

LAS Links to Texas English Language Proficiency Standards

Third-Party Independent Alignment Study Report Spring 2019

The findings in this study are those of the independent reviewing team.



Table of Contents

LAS

Executive Summary	3
LAS Links Assessment System	4
Structure of the Texas English Language Proficiency Standards	5
Alignment Study: Approach and Process	6
Overview of the Study	6
Alignment Study Participants	6
Alignment Methodology	14
Webb's Alignment Model and Modifications	14
Alignment Study Procedure	15
Alignment Criteria	19
Alignment Analysis of the Texas English Language Proficiency Standards and LAS Links Assessments	20
Alignment Results	20
Reliability among Reviewers	27
References	28
Appendix A: Linguistic Difficulty Levels	30
Appendix B: Linguistic Difficulty Level Consensus Values	37
Appendix C: Summary Tables	41
Appendix D: Results of Intra-class Correlation	99

Executive Summary

A third-party independent alignment study was conducted for the LAS Links assessments and the Texas English Language Proficiency Standards (ELPS) in Austin, Texas, from January 31 to February 1, 2019.

Eight reviewers (six Texas reviewers and two national reviewers) analyzed the LAS Links assessments at grades K through 12 for alignment to the Texas ELPS. Reviewers included English language learner curriculum experts and teachers, administrators, professors of higher education, and assessment specialists. A national alignment expert and psychometrician provided training and conducted the study. Reviewers analyzed the LAS Links assessments for each grade according to Dr. Norman Webb's alignment methodology, as modified by Dr. Gary Cook for English language proficiency assessments and standards (2007).

Data on the alignment of the LAS Links assessments were collected from the eight reviewers, following the methodology developed by Norman Webb, as modified by Dr. Gary Cook. The data collected were then statistically analyzed to determine whether each test form met the statistical criteria established by the alignment model.

Four test alignment criteria, or alignment dimensions, were examined for each LAS Links assessment. These were Categorical Concurrence, Linguistic Difficulty Level Consistency, Range-of-Knowledge Correspondence, and Balance of Representation. Three of these criteria have been defined and explained by Norman Webb in a series of publications describing his model of standards-based test alignment. Linguistic Difficulty Level Consistency has been defined by Gary Cook in his alignment studies of English language proficiency assessments and standards (2005).

Taken as a whole, the third-party independent results of the alignment of the LAS Links assessments to the Texas ELPS were very strong.

This report includes the results of the third-party independent alignment study. Given the results, the use of LAS Links as a tool of measurement for the Texas ELPS is warranted.

LAS Links Assessment System

In U.S. public schools, one in ten students receives English language services—and that number continues to grow. In Texas public schools, over 16% of students are English language learners (McFarland). Additionally, other students may also need additional language instruction to meet academic proficiency targets. LAS Links is an integrated suite of English language proficiency assessments and instructional tools designed to strengthen an English language learning program.

The LAS Links forms and placement tests correspond to the goals and objectives of states' rigorous content standards in language arts, mathematics, and science. For initial placement of students, LAS Links provides scale scores and proficiency levels. The LAS Links suite of assessments measures social and academic language in all four domains (Listening, Speaking, Reading, and Writing) for the major academic strands:

- Social, Intercultural, and Instructional
- Language Arts, Social Science, and History
- Mathematics, Science, and Technical Subjects

LAS Links uses realistic illustrations, reading passages set in context, and culturally relevant test content to engage, challenge, and encourage students throughout the learning process. LAS Links K–12 assessments are grade-level appropriate and available in five grade bands:

- Primary (K–1)
- Early Elementary (2–3)
- Elementary (4–5)
- Middle School (6–8)
- High School (9–12)

There is also a pre-K assessment available, but it was not used in this study. The LAS Links assessment forms C and D consist of multiple-choice (MC) items, short constructed-response (SCR) items, and extended constructed-response (ECR) items. The table below provides a general overview of the numbers and types of items that are contained in the LAS Links assessments. The number of items varies by grade band, as demonstrated by the use of ranges for the number of items.

	Listening	Reading	Speaking	Writing
Number of Items	20–23	30	18	17–20
Type of Items	MC	MC and SCR	SCR and ECR	MC, SCR, and ECR

Structure of the Texas English Language Proficiency Standards

The English Language Proficiency Standards (ELPS), as required by 19 Texas Administrative Code, Chapter 74, Subchapter A, §74.4, outline English language proficiency level descriptors and student expectations for English language learners (ELLs). School districts are required to implement the ELPS as an integral part of each subject in the required curriculum. The ELPS are part of the state-required curriculum, the Texas Essential Knowledge and Skills for K–12.

The ELPS in this section outline English language proficiency level descriptors and student expectations for ELLs. The cross-curricular second-language acquisition skills in subsection (c) of this section apply to ELLs in Kindegarten through Grade 12.

The Texas ELPS in subsection (c) are organized into five domains: learning strategies, listening, speaking, reading, and writing. Reviewers aligned to the skills detailed in this subsection. For the purpose of this alignment study, each skill within subsection (c) for each domain was given a code. A number was assigned for each domain, and a letter was assigned for each skill within that domain. See the range of codes for each domain below.

- Learning Strategies: 1.A–1.H
- Listening: 2.A–2.I
- Speaking: 3.A–3.J
- Reading: 4.A–4.K
- Writing: 5.A–5.G

Alignment Study: Approach and Process

Overview of the Study

The LAS Links and Texas English Language Proficiency Standards (ELPS) alignment study was conducted in Austin, Texas, from January 28 to February 1, 2019. The study involved a review of the LAS Links assessments at grades K through 12 for alignment to the Texas ELPS.

