Similar findings in grades 7

and 8 are presented in the boxplots in the appendix

Figure 1

ISASP ELA Distributions at Five Performance Levels of ELPA21 Reading in Grade 6
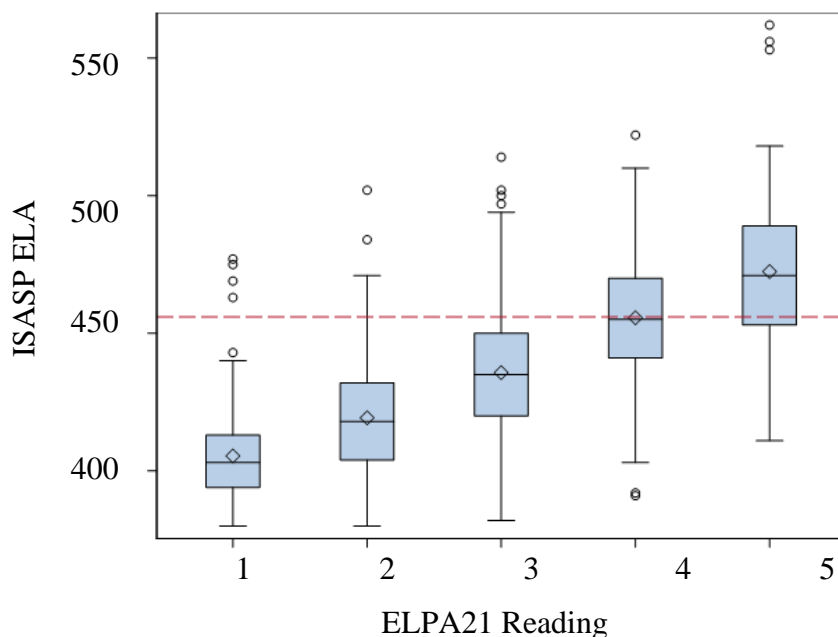
Figure 2

ISASP Reading Distributions at Five Performance Levels of ELPA21 Reading in Grade 6
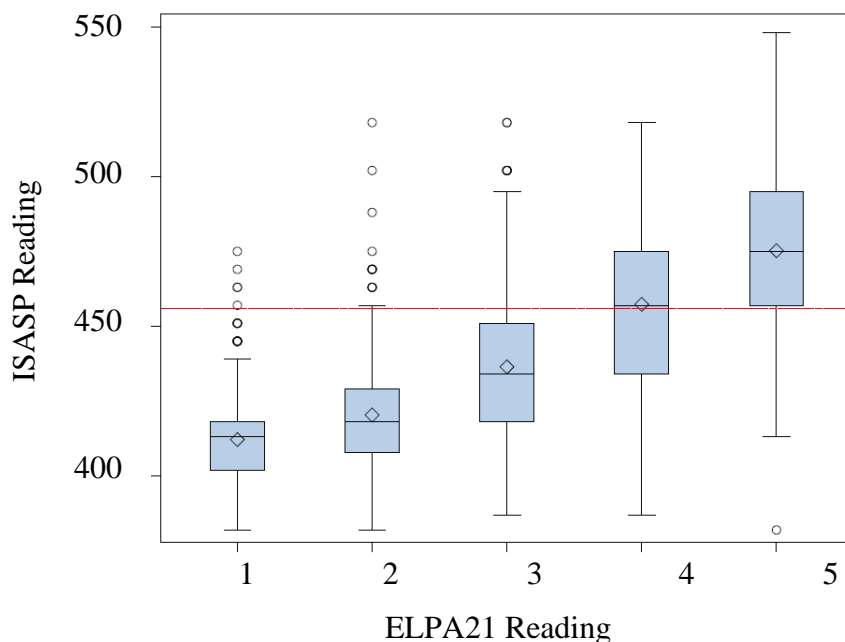
Construct Considerations in Reading

Recognizing that there are students who were proficient on ISASP ELA and Reading but not proficient on ELPA21 Reading, concerns can be raised about the comparability of ISASP and ELPA21 constructs given differences in certain administrative conditions of the two assessment programs. Perhaps most salient in this regard are the ISASP text-to-speech (TTS) tools. Text-to-speech is a designated feature in grades 6 and above that provides students with the option to listen to pre-generated text-to-speech content that is synchronized with text highlighting on the screen. This tool is available on the ISASP assessments in Reading and Language/Writing in grades 6 to 8 because the Iowa Core Standards do not explicitly include the Foundations of Language strand beyond fifth grade. However, ELPA21 does not provide text-to-speech (National Center for Research on Evaluation, Standards, and Student Testing, 2021) or

the human parallel of read aloud protocols in the reading component of the assessment (read-aloud protocols are available in the other three ELPA21 tests as a designated feature). If ELs were able to access and use the TTS feature on ISASP Reading, the contribution of listening to their scores could effectively reduce score variance related to the processing of printed text, which is still required on ELPA21 Reading, possibly causing them to achieve proficiency in reading on ISASP but not on ELPA21.

Table 9 shows, for the ELs in grades 6 to 8 who were proficient on ISASP and not proficient on ELPA21, the TTS usage on the ISASP Reading assessment. The total numbers of students who used the TTS function and those who did not are 323 and 249, respectively. Specifically, the number of students in grade 6 who were proficient on 2021 ISASP Reading but are not proficient on ELPA21 Reading is 197. Among them, 112 students used the TTS function, while 85 did not. The number of students in grade 7 and 8 who were proficient on ISASP Reading but not proficient on ELPA21 is 135 and 240, of which 81 and 130 students used text-to-speech, while 54 and 110 did not, respectively.

Table 9

Number (Percent) of Students Receiving Text-to-Speech on ISASP Reading

Grade	Text-to-Speech	No Text-to-Speech	Total
6	112 (57)	85 (43)	197
7	81 (60)	54 (40)	135
8	130 (54)	110 (46)	240
Total	323 (57)	249 (43)	572

Figure 4 illustrates the essential result in the grade 6 sample of a logistic regression analysis to investigate the extent to which using the TTS functionality on ISASP Reading influenced the probability that a student was proficient given the score on ELPA21 Reading. In

the logistic regression model, main effects for ELPA21 Reading and TTS and their interaction

were independent variables predicting the ISASP proficiency as a binary outcome. In the model,

both the TTS main effect and the interaction effect were non-significant. Figure 4 displays the

separate regression functions for the model with main effects for ELPA21 Reading score and

TTS use. As can be seen in the figure, the separate functions are virtually coincident, and the

non-significant main and interaction effects involving TTS suggest that they are coincident from

the standpoint of inference to a population of test-takers. Similar analyses and results (see

Appendix A) were obtained for the samples in grades 7 and 8 as well, supporting the argument

that TTS did not introduce a degree of construct irrelevant variance sufficient to account for the

discrepancy in proficiency classifications for the two instruments.

Figure 3

Conditional Probability of a Proficient ISASP Score in Reading given ELPA21 Reading Score in Grade 6 by TTS Use.
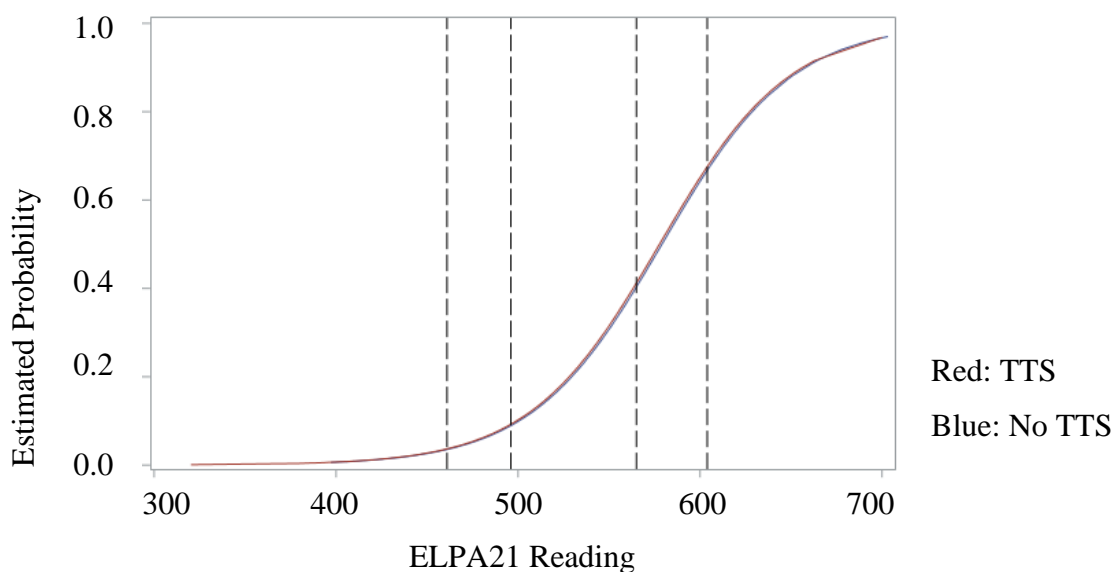
Note: Dashed lines indicate boundaries between ELPA21 achievement levels

