



NeMTSS
FRAMEWORK



April 2020

NeMTSS Research Brief

English Language Proficiency Assessment for the 21st Century (ELPA21)

Abril Rangel-Pacheco, M.A. & Amanda Witte, Ph.D.



**NEBRASKA CENTER FOR RESEARCH ON
CHILDREN, YOUTH, FAMILIES & SCHOOLS**

English Language Proficiency Assessment for the 21st Century (ELPA21): An NeMTSS Research Brief.

Key Points:

- The ELPA21 is an online assessment of English language proficiency for K-12. The assessment measures performance in the domains of Reading, Writing, Speaking, and Listening.
- ELPA21 results are used to measure ELL students' progress in English language development, determine placement in ESL services and reclassification of ELL students, and to design professional development and formative assessments.
- Strengths of the test include gender, racially, and culturally diverse artwork, images, and names of the characters in the prompts and instructions for providing accommodations for ELL students with special needs.
- Initial reviews of the assessment yielded support for the assessment's content validity. however, there is little evidence to support other types of validity and reliability, indicating a potential weakness in the test's psychometric properties.
- Other concerns regarding the ELPA21 include issues with the practicality of using an online test in schools with a lack of technological resources and with students who lack familiarity with computer technology.

Overview

ELPA21 is an online test of English language proficiency based on the specific language skills students need to be able to interact with grade-level academic courses and become college and career ready by the 12th grade. A consortium of states (including Nebraska) developed ELPA21 to identify and measure the grade-level academic language demands corresponding to states' rigorous college and career readiness standards (Anderson, 2015). The assessment system is available for students K-12 and measures performance in the four domains of Reading, Writing, Speaking, and Listening, and produces an aggregate Proficiency Determination classification of English language student proficiency. The aggregate Proficiency Determination is based on profiles of performance across the four domains and is reported as Proficient, Progressing, or Emerging.

ELPA21 results are used to measure ELL students' progress in English language development, determine placement in ESL services and reclassification of ELL students, and to design professional development and formative assessments (The ELPA21 Framework: Summative Assessment, 2015). The ELPA21 test is generally administered in winter/spring each school year (Nebraska Department of Education, 2016, ELPA21, n.d.). Paper assessments, large-print assessments, and blind or low vision test forms are also available to provide accommodation for ELL students with special needs (UCLA CRESST, 2020).

Test Construct

The ELPA21 assessment system is aligned with the 10 ELP Standards developed by the Council of Chief State School Officers (CCSSO) in conjunction with WestEd and the Understanding Language initiative of Stanford University. The 10 ELP standards correspond to the Common Core State Standards in English language arts and mathematics as well as the Next Generation Science Standards (Huang and Flores, 2018).

According to Huang and Flores (2018), the development of the test was guided by an eclectic conceptualization of language proficiency that took into account structural linguistics and sociolinguistics frameworks of English language development. The test construct distinguishes between conversational language, classroom navigational language, and higher-order language required for accessing content area. It also reflects current thinking and beliefs about ELL education (Huang & Flores, 2018).

Assessments

ELPA21 offers both a screener for newcomer students and an annual summative assessment. Both the screener and summative assessments are available for K-12. Screener assessments are given to students entering a school system to determine if they need English language services. If the results indicate the student is non-English proficient, the levels of services they need are identified (ELPA21, n.d.). Summative assessments are administered annually. The first year of the summative assessment identifies the student's baseline and consecutive years monitor and measure their progress in English language proficiency (ELPA21, n.d.).

ELPA21 is built around a set of machine- and hand-scored task types (Anderson, 2015). Some tasks are specific to individual grade levels or domains, while others apply to all grade bands and domains. Table 1 lists subtests assorted by domain. Specific tests administered by grade band and domain are listed in the ELPA21 Framework: Summative Assessment (2015).