The purpose of the alignment study was to determine the degree of alignment between the ELPS and the operational test items found on the LAS Links assessments. The study was based on Webb's alignment model, a model developed by Dr. Norman Webb of the Wisconsin Center for Educational Research. Dr. Gary Cook adapted Webb's model for English language learner (ELL) assessments and standards, and his adaptations are used in this study (2005). The Webb model requires a balanced alignment-study review approach, which brings together in-state alignment and/or subject-area experts and national alignment and/or subject-area experts with the goal of ensuring that the alignment study is valid and reliable. The primary roles of the independent reviewers are to judge the linguistic difficulty level of each item and to identify the primary, and possibly secondary and tertiary, standard to which each item is aligned. Descriptions of the third-party independent reviewers for each study can be found below, in the Alignment Study Participants section of this report. A detailed description of the alignment process used with the study including summary tables showing the result, can also be found in this report. Overall, the alignment relationships between Texas ELPS and the LAS Links assessments for each study are strong and demonstrate that the items are well aligned to the respective standards.

Alignment Study Participants

Eight independent alignment experts were engaged in the study (six Texas reviewers and two national reviewers). Both the Texas reviewers and the national reviewers were individuals who had not been involved in the LAS Links item and assessment development process but who had teaching experience and/or extensive background and expertise in English language learner (ELL) curriculum. The study was planned so that four individuals on each panel were educators from school districts within Texas who have worked with ELLs. The remaining four educators were selected from various locations across the state and the country for their expertise in English language proficiency.

Information on the backgrounds and qualifications of the panelists is provided in the following section. One of the national content experts had been identified prior to the alignment study meeting to serve as the facilitator of the panel's discussions and other proceedings. The facilitator, who had served on alignment study panels in other states in the past, was briefed prior to data collection on her facilitator role in this particular study.

An independent trainer experienced in conducting alignment studies provided the initial group training and individual training as necessary throughout the alignment study proceedings. The

trainer also monitored the panel's progress throughout the five days of data collection in Austin, Texas. Dr. James Augustin served in this role. He has experience in conducting alignment studies using Webb's alignment model and served as the main facilitator of the thirdparty independent review process. Dr. Augustin is a nationally known alignment expert who has participated in a number of alignment studies as a reviewer and as a facilitator. As such, he has broad experience in conducting alignment studies using the Webb model. His role in this third-party independent alignment study was to oversee the entire alignment process, ensuring that the review was conducted correctly. Dr. Augustin was responsible for analyzing the results and providing the interpretation of the alignment results. He also provided reviewers with alignment training, including understanding Cook's linguistic difficulty levels and understanding the alignment process.

An independent auditor observed the training and the work progress of the panelists during their participation in the study, ensuring the study was not compromised in any way.

The section below provides additional information regarding the national alignment study thirdparty facilitator/trainer, the national alignment study third-party independent auditor, and the ELL third-party independent reviewers.

National Alignment Study Third-Party Facilitator/Trainer

James Augustin, PhD

Dr. James Augustin has extensive experience serving as the overall alignment process trainer and main facilitator of the third-party independent review process. He also serves to analyze the alignment data, and he provides written conclusions based upon the data. Dr. Augustin is uniquely qualified to serve in this role, having participated as a national alignment expert for several state alignment studies, including studies for Alabama, Alaska, Idaho, Louisiana, Pennsylvania, Nebraska, Maryland, and Wisconsin. For these studies, Dr. Augustin served as a trainer, lead facilitator, report writer, and/or reviewer for the alignment studies, which were based on Dr. Norman Webb's methodology. He has also consulted with the Commonwealth of Puerto Rico on alignment study procedures for the Commonwealth's testing program. Dr. Augustin has contributed numerous research, evaluation, and program-development reports on curriculum and measurement topics published by the University of Wisconsin and other education and government agencies. He served as guest editor of a special issue of *Measurement and Evaluation in Counseling and Development* (2002).

In addition to Dr. Augustin's alignment study experience, he has served as a measurement consultant, providing support for the development of a number of large-scale assessment programs. He was also most recently the director of Large-Scale Assessment with Educational Testing Service (ETS), where he was responsible for overseeing the development of multiple and complex assessments for large-scale assessment programs. Dr. Augustin was also a test development director for the Psychological Corporation/Harcourt Educational Measurement. In addition, he was with the Wisconsin Assessment Center at the University of Wisconsin and the Department of Psychology at North Carolina State University.

Dr. Augustin received his PhD degree in human resource development psychology, with advanced study in measurement and testing, from North Carolina State University at Raleigh. He received an MA degree in psychology from Marquette University and a BA degree in psychology from Trinity University in San Antonio, Texas.

National Alignment Study Third-Party Independent Auditor

Barbara Kapinus, PhD

Dr. Barbara Kapinus is a consultant in education, having recently consulted for such organizations as the Partnership for Assessment of Readiness for College and Careers (PARCC), the Smarter Balanced Assessment Consortium (SBAC), the Literacy Design Collaborative (LDC), Educational Testing Service (ETS), and the Stanford Center for Assessment, Learning, and Equity (SCALE). She has also consulted on several projects for the U.S. Department of Education, most notably the National Center for Education Statistics (NCES). In addition, she has worked on several state reading assessments, standards development projects, curriculum efforts, and staff development programs through which she has gained extensive alignment experience. She has served as a national alignment auditor and/or group facilitator for alignment studies using Webb's methodology in Alabama, Alaska, Nebraska, Pennsylvania, and Wisconsin.

Most recently, Dr. Kapinus served as the director of English Language Arts for the Smarter Balanced Assessment Consortium (SBAC). Prior to her work at the SBAC, she retired from the National Education Association, where she was a senior policy analyst for over thirteen years. Dr. Kapinus also served as the director of the Curriculum and Instructional Improvement Program at the Council of Chief State School Officers, where she worked on projects and state collaborations related to standards implementation, assessment, reading, workplace readiness, early learning, and Title I. Her experience also includes eight years as a Specialist for Reading and Communication Skills at the Maryland State Department of Education and sixteen years in Prince George's County Public Schools in several roles, including classroom teacher, reading specialist, and curriculum specialist.