Discussion

There is obvious utility in being able to make a distinction between an English learner's

ability within a certain content domain and overall proficiency in the English language

depending on the purpose and use of assessment results. Without an accurate overall assessment

of language proficiency, a valid interpretation of the results of other benchmarks is difficult if not

impossible. Yet, there is also reason for concern when EL students are assessed repeatedly by

multiple instruments that may have been developed based on overlapping construct definitions.

Pandya (2011) points out that educators and administrators having access to such multiple

sources of similar information does not necessarily support them in providing help to their

students (p. 24). Rather, he argues the overlapping scores produce confusion as it is not always

clear which assessment is most accurately measuring their students' abilities. This could lead, for

example, to greater emphasis on instruction in reading when it is not needed or less focus on

reading when it is needed.

Of particular interest with respect to potential overlap between ELPA21 and ISASP scores

in this study is the strong evidence of concurrent validity in grades 6 to 8 presented in the

regression analyses. In all grades, the ELPA21 Reading and Writing scores were significantly

related to the ISASP ELA composite score, on which proficiency designations are based for all

students, whereas the ELPA21 Listening and Speaking scores were not. One explanation for

these findings can be derived from a more detailed examination of the content correspondences

between the ELP Standards and the content domains and stem types used by the two

assessments. These correspondences are presented in Table 10 in which the rows identify the

ELP Standards, and the columns describe either the ISASP domains (Welch & Dunbar, 2019) or

the descriptions of stem types used in the ELPA21 Reading and Writing tests (Office of

Teaching, Learning and Assessment, 2021). The shaded cells represent the commonalities of

content in the Reading and Writing tests in terms of ISASP domains and ELPA21 stem types for

each ELP standard.

The strongest content correspondence in grades 6 to 8 between ISASP and ELPA21 was

observed in reading. As can be seen in Table 10, the reading assessments show considerable

overlap in content with respect to ELP Standards 1 and 8. The bulleted items in reading

constitute the bulk of the content coverage for ISASP and ELPA21, and the correspondence

between ISASP domains and ELPA21 stem types is strong. For example, ISASP's Key Ideas and

Details items include questions about main ideas of passages and paragraphs within them, about

point of view and argumentation, and about inferential meaning (Iowa Testing Programs, 2019).

Similar correspondences hold for the other ISASP domains, e.g. structural features used by

authors to convey purpose and meaning, as well as connections between graphical information in

sidebars and ideas in the text.

Although the ELPA21 Writing test was significantly related to the ISASP ELA composite

score, the association was not as strong as that of the ELPA21 Reading test. Nevertheless, there

is evidence in the specifications of ELPA21 Writing and ISASP Language/Writing to support

similar lines of reasoning as to why the significant relationship was observed. In the writing

strands of the ELP Standards, three of the four ISASP domain scores emphasize the productive

aspects of writing skills articulated in the evidenced-based prompts and the scoring rubrics,

which concentrate on discourse aspects of student responses (Iowa Testing Programs, 2019),

including organizing and developing an argument or using evidence to support an idea.

Similarly, four of the five stem types in ELPA21 Writing emphasize productive skills in the

creation of text to serve a purpose in writing or to respond to a point provided in stimulus

Table 10

Content Correspondences among ELP Standards and ISASP-ELPA21 Domains and Stem Types

English Language Proficiency Standards	ISASP Reading	ELPA21 Reading	ISASP Language/Writing	ELPA21 Writing
1. Construct meaning from oral presentations and literary and informational text through grade-appropriate listening, reading, and viewing	• Key Ideas and Details • Craft and Structure • Integration of Knowledge and Ideas	• Main Idea, Argument, Key Detail, Inference • Author's Purpose, Rhetorical Structure • Summarize Graphics		
3. Speak and write about grade-appropriate complex literary and informational texts and topics			• Production and Distribution of Writing • Text Types and Purposes	• Respond to Peer • Storyboard • Construct a Claim
4. Construct grade-appropriate oral and written claims and support them with reasoning and evidence			• Text Types and Purposes • Production and Distribution of Writing	• Construct a Claim • Respond to Peer

6. Analyze and critique the arguments of others orally and in writing			• Text Types and Purposes	• Construct a Claim
7. Adapt language choices to purpose, task, and audience when speaking and writing			• Knowledge of Language • Production and Distribution of Writing	• Discrete Editing
8. Determine the meaning of words and phrases in oral presentations and literary and informational text	• Vocabulary Acquisition and Use	• Vocabulary in Context		
9. Create clear and coherent grade-appropriate speech and text			• Knowledge of Language • Production and Distribution of Writing	• Discrete Editing • Writing Questions • Respond to Peer • Storyboard
10. Make accurate use of standard English to communicate in grade-appropriate speech and writing			• Conventions of Standard English	• Discrete Editing

materials such as a peer's email inquiry or a stated position on an issue (Office of Teaching, Learning and Assessment, 2021). In addition, both assessments include multiple choice items that measure knowledge of language and discrete language skills related to the use of Standard English conventions. In the writing strands of the ELP Standards, the content overlap between ELPA21 Writing and the ISASP Language/Writing assessment is substantial.

The ELPA21 tests not included in Table 10, Listening and Speaking, do overlap with some of the ISASP ELA domain scores, although to a much lesser degree. For example, the stem types in ELPA21 Listening are the same as those in ELPA21 Reading, and from the perspective of comprehension represent comparable and somewhat overlapping receptive language skills, although the difference between spoken text and written text is likely responsible for the stronger evidence of concurrent validity between ELPA21 Reading and ISASP ELA. Although the same can be said of certain stem types used in ELPA21 Speaking (Office of Teaching, Learning and Assessment, 2021), again the difference between spoken and written language supports a weaker argument with respect to concurrent validity. Small or non-significant effects for ELPA21 Listening and Speaking in the regression analyses are consistent with this idea.

With respect to the determination of proficiency in English, misclassification of ELs has been known to have serious effects on their learning opportunities (Carroll & Bailey, 2016). A non-proficient student who is classified as proficient may lose access to needed language supports. Conversely, a proficient EL who is classified as non-proficient may continue to receive dedicated language instruction when it is not needed, have fewer opportunities to engage with the same level of subject matter content as native English speaking peers, and ultimately experience reduced access to advanced courses in the regular school curriculum (Gandara et al., 2003).

Classification of ELs as fully English proficient is consequential, affecting the type of instruction the students are exposed to, the standards they are held to, and the amount of testing they are subjected to (Solorzano, 2008). Evidence suggests ELs spend many years in EL programs before they are classified as fully English proficient. Data from 2007 in Los Angeles Unified School District showed that only 61 percent of EL students were classified as fully proficient after more than six years in the program (Solorzano, 2008). Similarly, ELPA21 results for ELs in Oregon showed that students with an initial domain proficiency at the lowest level can take as long as six years to exit an EL program, and students with an initial proficiency level in the middle category can take as many as three years to exit (Office of Research, Analysis, and Accountability, 2019). Any errors in classifying students as not fully proficient only add to the amount of time that they are kept in these programs, which can have deleterious effects on future academic progress.