Table 1. ELPA Subtests by Domain

Listening	Reading	Speaking	Writing
Academic Debate	Argument & Support Essay Set	Academic Debate	Complete a Word
Academic Lecture & Discussion	Discrete Items	Analyze a Visual & a Claim	Construct a Claim
Academic Lecture or Discussion	Informational Set	Classroom Tableau	Copy a Word
Follow Instructions	Short Informational Set	Conversation	Opinion
Interactive Student Presentation	Extended Informational Set	Language Arts Presentation	Picture Caption
Listen & Match	Literary Set	Observe & Report	Respond to Peer Email
Listen for Information	Short Literary Set	Opinion	Storyboard
	Extended Literary Set	Picture Description/Compare	Write a Sentence

Long Conversation	Short Literature Set	Pictures	Write a Word
Read-Aloud Story	Extended Literature Set	Read Aloud	Writing Questions
Short Conversations	Match Picture to Word & Sentence	Show & Share	
Student Discussions	Procedural Text	Student Discussion	
Teacher Presentation	Read & Match		
Teacher Presentation: Read-Aloud	Read for Details		
	Read-Along Sentence		
	Read-Along Story		
	Short Correspondence		
	Short Correspondence Set		

Scoring

The ELPA21 assessment system measures performance across the four domains of Reading, Writing, Speaking, and Listening and in three different modalities (productive, receptive, and interactive). Domain scores are reported in terms of scale scores and achievement levels (Level 1 – Beginner, Level 2 – Early Intermediate, Level 3 – Intermediate, Level 4 – Early Advanced, and Level 5 – Advanced). The achievement levels in each of the domains are used in concert to create a “profile” and determine a Proficiency Determination. The Proficiency Determination is reported in three levels—Emerging, Progressing, and Proficient (Nebraska Department of Education, 2016).

Domain Scores

A numeric three-digit scale score describes performance on the four domains of listening, reading, speaking and writing. Each score is classified into one of five performance levels, where each level corresponds to a text descriptor stating what students in each level know and can do. The cut scores defining each level are documented in the ELPA21 Standard Setting Report and in ELP21 2016 Summative Assessment Scoring and Scaling Specifications and the corresponding descriptors (ALDs), available at www.education.ne.gov/NATLORIGIN/ELPA21.html

Achievement Profiles

A level 4 on each domain represents the English language knowledge, skills and abilities that are required to interact with and engage in grade-level content instruction at the same level as non- ELLs and is referred to as the “assessment target” for each domain. Once the assessment target is met on all non-exempt domains (e.g., a student scores “4444” (a four on each domain), ELPA21 recommends the student be eligible for reclassification. Below are examples of profiles and their classifications based on ELPA21 Proficiency rules (Nebraska Department of Education, 2016; The ELPA21 Framework: Summative Assessment, 2015).

Proficiency Determination

Once the domain scores are organized into an achievement “profile”, a Proficiency Determination is calculated based on the rules for proficiency. For example, ELPA21 indicated that a profile of 4s and 5s meets assessment targets and indicates overall proficiency, which means a child’s profile containing domain scores of 4s and 5s would receive the Proficiency Determination of “Proficient”. See Table 2 for additional profile examples and Tables 3 and 4 for descriptions of domain performance scores and Proficiency Determinations (Nebraska Department of Education, 2016; The ELPA21 Framework: Summative Assessment, 2015).

Table 2. ELPA Student Profile Examples

Profiles	Rules	Proficiency Determination
4444, 5555, 4545, 5454, 4455, 5544, 4445, 4454, 4544, 5444, 5554, 5545, 5455, 4555	A profile of 4s and 5s meets assessment targets and indicates overall proficiency	Proficient
3333, 1333, 3353, 3233, 2242, 1234, 1114, 2232	A profile with one or more domain scores above Level 2 that does not meet the requirements to be Proficient	Progressing
1122, 1212, 2222	A profile of 1s and 2s indicates an “Emerging” level of proficiency.	Emerging

Table 3. Domain Performance Descriptions

Level	Description
Level 1: Beginning	Displays few grade-level English language skills and will benefit from EL Program support.
Level 2: Early Intermediate	Presents evidence of developing grade level English language skills and will benefit from EL Program support.
Level 3: Intermediate	Applies some grade-level English language skills and will benefit from EL Program support.
Level 4: Early Advanced	English language skills required for engagement with grade-level academic content instruction.
Level 5: Advanced	Exhibits superior English language skills, as measured by ELPA21.

Table 4. Proficiency Determination Descriptions

Proficiency Determination	Description
Proficient	Students are Proficient when they attain a level of English language skill necessary to independently produce, interpret, collaborate on, and succeed in grade-level content-related academic tasks in English. This is indicated on ELPA21 by attaining a profile of Level 4 or higher in all domains. Once Proficient on ELPA21, students can be considered for reclassification.
Progressing	Students are Progressing when, with support, they approach a level of English language skill necessary to produce, interpret, and collaborate on grade-level content-related academic tasks in English. This is indicated on ELPA21 by attaining a profile with one or more domain scores above Level 2 that does not meet the requirements to be Proficient. Students scoring Progressing on ELPA21 are eligible for ongoing program support.
Emerging	Students are Emerging when they have not yet attained a level of English language skill necessary to produce, interpret, and collaborate on grade-level content-related academic tasks in English. This is indicated on ELPA21 by attaining a profile of Levels 1 and 2 in all four domains. Students scoring Emerging on ELPA21 are eligible for ongoing program support.