Dr. Kapinus has published works on reading research, research applications, assessment, and education policy and instruction. She has served on numerous committees of the International Literacy Association, the National Assessment of Education Progress, and the National Reading Conference, including those committees responsible for alignment.

Dr. Kapinus received a BA degree in history from the University of California at Berkeley and MA and PhD degrees in reading from the University of Maryland at College Park.

English Language Learner Third-Party Independent Reviewers

Becky Huang, PhD

Dr. Becky Huang is an associate professor in the Department of Bicultural-Bilingual Studies, College of Education and Human Development, at the University of Texas San Antonio. She is also the coordinator/graduate advisor of record of the MA program in applied linguistics/teaching English as a second language.

Dr. Huang's research areas span applied linguistics, psychology, and education. Her research program focuses on two interrelated areas that address the goal of promoting language and education outcomes for bilingual/English learner (ELL) students: language/literacy development and assessment of bilingual/ELL students. The first line of research focuses on language and literacy development among bilingual/dual-language learners, particularly those who learn English as a second language. She examines the cognitive and environmental predictors of bilingual language and literacy outcomes. Her second line of research investigates the reliability, validity, and fairness of language/literacy assessments for bilingual learners. For this line of work, she focuses on the validity of placement and language proficiency tests as well as rater reliability in assessments that involve human judgments (e.g., speaking assessments and formative assessments).

Dr. Huang has published in leading applied linguistics, education, and psychology journals, including the *Teachers College Record*, *International Journal of Bilingualism*, *Studies in Second Language Acquisition*, *Reading Research Quarterly*, *Journal of Psycholinguistic Research*, *System*, *Language Testing*, and *Language Assessment Quarterly*. She has also published conference proceedings, handbook chapters, research reports, and encyclopedia entries. She recently co-edited two special issues for *System* and *International Journal of Bilingualism*. Dr. Huang holds a teaching foreign/heritage language certificate from the University of California, Los Angeles (UCLA) and an English language teacher certification in secondary education from the Ministry of Education in Taiwan. She is a former middle school teacher and has also taught English as a second/foreign language (ESL/EFL) to various age groups.

She received her doctorate in educational psychology from UCLA and completed the Linguistic Society of America's Summer Institute in Linguistics at the University of California at Berkeley. She did her postdoctoral training at Harvard University and at Educational Testing Services English Language Learning and Assessment research division.

Jennifer Edstrom

Currently, Ms. Edstrom serves as a principal educational consultant, providing instructional seminars on curriculum development, assessment, and best practices for teachers. Additionally, she serves as a field supervisor for the University of San Diego, where she observes, assesses, and instructs teaching candidates during the California state teaching certification process. Ms. Edstrom has also authored several publications in the field of education.

Ms. Edstrom's prior experience includes teaching and staff development. Her alignment experience includes work with large-scale assessment programs, for which she has overseen content development for a variety of K–8 reading/language arts state and district-implemented assessment programs in accordance with state standards and benchmarks, including materials for English language learners (ELLs). She has also participated as an English language arts national expert for alignment studies based on Dr. Webb's methodology in Pennsylvania and Nebraska. In addition to her development experience, she has trained Washington State educator committees in all aspects of assessment development, including item writing, bias and sensitivity review, ELL review, rubric writing, item review, and alignment of items to standards. Ms. Edstrom has also served as a consultant to Educational Testing Services (ETS).

Ms. Edstrom received her BS degree from Wellesley College in Wellesley, Massachusetts, and her MS degree in curriculum and instruction from the University of San Diego in San Diego, California. At the University of San Diego, she received an award of distinction from the department of learning and teaching and was a merit scholar and a graduate research fellow.

Sarah De La Garza, EdM

Sarah De La Garza is a graduate research assistant and doctoral student at the University of Texas at Austin. Prior to entering the doctoral program, she was involved in English language learning (ELL) education in various capacities. Ms. De La Garza was a high school teacher and instructional specialist at Brooks Academy of Science & Engineering in San Antonio, Texas. In this role, she coached teachers in instructional strategies, created district benchmarks, and analyzed data to revise curriculum guides. Subsequently, she served as assistant principal at Jubilee Academies in San Antonio, Texas. She also served as an English learner specialist and assistant director of bilingual/ESL programs at Jubilee Academies in Austin, Texas. In these roles, she supervised compliance and instructional initiatives related to bilingual/ESL services in San Antonio, Austin, and Kingsville regions. She also initiated the first district program evaluation and analyzed student data from classroom observations and state assessments to provide program recommendations for the Language Proficiency Assessment Committee. Ms. De La Garza holds a Texas teaching certificate in history (grades 8–12), social studies (grades 4–8), English as a second language supplemental (grades 4–12), English language arts and reading (grades 8–12), and principal (grades EC–12).

Ms. De La Garza received a BA in comparative studies in race & ethnicity with departmental honors from Stanford University. She received an EdM in school leadership from Trinity University in San Antonio, Texas, and is currently pursuing a PhD in educational leadership in policy from the University of Texas at Austin.

Stacy Reeves, PhD

Dr. Stacy Reeves is an associate professor (tenured) at the University of Southern Mississippi in Hattiesburg, Mississippi. In this position, she teaches undergraduate and graduate classroom-



management courses and literacy courses for initial and advanced teacher licensure. Dr. Reeves is chair and co-chair of undergraduate honor students' theses, specialist students' action research field projects, and doctoral candidates' dissertations in literacy and related areas. She is also a professional-development trainer for area schools and other groups that work with students. Dr. Reeves serves on a variety of committees and other university-based decision-making groups, and she is on the board of a multinational school in Limuru, Kenya, providing support for the assessment of students in literacy and English language acquisition and making suggestions to teachers, parents, and other parties for students' growth in literacy and English language acquisition.