As a matter of assessment policy and school accountability, it is difficult to argue solely on the evidence of concurrent validity presented in this report that ISASP and ELPA21 duplicate each other to such a degree that policymakers should consider eliminating one or the other for students in the middle grades analyzed here. The federal mandates driving the two assessments are intended to address a broad array of program- and school-related evaluation concerns. Nevertheless, the evidence supporting substantial overlap in both the constructs and specific assessment tasks presented to ELs, particularly in reading, suggest some instructional time for ELs could be restored without loss of information regarding English language skills important to student growth. For example, many students in the ELPA21 test taking population in the middle grades have achieved proficiency on some, though not all, ELPA21 tests, or they remain ELPA21 test takers for purposes of monitoring skills rather determining proficiency. Perhaps an ISASP assessment in reading, required of such students anyway, would be sufficient for future

ELA proficiency determinations or for continued monitoring of language skills. The results of this study argue that considerations such as these, which would reduce testing time and student anxiety as well as conserve both human and financial assessment resources, should be weighed carefully by policymakers interested in assessment best practice and the general welfare of the EL student population.

References

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.

American Educational Research Association, American Psychological Association, & National Council on Measurement in Education (Eds.). (2014). *Standards for educational and psychological testing*. American Educational Research Association.

Cambium Assessment (2020). *Technical report Part I – Summative assessment, English language proficiency assessment for the 21st century – Listening, Reading, Speaking, Writing*. Cambium. https://www.louisianabelieves.com/docs/default-source/assessment/2019-2020-elpt-and-elps-technical-report-part-i-assessment-overview.pdf?sfvrsn=2c036718_2

Bailey, A. L., & Wolf, M. K. (2020). The construct of English language proficiency in consideration of college and career readiness standards. In M.K. Wolf (Ed.), *Assessing English language proficiency in US K–12 schools* (pp. 36-54). Routledge.

Beddow, P. A. (2012). Accessibility theory for enhancing the validity of test results for students with special needs. *International Journal of Disability, Development and Education*, *59*(1), 97–111.

Carroll, P. E., & Bailey, A. L. (2016). Do decision rules matter? A descriptive study of English language proficiency assessment classifications for English-language learners and native English speakers in fifth grade. *Language Testing*, *33*(1), 23–52. https://doi.org/10.1177/0265532215576380

Cook, H. G., Linquanti, R., Chinen, M., & Jung, H. (2012). *National evaluation of Title III implementation supplemental report: Exploring approaches to setting English language*

proficiency performance criteria and monitoring English learner progress. U.S.

Department of Education.

Dickinson, E. R., Michaels, H. R. & Thacker, A. A. (2018). *Evaluation of the Iowa Statewide Assessment of Student Progress ELA, mathematics, and science assessments: Grades 3-11 alignment report*. Human Resources Research Organization Report 2018 No. 070. HumRRO.

Every Student Succeeds Act, 20 U.S.C. § 6301 (2015). https://www.congress.gov/bill/114th-congress/senate-bill/1177 Section 111 (b)(2)(F).

Gandara, P., Rumberger, R., Maxwell-Jolly, J., & Callahan, R. (2003). English learners in California schools: Unequal resources, unequal outcomes. *Education Policy Analysis Archives*, *11*(36), 1–54.

Haladyna, T. M., & Rodriguez, M. C. (2013). *Developing and validating test items*. Routledge.

Huang, B. H., & Flores, B. B. (2018). The English Language Proficiency Assessment for the 21st Century (ELPA21). *Language Assessment Quarterly*, *15*(4), 433-442. https://doi.org/10.1080/15434303.2018.1549241

Iowa Department of Education (2018). *Every student succeeds act in Iowa: Final plan*. Des Moines, IA: Author. https://educateiowa.gov/sites/files/ed/documents/2019-12-23ESSAPlan_508.pdf.

Iowa Testing Programs (2019). *Iowa statewide assessment of student progress: English language arts test specifications, grades 3-11*. University of Iowa. https://iowa.pearsonaccess.com/resources/research/ISASPTestSpecifications_ELA.pdf

Iowa Testing Programs (2021). *Iowa statewide assessment of student progress: Accessibility and accommodations manual*. University of Iowa. https://iowa.pearsonaccess.com/resources/manuals/IA1133018_Accom_Sp22_WEB.pdf

Koenig, J. A., & Bachman, L. F. (2004). Executive Summary. In Koenig, J. A. & Bachman, L. F.

    (Eds.) *Keeping Score for All: The Effects of Inclusion and Accommodation Policies on*

    Large-Scale Educational Assessment (pp. 1-9). National Academies Press.

Lee, O. (2018). English language proficiency standards aligned with content standards.

    Educational Researcher, *47*(5), 317-327.

Lee, D., Buzick, H., Sireci, S. G., Lee, M., & Laitusis, C. (2021). Embedded accommodation and

    accessibility support usage on a computer-based statewide achievement test. *Practical*

    Assessment, Research & Evaluation, *26*(25). http://pareonline.net/getvn.asp?v=26&n=25

National Center for Research on Evaluation, Standards, and Student Testing. (2021). *ELPA21*

    accessibility and accommodations manual. https://iowaelpa21.portal.cambiumast.com/-

    /media/project/client-portals/iowa/pdf/accessibility-and-accommodations-

    manual_updated.pdf

National Governors Association Center for Best Practices & Council of Chief State School

    Officers. (2010). *Common Core State Standards*. Washington DC: Authors.

Office of Research, Analysis, and Accountability. (2019). *Accountability details policy and*

    technical manual. Oregon Department of Education.

    https://www.oregon.gov/ode/schools-and-

    districts/reportcards/reportcards/Documents/Accountability_Policy_Technical_Manual_1

    819.pdf

Office of Teaching, Learning and Assessment (2021). *ELPA21 test specifications and blueprints:*

    Grades 6-8. Oregon Department of Education.

    https://digital.osl.state.or.us/islandora/object/osl%3A977733/datastream/OBJ/view

Pandya, J. Z. (2011). *Overtested: How high-stakes accountability fails English language*

    learners. Teachers College Press.

Parker, C. E., Louie, J., & O'Dwyer, L. (2009). *New measures of English language proficiency*

    and their relationship to performance on large-scale content assessments. Issues and

    Answers. Report, REL 2009–No. 066. Washington, DC: U.S. Department of Education,

    Institute of Education Sciences, National Center for Education Evaluation and Regional

    Assistance, Regional Educational Laboratory Northeast, and Islands.

    http://ies.ed.gov/ncee/edlabs.

Solorzano, R. W. (2008). High stakes testing: Issues, implications, and remedies for English

    language learners. *Review of Educational Research*, *78*(2), 260–329.

    https://doi.org/10.3102/0034654308317845

Welch, C. & Dunbar S. (2022). *Iowa statewide assessment of student progress: Technical*

    manual, Version 2.0. Iowa Testing Programs.

    https://iowa.pearsonaccess.com/resources/research/ISASP_TechnicalManual_Version2_0

    .pdf

Wheeler, K. & Davis-Becker, S. (2022). *Evaluating correspondence: A process for evaluating*

    the alignment of ELP standards and academic content standards. Paper presented at the

    annual meeting of the National Council on Measurement in Education, San Diego.

Wolf, M. K., Farnsworth, T., & Herman, J. (2008). Validity issues in assessing English language

    learners' language proficiency. *Educational Assessment*, *13*(2-3), 80-107

Wolf, M. K., Kao, J., Griffin, N., Herman, J. L., Bachman, P. L., Chang, S. M., & Farnsworth, T.

    (2008). *Issues in assessing English language learners: English language proficiency*

    measures and accommodation uses. Practice Review (Part 2 of 3). CRESST Report 732.

    National Center for Research on Evaluation, Standards, and Student Testing (CRESST).

Appendix A

Tables and Figures for Grade Groups 4-5 and 9-11

This appendix contains empirical results for ELPA21 grade groups 4-5 and 9-11 that parallel the results for grade group 6-8 presented in Tables 4 through 9 and Figures 1 through 3 in the report. In general, the results reported here support the empirical findings related to concurrent validity for grades 6 though 8, as well as the validity argument presented related to comparability between the ELP assessment and the general ELA assessment operationalized in the ELPA21 and ISASP testing programs.