Score Uses

The ELPA21 assessment scores serve multiple uses. Specifically, they are intended to meet three objectives: Measuring Progress, Reclassification, and Accountability (Nebraska Department of Education, 2016; The ELPA21 Framework: Summative Assessment, 2015). Individual student scores can inform teachers and other stakeholders of what domain areas and/or language modalities ELL students need additional instruction or support in. Assessment scores can also inform ELL program eligibility decisions, provide a means to monitor English proficiency progress, determine proficiency for program exit decisions, inform teachers of instructional needs of ELs, identify resource needs, and provide evidence of program effectiveness and accountability.

Psychometrics and Technical Qualities

Reliability

Huang and Flores (2018) reported that information about the internal consistency and reliability of the test items were not available to the public, possibly because of the newness of the test. After exhaustive research, they were only able to locate a field test report online, which included extensive descriptive results about the field test items and administration but lacked psychometric information about the items. Additional research also failed to yield any information regarding the validity of the ELPA21 assessment system.

Validity

Huang and Flores (2018) reviewed technical reports, practice items, and correspondence from various state representatives and determined that the ELPA21 tests appear to have high content validity. Further, they reported that both the test development and design of ELPA21 consider current language learning theories and the items are aligned with the 10 ELP standards developed by the consortium. Regarding construct validity, Huang and Flores (2018) reported that the practice questions and prompts are overall grade-level appropriate and seem to measure language skills required for learning content area knowledge. However, they noted that a few practice items appear in multiple grade levels. According to their research, while including these items across grade levels can help track ELLs' language growth over time, they may not be developmentally appropriate for evaluating ELLs in all grade levels. Overall, due to the lack of critical information, Huang and Flores (2018) were unable to further confirm the validity of inferences made from the ELPA21 tests.

Authenticity

Authenticity refers to the degree of correspondence between the features of a particular ELP task and those of a target language task. Test authenticity can be measured by: (1) the degree of naturalness in the test language, (2) test items' degrees of contextualization, (3) the meaningfulness and relevance of the topics in the test, (4) the inclusion of some thematic organization to items, and (5) tasks that replicate real-world tasks (Huang and Flores, 2018). Huang and Flores (2018) utilized the practice items available online to determine that the ELPA 21 test items are generally authentic and grade-level appropriate. Further, they concluded that the items and prompts are presented in real-world scenarios and relate to the students' personal or schooling experiences. However, the researchers stated that although the items and prompts are overall authentic, they had some reservations about the authenticity of the online delivery mode, especially for students who have little familiarity and/or limited experience with computers (Huang and Flores, 2018).

Practicality

Huang and Flores (2018) conceptualized practicality as the "logical, down-to-earth, administrative issues involved in making, giving, and scoring an assessment instrument". They evaluated the practicality by considering multiple aspects, including the cost, the amount of administration time, ease of scoring, and ease of interpreting/reporting the results. Unfortunately, the cost of the test is not available to the public, but the amount of time it takes students to complete the test and the flexibility in test administration time appear to be practical. The test is not timed, but the estimated time for completing the four domains increases with grade level and ranges from 69 min for kindergarteners to 155 min for students in Grades 9–12. Huang and Flores (2018) noted that the benefits of practicality are contingent upon participating schools' readiness and capacity to provide the required technology and facilities, like digital devices that meet certain specifications and headphones for the Speaking tasks. Therefore, the ELPA 21 assessment system is not practical for schools without such resources.