Before this position, Dr. Reeves taught for many years, including at William Carey University in Hattiesburg, Mississippi, and at the University of Southern Mississippi. She was also an elementary school teacher for Hattiesburg Public Schools. Dr. Reeves is a member of the Mississippi Reading Association, having served as state president from 2010 to 2011, and is an active member of the International Reading Association. She has served as an English language arts national alignment study expert for state assessment alignment studies based on Dr. Norman Webb's methodology in Alabama, Alaska, Nebraska, Pennsylvania, and Wisconsin.

Dr. Reeves received a BS degree in elementary education and an MS degree in education, with an emphasis on reading/literacy, from the University of Southern Mississippi. She received a PhD degree in curriculum and instruction, with a minor in technology, from Mississippi State University.

Rebecca Callahan, PhD

Dr. Rebecca Callahan is an associate professor of education and faculty research associate at the Population Research Center at the University of Texas at Austin. Dr. Callahan has published numerous books, peer-reviewed articles, book chapters, encyclopedia entries, book reviews, reports, manuscripts, and other publications. Her most recent book is titled *The Bilingual Advantage: Language, Literacy, and the U.S. Labor Market*. She received a Reviewer Award for the *American Educational Research Journal* in 2014 and an Early Career Award in Bilingual Education Special Interest Group in 2011, both from the American Educational Research Association.

Dr. Callahan has also presented at numerous conferences and invited speaker sessions. Most recently, she was an invited speaker at the TITLE I English Learner State Collaborative on Assessment and Student Standards at the Chief Council of State School Officers Conference in San Diego, California, in June 2018. She also recently presented "School Structure and the Dynamics of Power: Postsecondary Opportunities for Bilingual English Learner Students" at the American Educational Research Association Annual Conference in New York, New York, in April 2018. Dr. Callahan was formerly a bilingual teacher for ten years in the Jamul-Dulzura Union, Woodland Joint Unified, and Grant Joint Union School Districts.

Dr. Callahan received a BA in anthropology and a Bilingual Certificate of Competence (BCC) bilingual teaching certificate from the University of California at San Diego and an MA in

education from the University of California at Davis. She received her PhD of Education in language and literacy: second language acquisition from the University of California at Davis.

Virginia Acevedo

Virginia Acevedo has over fifteen years of experience in bilingual and English as a second language (ESL) instruction. As a bilingual teacher, Ms. Acevedo planned and delivered lessons to English language learning (ELL) students and created tools and support used for informal and formal assessments to evaluate students, including for the Texas English Language Proficiency Assessment System (TELPAS). As the district initial bilingual placement and TELPAS testing coordinator, she was responsible for all district testing materials, procedures, and support. As the current district bilingual and ESL instructional facilitator, Ms. Acevedo oversees initial placement testing for all ELLs in the Socorro ISD. She also facilitates TELPAS training, testing, and data collection for the district, in addition to evaluating and providing documentation to the program director regarding progress of bilingual/ESL students, teachers, and campuses. Additionally, Ms. Acevedo is currently enrolled in a Principal Internship at Lamar University and is pursuing a Principal Certification.

Ms. Acevedo received a BA in interdisciplinary studies with a minor in elementary bilingual/ESL education from the University of Texas at El Paso. She received her MA in teacher leadership from the University of Phoenix.

Blanca Florencia

Blanca Florencia is an experienced professional, is fluent in both English and Spanish, and has worked in public education for over forty years. Ms. Florencia was an instructor for thirty-one years in the Houston ISD and Minneapolis Public Schools, teaching English as a second language and physical education. She also served as a program coordinator, coordinating a language grant that successfully improved academic scores for targeted Spanish-speaking students and attending the National Association for Bilingual Education conferences. Ms. Florencia also served as a program facilitator of the English language learning (ELL) Department at Minneapolis Public Schools. In this role, she was responsible for overseeing the development and implementation of English as a second language (ESL) programs at seven elementary schools. Most recently, she worked as an Organizer for Texas American Federation of Teachers for Region 5 AMP.

Ms. Florencia received a BS in physical education/health K–12 from the University of Texas– Pan American. She received a Developmental and Adapted Physical Education endorsement from Minnesota State University, Mankato. She received an MA in curriculum and instruction from St. Thomas University and ESL licensure from Hamline University.



Leo Izaguirre

Leo Izaguirre has worked in public education for over thirty-four years, holding teaching and leadership positions. Ms. Izaguirre taught in the McAllen ISD and Pharr-San Juan-Alamo ISD for twenty years in elementary school, junior high school, and high school settings. She also served as the bilingual/English as a second language (ESL) director for fourteen years. In this role, she was responsible for overseeing the implementation of the district's ESL services, including instruction, placement, and testing. She also provided support to teachers of English language learners (ELLs) in bilingual and ELL settings. This included providing ESL resources for teachers at all grade levels and leading professional development. Ms. Izaguirre has served on various state and district ELL committees, including the Language Proficiency Assessment Committee, Region One Bilingual Task Force Committee, and Texas Education Agency Bilingual Task Committee.

Ms. Izaguirre received a BS in education with a minor in reading and an MA in curriculum and instruction, both from the University of Texas–Pan American.