Tables and figures in the appendix are numbered to correspond to similarly numbered tables and figures in the report. For example, Table A4 in this appendix matches the results in Table 4 in the report, and Figure A1 matches the results in Figure 1, etc.

Table A4

Correlations Among ELPA21 and ISASP Tests

			ELPA21				ISASP	
			Reading	Writing	Listening	Speaking	Reading	ELA Total
Grade 4	ELPA21	Reading	(.83)					
		Writing	.91	(.86)				
		Listening	.87	.87	(.85)			
		Speaking	.71	.78	.75	(.86)		
	ISASP	Reading	.71	.64	.61	.47	(.86)	
		ELA Total	.76	.72	.67	.53	.89	(.83)
Grade 5	ELPA21	Reading	(.85)					
		Writing	.92	(.87)				
		Listening	.88	.89	(.86)			
		Speaking	.73	.79	.77	(.88)		
	ISASP	Reading	.66	.59	.56	.40	(.87)	
		ELA Total	.70	.65	.62	.49	.88	(.83)

			ELPA21				ISASP	
			Reading	Writing	Listening	Speaking	Reading	ELA Total
Grade 9	ELPA21	Reading	(.82)					
		Writing	.88	(.88)				
		Listening	.93	.92	(.89)			
		Speaking	.75	.80	.80	(.89)		
	ISASP	Reading	.59	.49	.55	.44	(.85)	
		ELA Total	.72	.64	.68	.55	.86	(.86)
Grade 10	ELPA21	Reading	(.84)					
		Writing	.90	(.86)				
		Listening	.94	.93	(.88)			
		Speaking	.75	.78	.79	(.87)		
	ISASP	Reading	.66	.59	.62	.48	(.85)	
		ELA Total	.75	.70	.71	.57	.90	(.87)
Grade 11	ELPA21	Reading	(.85)					
		Writing	.90	(.83)				
		Listening	.95	.93	(.87)			
		Speaking	.74	.76	.78	(.86)		
	ISASP	Reading	.67	.63	.63	.46	(.87)	
		ELA Total	.72	.69	.69	.52	.91	(.88)

Table A5

Regression Summary Table for Predicting ISASP ELA Score from ELPA21

Grade	N	Multiple R	Intercept	ELPA21 Regression Weights Reading	Writing	Listening	Speaking
4	2117	.77	257.36	.23*	.09*	-.00	-.03*
				*.60**	*.26**	*-.01*	*-.10**
5	1775	.70	278.37	.24*	.03	.00	-.03*
				*.66**	*.09*	*.01*	*-.08**
9	1715	.72	275.35	.34*	.01	.00	.01
				*.68**	*.02*	*.00*	*.02*
10	1765	.75	305.60	.31*	.05*	.01	-.01
				*.65**	*.11**	*.01*	*-.01*
11	1354	.73	326.70	.31*	.14*	-.04	-.04*
				*.61**	*.27**	*-.08*	*-.08**

Note. Standardized (beta) weights in italics.
*$p < .05$ as last variable entered

Table A6

Regression Summary Table for Predicting ISASP Reading Score from ELPA21

Grade	N	Multiple R	Intercept	ELPA21 Regression Weights Reading	Writing	Listening	Speaking
4	2117	.72	256.85	.30*	.02	.01	-.04*
				*.73**	*.04*	*.01*	*-.10**
5	1775	.67	289.80	.29*	-.01	.01	-.06*
				*.77**	*-.02*	*.03*	*-.17**
9	1715	.59	311.26	.32*	-.09*	.05	.01
				*.63**	*-.17**	*.10*	*.02*
10	1765	.66	321.76	.34*	-.02	.04	-.02
				*.64**	*-.03*	*.08*	*-.03*
11	1354	.68	334.75	.34*	.11*	-.03	-.06*
				*.62**	*.19**	*-.05*	*-.10**

Note. Standardized (beta) weights in italics.
*$p < .05$ as last variable entered

Table A7

Classification Results for ELPA21 Overall and ISASP ELA

ELPA21 / ISASP		Not Proficient	Proficient
Grade 4	Not Proficient	1425 (67)	*358 (17)*
	Proficient	72 (3)	262 (12)
Grade 5	Not Proficient	1449 (82)	*135 (8)*
	Proficient	96 (5)	95 (5)
Grade 9	Not Proficient	1411 (82)	*231 (13)*
	Proficient	20 (1)	53 (3)
Grade 10	Not Proficient	1293 (73)	*324 (18)*
	Proficient	35 (2)	113 (6)
Grade 11	Not Proficient	1050 (78)	*159 (12)*
	Proficient	61 (5)	84 (6)

Table A8

Classification Results for ELPA21 and ISASP in Reading

ELPA21 / ISASP		Not Proficient	Proficient
Grade 4	Not Proficient	1383 (65)	*169 (8)*
	Proficient	189 (9)	376 (18)
Grade 5	Not Proficient	1267 (71)	*135 (8)*
	Proficient	175 (10)	198 (11)
Grade 9	Not Proficient	1371 (80)	*208 (12)*
	Proficient	48 (3)	88 (5)
Grade 10	Not Proficient	1108 (63)	*415 (24)*
	Proficient	47 (3)	195 (11)
Grade 11	Not Proficient	951 (70)	*195 (14)*
	Proficient	77 (6)	131 (10)

Table A9

Number (Percent) of Students Receiving Text-to-Speech on ISASP Reading

Grade	Text-to-Speech	No Text-to-Speech	Total
9	108 (48)	116 (52)	224
10	155 (51)	152 (49)	307
11	72 (46)	84 (54)	156
Total	335 (49)	352 (51)	687

Note. Text-to-Speech is not available on ISASP Reading in Grades 4 and 5.

Figure A1.1

ISASP ELA Distributions at Five Performance Levels of ELPA21 Reading in Grade 4
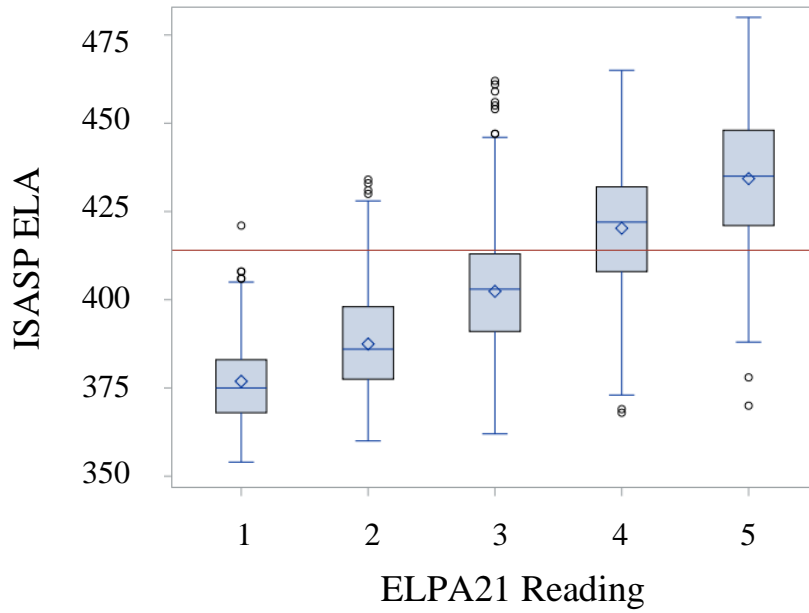
Figure A1.2