Test Bias

Huang and Flores (2018) report that although ELPA 21 has reviewed every item and analyzed for test bias, the lack of access to a full sample test precludes a definitive conclusion about the true fairness of the test. However, the researcher did review released practice items online and based their evaluation of test bias on those practice items. Overall, they found that the artwork, images, and names of the characters in the prompts represented gender, racial, and cultural diversity. However, a potential gender bias was found in the practice questions. With the exception of a few items (e.g., the reading passage about Sally Ride for Grades 6–8), boys in the questions tended to like playing sports and were generally shown with sports gears in the images, whereas girls were less likely to be associated with sports. Huang and Flores (2018) noted that these gender differences not only reinforce gender stereotypes but could also have negative effects on ELLs' social development. Additionally, they found that some practice questions or artwork could be biased against ELLs from lower SES households. For example, they noticed that a drag-and-drop item in the Grade 1 test showed a bear with fancy accessories like flippers and sunglasses, which may be unfamiliar to an ELL student from a low SES family.

Test Fairness

ELPA21 provides accessibility and accommodations for ELLs with special needs. The ELPA21 Accessibility and Accommodations Manual (UCLA CRESST, 2020) is an extensive manual that describes the accessibility and accommodation features in the designated features of the items (e.g., assistive technology and speech-to-text) and in administrative 4 considerations (e.g., extended time and re-reading directions) (Huang and Flores, 2018). However, given that the ELPA 21 tests are meant to be administered online, ELLs' test performances may partially be contingent their familiarity with technology. For example, ELLs who have little or no familiarity with technology (e.g. computers, tablets, etc.) and young ELLs who are still developing motor skills and learning to use computers and keyboards, may have difficulties navigating through the test or completing some tasks (Huang and Flores, 2018).

Summary & Conclusions

The ELPA21 assessment system is aligned with new ELP standards that are based on decades of accumulated research on English language development. The theories and core beliefs that undergird ELPA 21 not only take into account structural linguistics and sociolinguistics frameworks of English language development but also distinguish between conversational or navigational language and higher-order language for accessing content areas (Huang and Flores, 2018). ELPA21 also recognizes the diversity in ELLs and provides accessibility and accommodations for ELLs with special needs. However, critical evidence for evaluating ELPA 21's quality is missing. For example, there is no formal documentation of the reliability or consistency in administration and scoring within and across states for the tests.

Additionally, no released sample test is available online – which makes evaluation challenging. According to Huang and Flores (2018), because ELPA 21 is a fully online assessment and ELLs vary in their familiarity with and access to technology, a full sample test could help mitigate the effects of the variation in ELLs' technological familiarity and readiness on test outcomes. In short, the psychometric information required to fully support the validity and reliability of the ELPA21 assessment system is not currently available. Based on the documents and practice items available online, ELPA21 appears to be a useful instrument for evaluating and supporting ELL students' English language development. However, the lack of published validation research makes the test's quality unknown. Future research is necessary to ensure that the ELPA 21 test is properly serving ELL students.

References

- Anderson, D. P. (2015). ELPA21 Item Development Process Report. Available at <https://cdn.education.ne.gov/wp-content/uploads/2018/06/ELPA21-Item-Dev-Process-Report-FINAL-ETS-SUBMISSION-5-15-2015.pdf>
- ELPA21 (n.d.) "Assessment System." ELPA21, www.elpa21.org/assessment-system/
- Huang, B.H. & Flores, B.B. (2018) The English Language Proficiency Assessment for the 21st Century (ELPA21), Language Assessment Quarterly, 15:4, 433-442, DOI: 10.1080/15434303.2018.1549241
- Nebraska Department of Education (2016). NE ELPA21 Scoring Interpretation Guide Grades K-12. https://cdn.education.ne.gov/wpcontent/uploads/2017/07/4NE_ELPA21_Score_Interpretation_Guide.pdf
- The ELPA21 Framework: Summative Assessment (2015). Available at <https://cdn.education.ne.gov/wp-content/uploads/2018/06/ELPA21-Assessment-Framework-Summative-2.pdf>
- UCLA CRESST (2020) Accessibility and Accommodations Manual. The Regents of the University of California. Available at https://iowaelpa21.portal.airast.org/core/fileparse.php/3650/urlt/20190604_Accessibility-and-Accommodations-Manual-SY-2019-2020.pdf

Recommended Citation:

Rangel-Pacheco, A. & Witte, A. L. (2020). *English Language Proficiency Assessment for the 21st Century (ELPA21): An NeMTSS Research Brief*. Nebraska Multi-tiered System of Support (NeMTSS).

Authorship Information:

Abril Rangel-Pacheco, M.A.

School Psychology Doctoral Student

Graduate Research Assistant

Nebraska Center for Research on Children, Youth, Families and Schools

University of Nebraska–Lincoln

arangel-pacheco2@huskers.unl.edu