Alignment Methodology

As stated earlier in this report, the alignment study that was conducted for LAS Links and the Texas English Language Proficiency Standards (ELPS) was based on the work of Dr. Norman Webb, Wisconsin Center for Educational Research, University of Wisconsin–Madison. In his work, Webb states that the alignment of the standards or objectives for student learning with tests for measuring students' attainment of these expectations is an essential component for an effective standards-based education system. The alignment study was designed to model Webb's procedures, including the use of depth-of-knowledge levels and Webb's definition of alignment (Webb, 2005, 2007). The definition of alignment is as follows:

Alignment is defined as the degree to which expectations and assessments are in agreement and serve in conjunction with one another to guide the system toward students learning what they are expected to know and do. As such, alignment is a quality of the relationship between expectations and assessments and not a specific attribute of either of these two system components. Alignment describes the match between expectations and assessment that can be legitimately improved by changing either student expectations or assessments. Seen as a relationship between two or more system components, alignment can be determined by using the multiple criteria described in detail in a National Institute of Science Education (NISE) research monograph, *Criteria for Alignment of Expectations and Assessments* (Webb, 1997).

The Webb methodology was modified by Cook (2007) for use with English language proficiency standards and assessments, and this study was conducted using his modification.

Webb's Alignment Model and Modifications

Webb's alignment model as modified by Cook (2007) is based on four criteria: Linguistic Difficulty Level (LDL) Consistency, Categorical Concurrence, Range-of-Knowledge Correspondence, and Balance of Representation. Reviewers used these four criteria to assess the content agreement between the Texas ELPS and LAS Links. For each alignment criterion, an acceptable level was defined by what would be required to ensure that a student had met the standards. A brief description of the alignment criteria is provided below. Additional information can be found in the section of this report labeled "Alignment Criteria."

Linguistic Difficulty Level Consistency is a metric representing the percentage of items at the linguistic difficulty level (LDL). This measure is a replacement for Webb's depth of knowledge statistic. In this instance, LDL refers to linguistic instead of cognitive complexity. Each language proficiency standard is given an LDL of 1, 2, or 3. Level 1 stands for elementary linguistic features, level 2 represents standard linguistic constructions, and level 3 refers to complex linguistic formulations. During the alignment process, LDLs are also assigned to each test item. The purpose is to identify the connection between standards' LDLs and test items' LDLs. If an item is coded above or below a standard's LDL, little information would be available about how

that item samples students' behavior relative to the linguistic difficulty of that standard. The main idea of assessing ELL students is to help identify and monitor linguistic progress. If items do not match the linguistic levels of the standards, the ability to properly evaluate students is limited.

Categorical Concurrence refers to how well items match or cover standards. To evaluate this area, the statistic Categorical Concurrence is used. Categorical Concurrence is calculated by averaging the number of items assigned to specific English language proficiency standards by raters, or educators who participate in the alignment process. Raters select specific standards for each item on the tests being rated. The number of coded items is then averaged across raters and reported as Categorical Concurrence. This statistic is a proxy for the average number of items raters believe address specific standards. It is important to note that some items can address more than one standard, and raters are allowed to code accordingly.

Range-of-Knowledge Correspondence refers to how well a test's items cover a set of standards. Webb states that "this criterion is met if a comparable span of knowledge expected of students by a standard is the same as, or corresponds to, the span of knowledge that students need in order to correctly answer the assessment items/activities" (Webb, 2007).

Balance of Representation is the degree to which one standard or group of standards is given more emphasis on a test than another standard or group of standards. An index (Webb, 2002) is used to judge the distribution of the test items.

The Webb model has been used extensively in many alignment studies throughout the country and has been recommended for use by the Council of Chief State School Officers (CCSSO). The alignment criteria in the Webb model also adhere to the guidelines specified in the United States Department of Education's Standards and Assessments peer review documents, including the "Annotated Assessment Peer Review Guidance" provided on November 20, 2015.

Cook's modifications to Webb's model were published by CCSSO in the *Aligning Assessment to Guide the Learning of All Students* publication (Cook, 2005). The fifth chapter, "Aligning English Language Proficiency Tests to English Language Learning Standards," by H. Gary Cook, extends alignment research to an investigation of English language proficiency (ELP) test alignment (Cook, 2005). The report describes a process for ELP alignment, which matches an assessment's linguistic skills and acquisition levels to English language development standards.

Alignment Study Procedure

The panelists met in Austin for a period of five days during the last week of January 2019. Norman Webb's (1999) definition and model of alignment were followed in design of the data collection for the study. Aspects of Gary Cook's (2005) adaptation of Webb's model to English language proficiency (ELP) tests were also incorporated into the data collection process. At the beginning of the meeting, all panelists were provided with an orientation and training presentation. The presentation covered Webb's definition of alignment and his conceptualization of alignment criteria. A detailed discussion of LDLs, as defined by Cook (2005), was included. The alignment process all panelists would follow was also introduced. Further detailed instruction was provided to the panelists as they progressed through the five steps of the process. Individual panelists were given instruction on aspects of specific steps when this was requested or judged necessary by the trainer. The group facilitator also assisted with some instruction and guidance. Characteristics of the Texas English Language Proficiency Standards (ELPS) required some special instructions during the data collection phase of the study.

Task 1 focused on reaching consensus on the LDL of each Texas content standard/indicator after individual panelists had made initial LDL assignments. All panelists participated in the discussions so that the judgments of both Texas educators and national content experts could be considered in reaching a consensus on the LDL of each standard. The panel facilitator made sure that there was adequate discussion whenever needed before designating the consensus LDLs of the standards. The trainer and independent auditor noted that discussions were thoughtful and balanced, and no one panelist dominated the discussions.

Detailed instruction was provided by the trainer and facilitator at the beginning of each subsequent task (i.e., tasks 2 through 5). Panelists worked independently on tasks 2 through 5, which focused on making a series of judgments regarding alignment at the individual test-item and test-form levels. Panelists worked at a pace comfortable to them, and all panelists had sufficient time to thoroughly examine alignment of all five grade-level tests assigned to them. Panelists worked up the grades from level K–1 to level 9–12.

A high-level overview of the steps in the process is provided on the next page. The Texas ELPS, along with the LDL consensus values, can be found in Appendix B of this report. The alignment study process also involved the electronic capture of data. Information about the electronic data capture tool and its use in the process is provided below.