ISASP ELA Distributions at Five Performance Levels of ELPA21 Reading in Grade 5
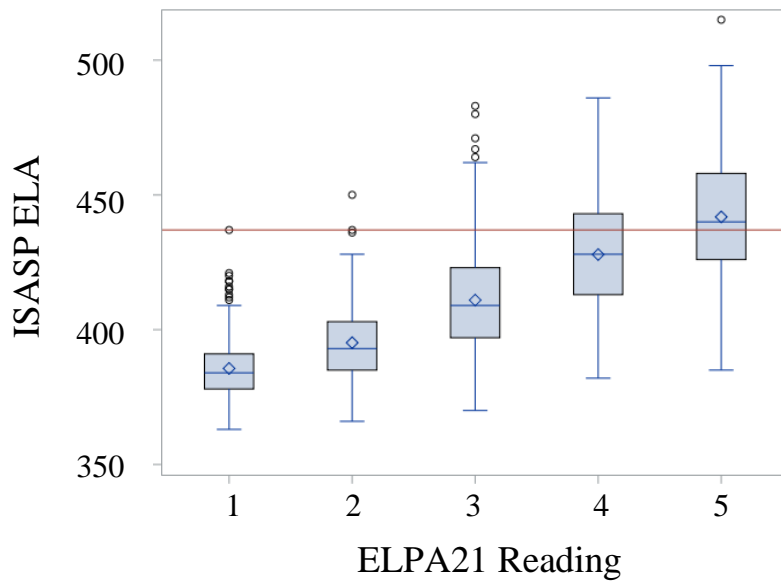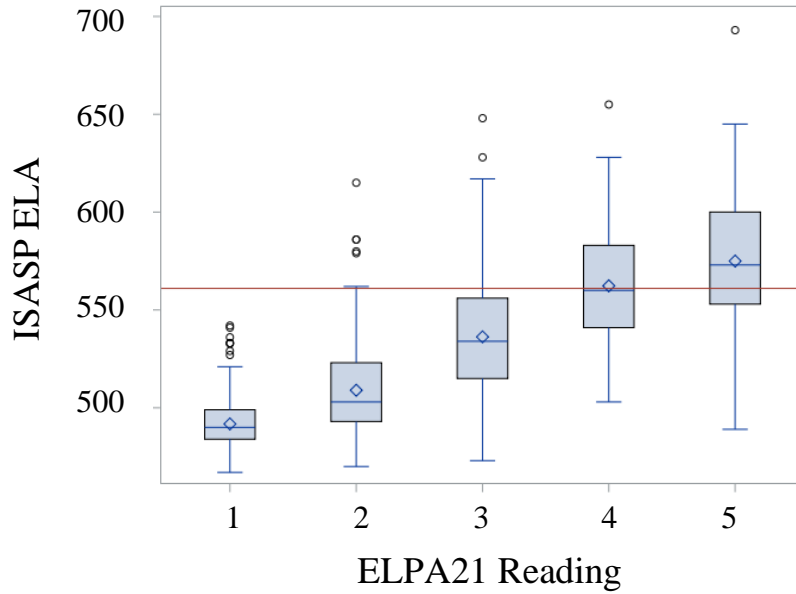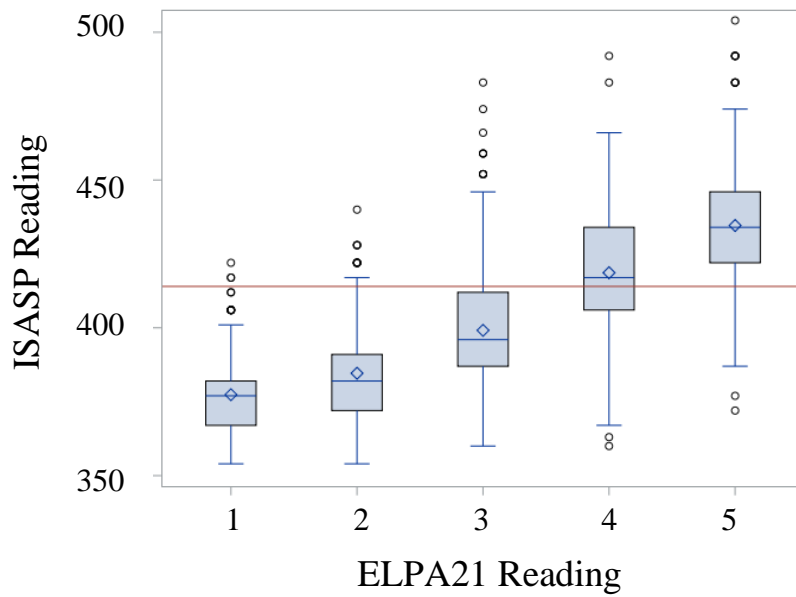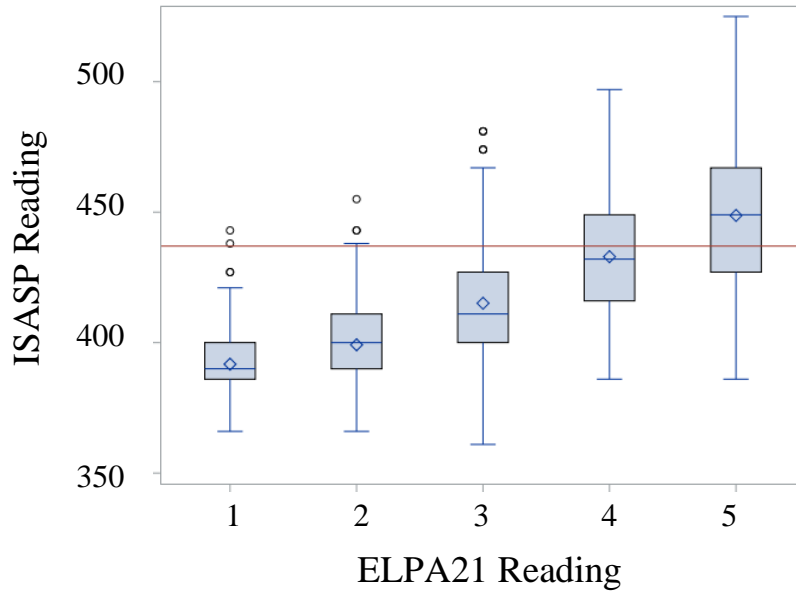
Figure A1.3

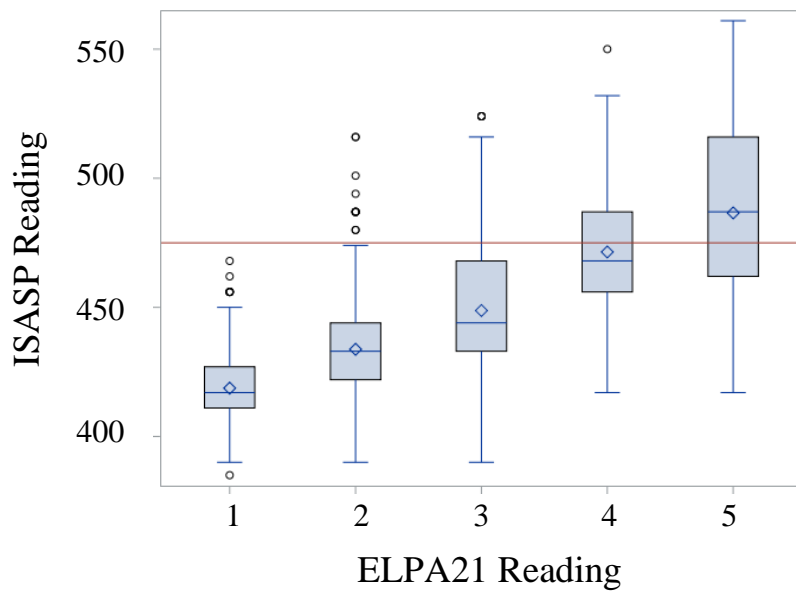*ISASP ELA Distributions at Five Performance Levels of ELPA21 Reading in Grade 7*
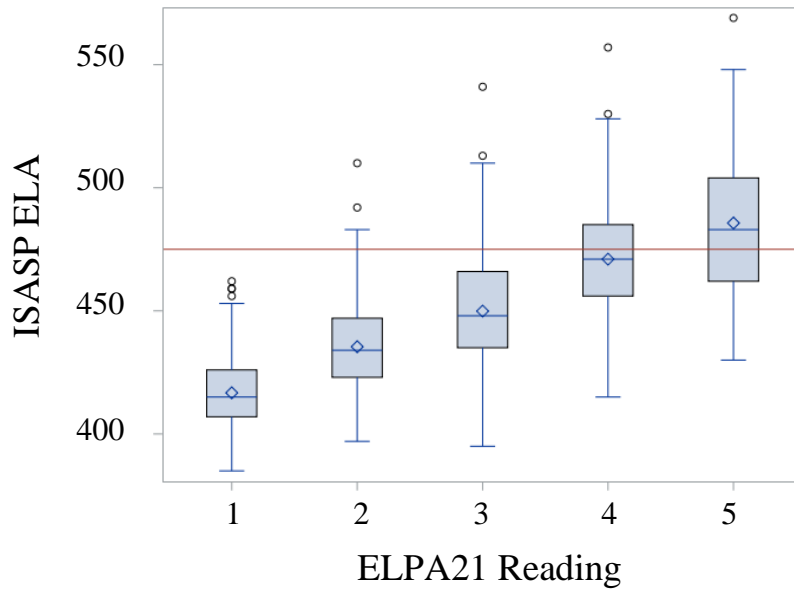
Figure A1.4