The Electronic Data Capture Tool

The electronic data capture tool was used in the third-party alignment study. The tool was designed specifically to facilitate the gathering of independent reviewers' judgments. For the LAS Links to Texas ELPS third-party alignment study, the application automated the process of aligning the Texas ELPS for a given content area and the test items found on the corresponding LAS Links assessment. The tool and its reports made it possible to gauge in a timely manner the alignment, based on Webb's alignment model, between the Texas ELPS and the items on the assessments. In addition, the tool also provided opportunities for reviewers to provide additional information regarding items, including comments related to source of challenge. The item-by-objective or standard codes by reviewers were then aggregated and analyzed.

As stated, the national alignment expert Dr. James Augustin provided training on the overall alignment process and LDLs and also served as the lead facilitator. Dr. Augustin has extensive experience training third-party independent review committee members in the use of electronic data capture software for alignment studies. The training provided information not only on understanding the LDLs but also on properly using the electronic data capture tool when

assigning an LDL to each standard and item. A high-level overview of the process is provided on the next page.

LAS Links

Alignment Study Process

Step 1: Determining the Linguistic Difficulty Level (LDL)

Reviewers individually determined the LDL for each skill in subsection (c) of the Texas English Language Proficiency Standards (ELPS). They discussed their LDL ratings to reach a group consensus.

Step 2: Taking the test

Reviewers reviewed the test and recorded their comments about the test items.

Step 3: Determining what each test item measured and the LDL for each test item

Step 3.1

Using the first three test items, reviewers independently determined what each item measured by assigning it to a primary standard (and a secondary standard and tertiary standard, if applicable). A group discussion took place; however, reaching consensus on what each item measured was not required.

Step 3.2

Reviewers independently determined the LDLs of the first three items. Reviewers were instructed to code only one LDL (1, 2, or 3) for each of the three items. Reviewers also independently noted any source of challenge for the first three items. A group discussion took place; however, reaching consensus on the LDLs of the first three items was not required.

Step 3.3

Reviewers continued to independently determine the primary standard (and the secondary standard and tertiary standard, if applicable) for the remainder of the test items.

Step 3.4

Reviewers independently determined the LDL for the remainder of the test items. Again, the reviewers were instructed to code only one LDL for each test item.

Throughout the alignment process, reviewers independently noted any source of challenge for each test item and provided written comments as necessary.

Step 4: Summarizing the alignment criteria of test items

Once reviewers determined the primary standard (and secondary standard and tertiary standard, if applicable) for each test item and the LDL for each test item, they analyzed the entire test for LDL Consistency, Categorical Concurrence, Range-of-Knowledge Correspondence, and Balance of Representation.

Step 5: Answering a debriefing questionnaire

Reviewers independently shared feedback about the process, the test items, and the standards.

Alignment Criteria

Judgments of the panelists were statistically analyzed according to Webb's model of alignment and Cook's specific application of the model to English language learning standards. All the model's statistical alignment criteria were applied, and the results were reviewed along with the panelists' written responses to a debriefing questionnaire (task 5) completed for each grade level test form, in preparation of this report.

In most instances, Cook's (2007) alignment criteria were applied to the resulting statistical analyses of the panel's judgments to describe the degree of alignment for each of the four alignment dimensions. Two changes were made that actually increased the rigor of the criteria; these changes were made in recognition of the characteristics of the LAS test series. First, Cook's (2007) statistical criteria for the Writing subtests were strengthened to match the statistical criteria for the alignment levels used for the other three subtests addressing Speaking, Listening, and Reading. This was done because the Writing subtest at each LAS Links test level has a sufficient number of test items to do so, in contrast to Cook's criteria established for a Writing test with only two writing prompts. Second, the statistical criteria required to attain the three levels of K–1 Range-of-Knowledge Correspondence were increased to match those used for the other four LAS Links test levels. Making these statistical criteria more rigorous seemed appropriate due to there being similar numbers of items on the LAS Links K–1 subtests as appear on the subtests of the other LAS Links levels. The statistical criteria used in this study to identify degree of alignment for all four subtests are presented in Table 2.

	Alignment Criteria				
Alignment Areas	Speaking, Listening, Reading, Writing				
	LIMITED	MODERATE	STRONG		
Categorical Concurrence	<4	≥4	≥6		
LDL Consistency	<40%	≥40%	≥50%		
Range-of-Knowledge	<40%	≥40%	≥50%		
Correspondence					
Balance of Representation	<0.6	≥0.6	≥0.7		

Table 2: Alignment Criteria for Speaking, Listening, Reading, Writing Domains

The results for each of the four criteria discussed in this section were calculated using Cook's adaptations to Webb's methodology, reviewers' averaged ratings, and reviewers' comments. The results for LDL consistency, categorical concurrence, range-of-knowledge correspondence, and balance of representation are included in Appendix C of this report.

Alignment Analysis of the Texas English Language Proficiency Standards and LAS Links Assessments

The overall results of the alignment study indicate that all five levels of the LAS Links tests are very strongly aligned with the Texas content standards for Speaking, Listening, Reading, and Writing within the state's English Language Proficiency Standards (ELPS). For the Categorical Concurrence and Range-of-Knowledge Correspondence dimensions of alignment, alignments were found to be consistently strong across all four content areas for all five levels (K–1, 2–3, 4– 5, 6–8, and 9–12) of LAS Links. Alignment of the tests along the Linguistic Difficulty Level (LDL) Consistency and Balance of Representation dimensions was similarly strong, with only a few exceptions where alignment with some content standards was found to be moderate.

Before discussing the details of the alignment of each specific LAS Links test level with the Texas ELPS, it is important to note that the LDLs assigned (through consensus discussions by the panel) to the standards within each content domain showed fewer LDL 1 standards and more LDL 2 and LDL 3 standards. This resulted in the LDL means of the content domains reported in Table 3 following a similar pattern and ranging from 2.1 for Writing to 2.3 for Speaking.

Table 3 below shows the results for this first major step in the alignment process involving reviewers' determinations of the LDLs of the Texas ELPS. Additional information regarding the various LDLs can be found in Appendix B of this report.

Domain	LDL 1 Standards	LDL 2 Standards	LDL 3 Standards	Domain Mean
Speaking	1	5	4	2.3
Listening	2	3	4	2.2
Reading	2	5	4	2.2
Writing	1	4	2	2.1
Learning Strategies	0	6	2	2.3

Table 3: Distribution and Means of Consensus Linguistic Difficulty Levels for the TexasEnglish Language Proficiency Standards

Alignment Results

Using the electronic data capture tool, reviewers independently determined what each item measured. They also entered the linguistic difficulty level (LDL) for each item. The calculation software provided the statistical analysis to determine whether each LAS Links assessment as a whole included items measuring content from each of the Texas ELPS domains in order to

evaluate categorical concurrence. The tool also provided the statistical analysis to determine LDL consistency, range-of-knowledge correspondence, and balance of representation.

A high-level summary alignment analysis for LDL consistency, categorical concurrence, rangeof-knowledge correspondence, and balance of representation is provided in Tables 4–8. The results indicate that the alignment relationship between the Texas ELPS and the LAS Links assessment is very strong, as noted in the "Interpretation of Content Alignment Results" section of this report. Additional detailed information is provided in Appendix C of this report.

The following is a summary of the degree of alignment found at each of the five levels of the LAS Links tests for the four alignment dimensions.

LAS Links Grades K–1 Alignment with Texas English Language Proficiency Content Standards						
Content Standards	Categorical Concurrence	LDL Consistency	Range of Knowledge	Balance of Representation		
Speaking	Strong	Moderate	Strong	Moderate		
Listening	Strong	Strong	Strong	Strong		
Reading	Strong	Strong	Strong	Strong		
Writing	Strong	Strong	Strong	Strong		

Table 4: Grades K–1 Content Alignment Summary

Table 5: Grades 2–5 Content Alignment Summary	Table 5:	Grades 2–3	Content Alignment Summary	7
---	----------	------------	----------------------------------	---

LAS Links Grades 2–3 Alignment with Texas English Language Proficiency Content Standards						
Content Standards	Categorical Concurrence	LDL Consistency	Range of Knowledge	Balance of Representation		
Speaking	Strong	Moderate	Strong	Strong		
Listening	Strong	Strong	Strong	Strong		
Reading	Strong	Strong	Strong	Moderate		
Writing	Strong	Strong	Strong	Strong		

LAS Links Grades 4–5 Alignment with Texas English Language Proficiency Content Standards						
Content Standards	Categorical Concurrence	LDL Consistency	Range of Knowledge	Balance of Representation		
Speaking	Strong	Moderate	Strong	Strong		
Listening	Strong	Strong	Strong	Strong		
Reading	Strong	Strong	Strong	Moderate		
Writing	Strong	Strong	Strong	Strong		

Table 6: Grades 4–5 Content Alignment Summary

 Table 7: Grades 6–8 Content Alignment Summary

LAS Links Grades 6–8 Alignment with Texas English Language Proficiency Content Standards						
Content Standards	Categorical Concurrence	LDL Consistency	Range of Knowledge	Balance of Representation		
Speaking	Strong	Moderate	Strong	Strong		
Listening	Strong	Strong	Strong	Strong		
Reading	Strong	Strong	Strong	Strong		
Writing	Strong	Moderate	Strong	Strong		

Table 8: Grades 9–12 Content Alignment Summary

LAS Links Grades 9–12 Alignment with Texas English Language Proficiency Content Standards						
Content Standards	Categorical Concurrence	LDL Consistency	Range of Knowledge	Balance of Representation		
Speaking	Strong	Moderate	Strong	Moderate		
Listening	Strong	Strong	Strong	Strong		
Reading	Strong	Strong	Strong	Strong		
Writing	Strong	Strong	Strong	Strong		

Interpretation of Content Standards Alignment Results

Level K-1

At level K–1, Categorical Concurrence and Range-of-Knowledge Correspondence were strong for all four subtests: Speaking, Listening, Reading, and Writing. This indicates a sufficient number of items measuring the standards within each content area and good coverage of the collection of Texas standards within each English language proficiency subtest or domain.



LDL Consistency and Balance of Representation also were strong for the Listening, Reading, and Writing subtests at the K–1 level. For the Speaking subtest, alignment for these two dimensions was judged by the panel to be moderate—perhaps, as one panelist noted in her debriefing comments, due to the challenge of measuring conversational speaking as part of a written standardized test battery. Overall, the written comments of the panel were positive concerning coverage of the standards at appropriate LDLs, including a wide range of item difficulties, and the potential of the test to measure a broad range of students' English proficiency from low to high.

Level 2–3

Similar to level K–1, Categorical Concurrence and Range-of-Knowledge Correspondence were strong for all four subtests. LDL Consistency was strong for Listening, Reading, and Writing. LDL Consistency was moderate for Speaking— again, probably as a result of the challenge of measuring conversational speaking with a written standardized test.

Balance of Representation was strong for Speaking, Listening, and Writing. Balance of Representation was moderate for the Reading subtest. One panelist noted that there was "a lack of critical reading analysis in this assessment but that is not entirely inappropriate based on where 2–3 graders lie on the developmental continuum for language." Another panelist confirmed this observation with the comment that analytic skills within Reading were not as well represented in the Reading questions as were basic and inferential comprehension skills.

In their overall comments, panelists stated that the assessment covered the full range of content specified in the standards, with a full range of cognitive complexity and difficulty reflected in the test items.