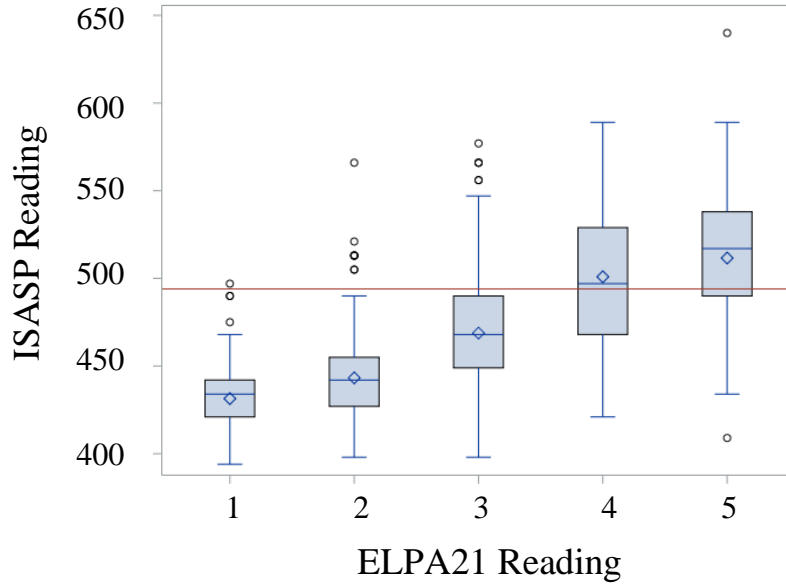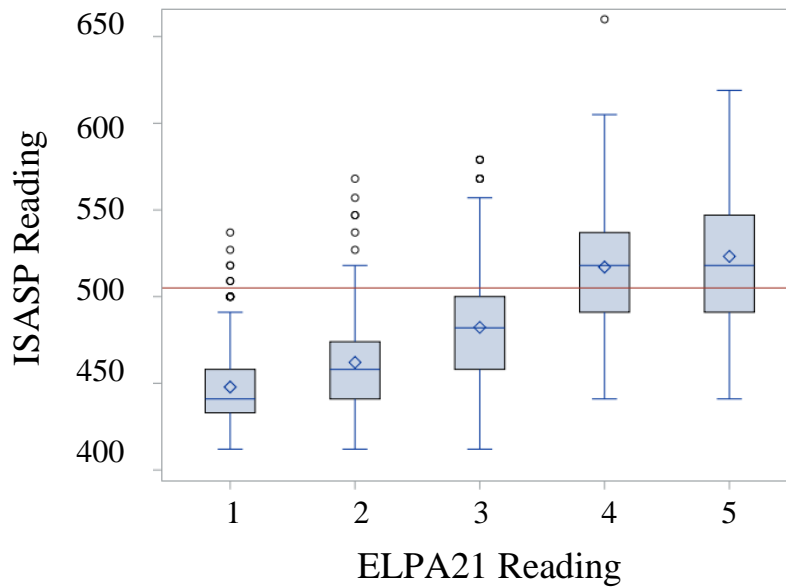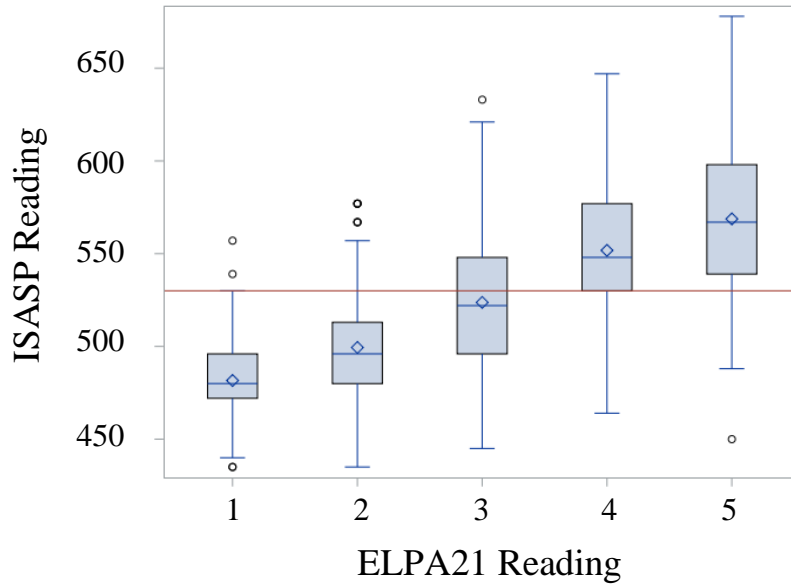*ISASP ELA Distributions at Five Performance Levels of ELPA21 Reading in Grade 8*
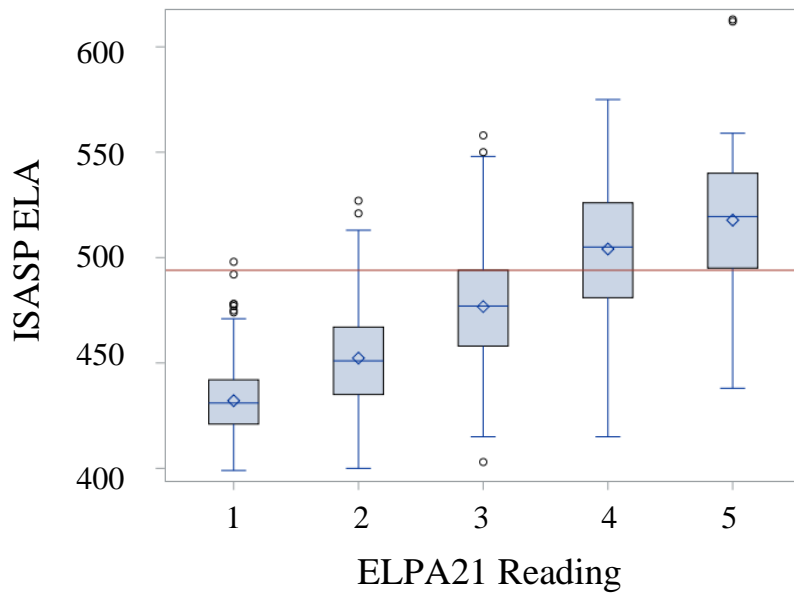
Figure A1.5

ISASP ELA Distributions at Five Performance Levels of ELPA21 Reading in Grade 9
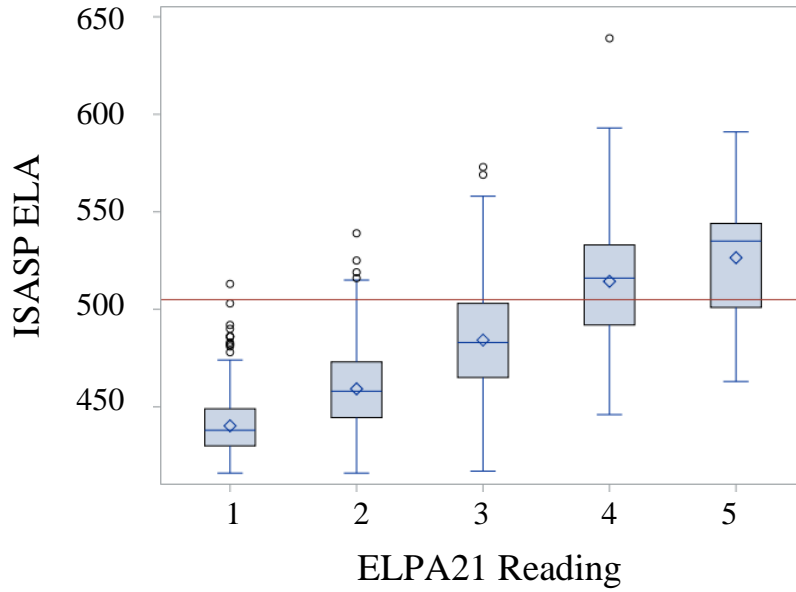
Figure A1.6

ISASP ELA Distributions at Five Performance Levels of ELPA21 Reading in Grade 10
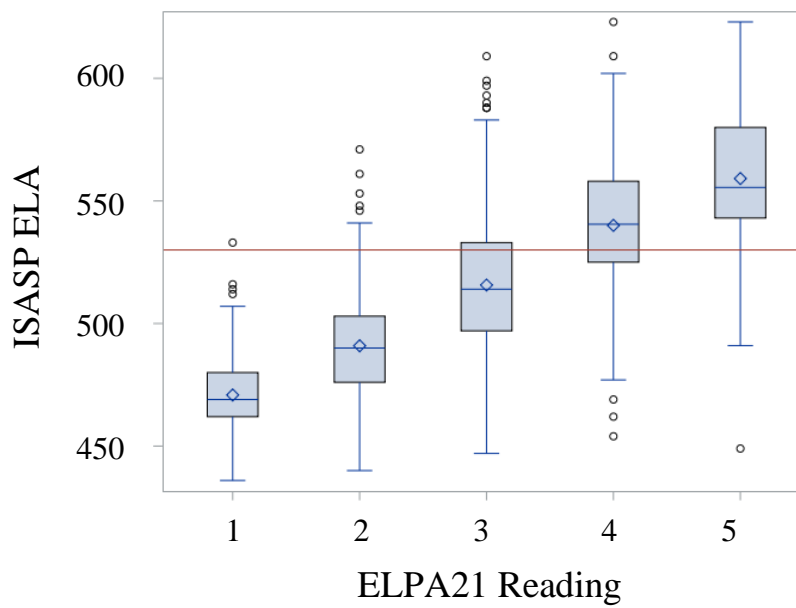
Figure A1.7

*ISASP ELA Distributions at Five Performance Levels of ELPA21 Reading in Grade 11*

Figure A2.1

ISASP Reading Distributions at Five Performance Levels of ELPA21 Reading in Grade 4

Figure A2.2

ISASP Reading Distributions at Five Performance Levels of ELPA21 Reading in Grade 5

Figure A2.3

ISASP Reading Distributions at Five Performance Levels of ELPA21 Reading in Grade 7

Figure A2.4

ISASP Reading Distributions at Five Performance Levels of ELPA21 Reading in Grade 8

Figure A2.5

ISASP Reading Distributions at Five Performance Levels of ELPA21 Reading in Grade 9

Figure A2.6

ISASP Reading Distributions at Five Performance Levels of ELPA21 Reading in Grade 10

Figure A2.7

ISASP Reading Distributions at Five Performance Levels of ELPA21 Reading in Grade 11
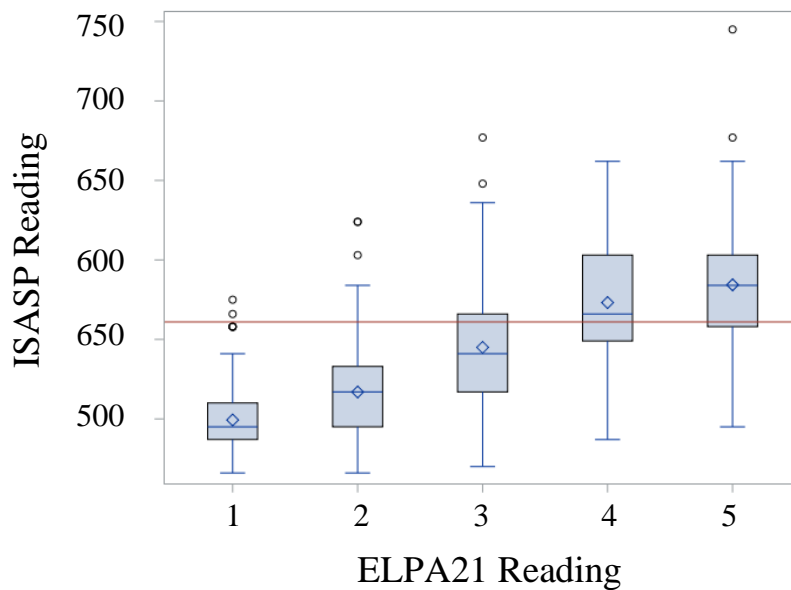
Figure A3.1

Conditional Probability of a Proficient ISASP Score in Reading given ELPA21 Reading Score in Grade 7 by TTS Use
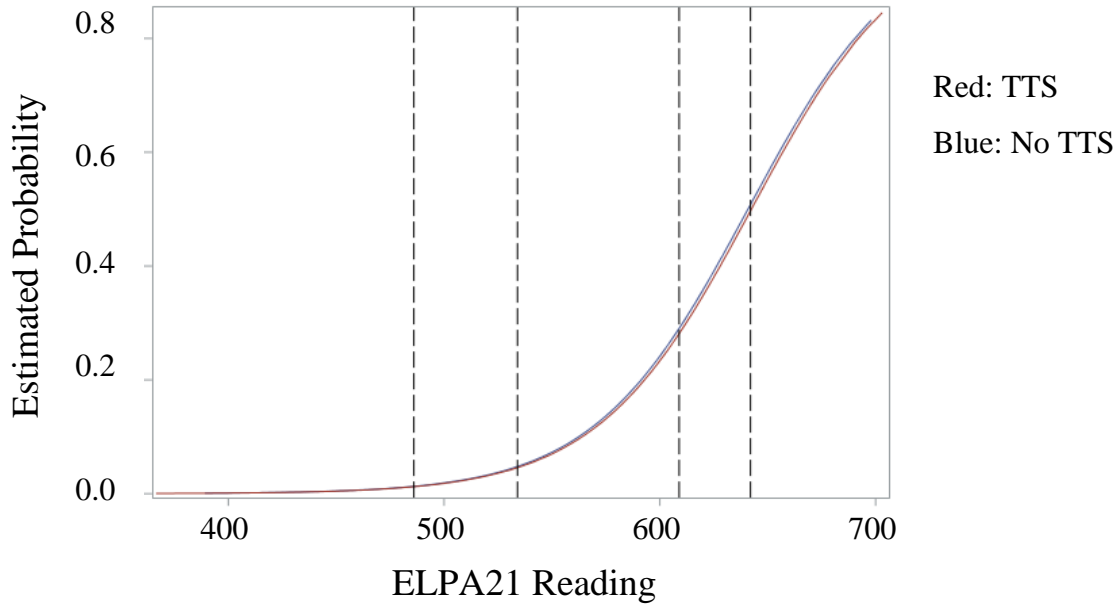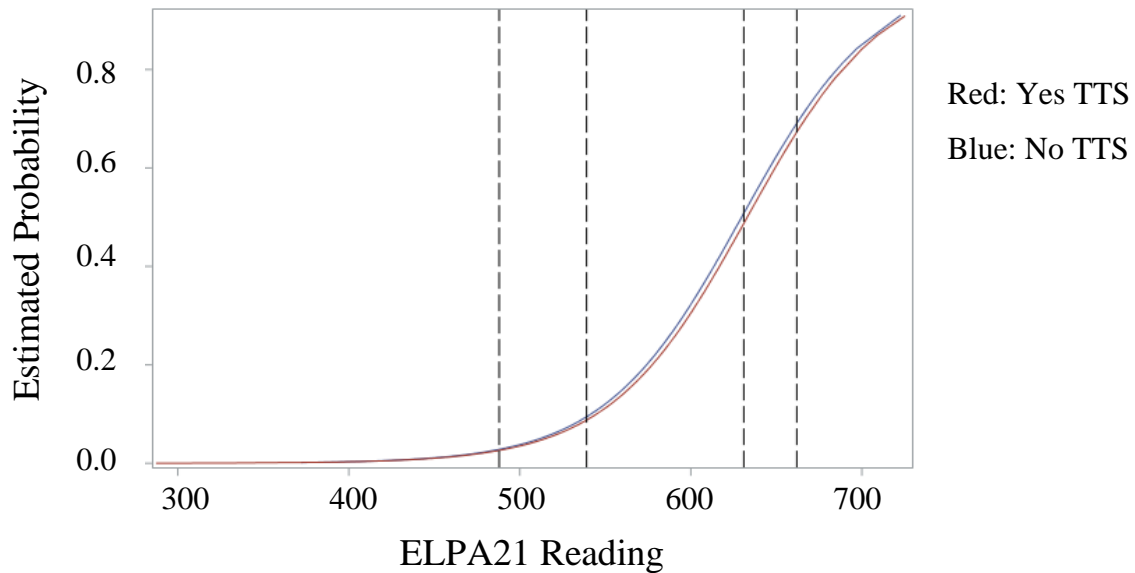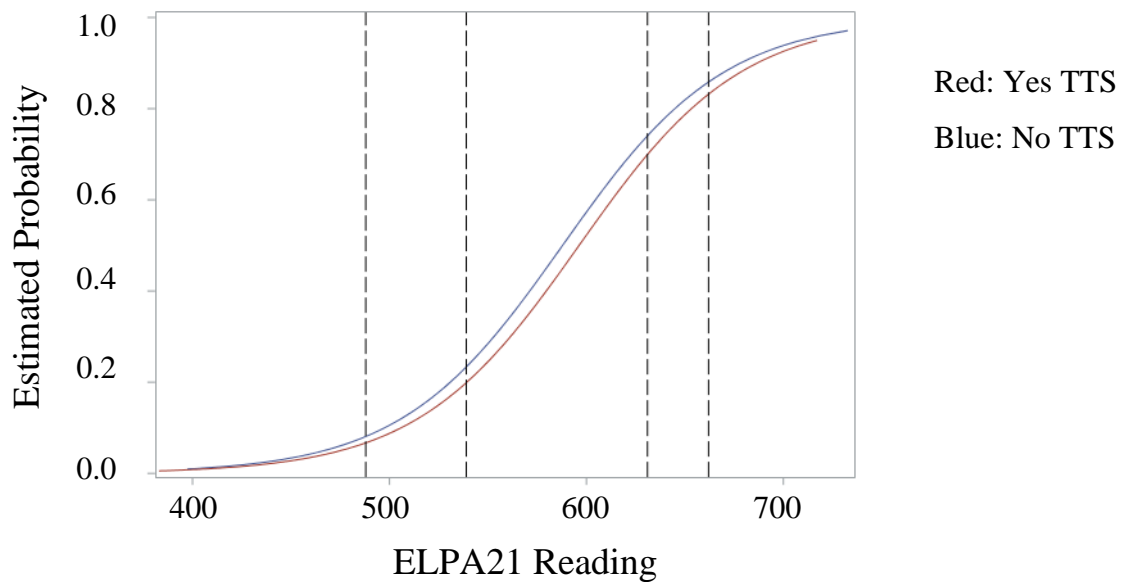

Figure A3.2

Conditional Probability of a Proficient ISASP Score in Reading given ELPA21 Reading Score in Grade 8 by TTS Use
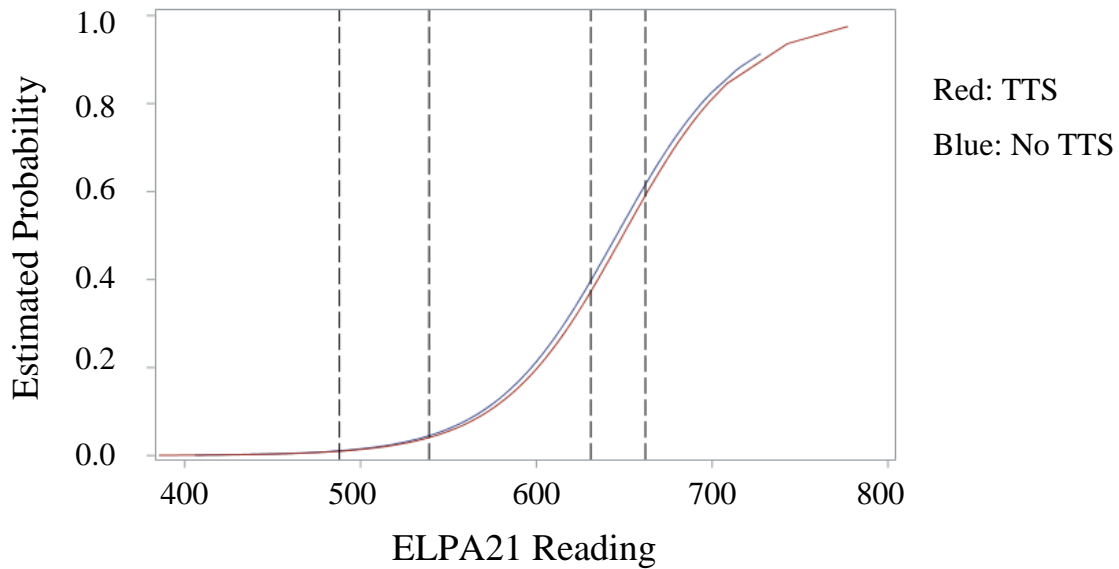
Note: Dashed lines indicate boundaries between ELPA21 achievement levels

Figure A3.3

Conditional Probability of a Proficient ISASP Score in Reading given ELPA21 Reading Score in Grade 9 by TTS Use

Note: Dashed lines indicate boundaries between ELPA21 achievement levels
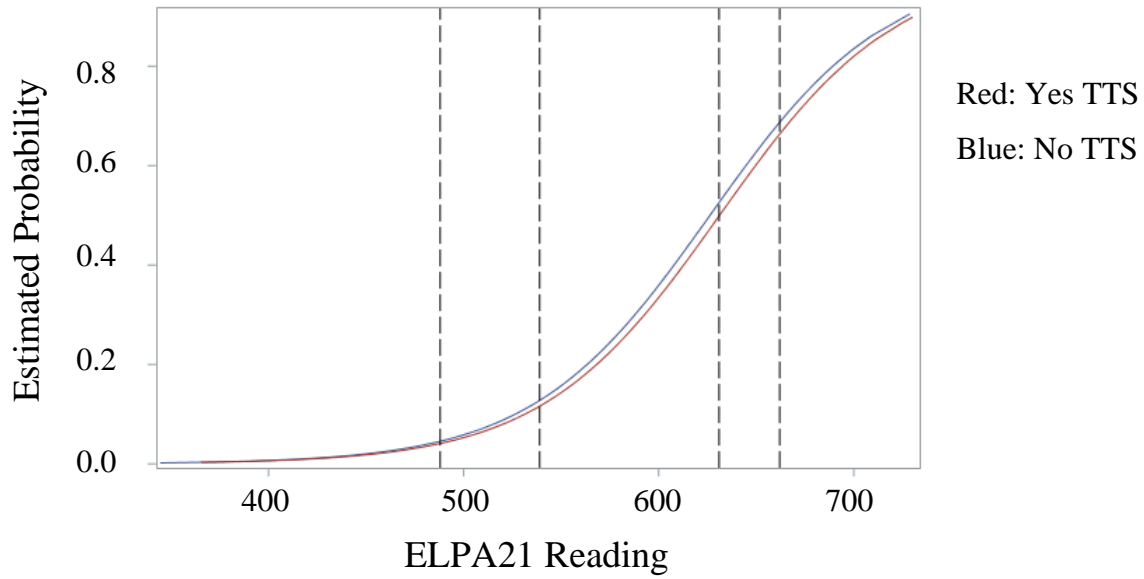
Figure A3.4

Conditional Probability of a Proficient ISASP Score in Reading given ELPA21 Reading Score in Grade 10 by TTS Use

Note: Dashed lines indicate boundaries between ELPA21 achievement levels

Figure A3.5

Conditional Probability of a Proficient ISASP Score in Reading given ELPA21 Reading Score in Grade 11 by TTS Use

Red: Yes TTS

Blue: No TTS

Note: Dashed lines indicate boundaries between ELPA21 achievement levels