Level 4–5

As was the case with the previous two levels, Categorical Concurrence and Range-of-Knowledge Correspondence alignment were strong for all four content domains. Some judges wrote comments elaborating that the test covered the range of standards appropriate to a written test for level 4–5, while some standards were most appropriately assessed at earlier grades or using performance-based assessment tools.

LDL Consistency was strong for Listening, Reading, and Writing content standards. For the Speaking standards, LDL Consistency was moderate. In their written comments, several panelists observed that the test items as a collection were more cognitively complex and linguistically difficult than the items in the levels K–1 and 2–3 tests.

Balance of Representation was strong for the Speaking, Listening, and Writing subtests. For the Reading subtest, Balance of Representation alignment was moderate.

Level 6-8

For the middle school grades (6-8) test, Categorical Concurrence, Range-of-Knowledge Correspondence, and Balance of Representation were strong for all four content domains. Written comments of the panelists tended to specify that the test items did a sufficient job covering the range and emphasis of content domain standards appropriate to middle school (i.e., grades 6–8) students, recognizing that some standards seemed not to apply to this level.

Panelists judged LDL Consistency as strong for the Listening and Reading content domains. For the Speaking and Writing subtests, LDL Consistency was moderate.

Level 9-12

Categorical Concurrence and Range-of-Knowledge Correspondence were strong across all four content domains for the high school level test. As with the lower levels of LAS Links, panelists noted in their debriefing comments that the test sufficiently covered the standards appropriate to English language learners (ELLs) in grades 9–12. It was observed by some panelists that standards that focused on foundational language skills for younger or beginner ELLs were slightly less covered.

LDL Consistency and Balance of Representation were strong for Listening, Reading, and Writing. LDL Consistency and Balance of Representation alignment were moderate for the Speaking subtest. Although the Speaking subtest was not specifically referenced in regard to complexity and difficulty of the items, panelists did notice that most items at the high school level were LDL 2 or LDL 3.

Regarding Balance of Representation, two panelists specifically observed in their debriefing comments that the subtests at the high school level did seem to focus on three or four standards, though most grade-appropriate standards were represented by at least one item. However, they did not specifically cite the Speaking subtest in making this observation.

LAS Links Level	Domain	Categorical Concurrence	LDL Consistency	Range of Knowledge	Balance of Representation
Grades K–1	Speaking	S	S	М	S
	Listening	S	S	L	S
	Reading	S	М	L	S
	Writing	S	М	L	S
Grades 2–3	Speaking	S	М	S	S
	Listening	S	S	М	S
	Reading	S	М	М	S
	Writing	S	S	М	S
Grades 4–5	Speaking	S	М	S	S
	Listening	S	М	М	S
	Reading	S	М	М	S
	Writing	S	S	М	S
			·		·
Grades 6–8	Speaking	S	М	S	S
	Listening	S	S	М	S
	Reading	S	S	S	S
	Writing	S	М	М	S
	·		· ·	·	
Grades 9–12	Speaking	S	L	S	S
	Listening	S	S	М	S
	Reading	S	S	М	S
	Writing	S	М	М	S

 Table 9: LAS Links Alignment with Texas English Language

 Proficiency Learning Strategies

Note: Alignment with Learning Strategies is as follows: S = Strong; M = Moderate; L = Limited.

Interpretation of Learning Strategies Alignment Results

During task 3 of the data collection process, panelists were encouraged to identify a Texas Learning Strategy standard measured (if one in particular could be identified) in addition to the primary content domain standard measured by each test item. The results of these panel judgments are summarized for all four domains across all five test levels in Table 9. What is apparent in examining the 80 alignment degree information cells in the table (4 domains X 4 alignment dimensions X 5 test levels) is that there is a pattern of fairly solid alignment of the subtests with the Texas Learning Strategies. A total of 58 (72.5%) alignments are strong, 18 (22.5%) are moderate, and only 4 (5.0%) are limited. The strong alignments include Categorical Concurrence across all subtests at all five grade levels. Similarly, Balance of Representation alignment with the Learning Strategies standards is strong across all subtests at all five grade levels.

Considering LDL Consistency and Range-of-Knowledge Correspondence, most alignments with the Learning Strategies standards were strong or moderate in degree. Range-of-Knowledge Correspondence for the Learning Strategies standards was limited for the Listening, Reading, and Writing subtests within the level K–1 test form. At the high school level (i.e., grades 9–12), LDL Consistency alignment with the Learning Strategies standards was limited for the Speaking subtest items.

The overall strength of the LAS Links tests' alignment with the Learning Strategies standards is noteworthy because panelists commented on the challenge of making these specific alignment judgments. Some panelists questioned the appropriateness of examining the alignment of written standardized test items with the <u>activities</u> represented in the Learning Strategies standards. Upon completing her alignment review of this last level in the series of five LAS Links test forms, one panelist wrote several perceptive comments about her cumulative experience working with the Learning Strategy standards. She noted that she aligned test content to the Learning Strategies standards in a more sparing way as she progressed up the levels of LAS Links. She stated that the Learning Strategies can be aligned tangentially to the assessment but, in her professional opinion, are not easily measured on a paper-and-pencil standardized test, but could be measured by a performance test. This general sentiment was echoed by other panelists during the course of the four days of data collection for the LAS Links alignment study.

Reliability among Reviewers

The intra-class correlation is based on the mean squares from the analysis of variance of a twoway random effects model, reviewers crossed with items (Shroud and Fleiss, 1979), as described in Appendix D. The overall intra-class correlation among the reviewers' assignment of linguistic difficulty levels to items was reasonably high. If there is a low variance among the reviewers' coding in assigning linguistic difficulty levels to items, the intra-class correlation has greater error. Table 10 provides a summary of the intra-class correlation and the percentage of items coded as the same linguistic difficulty level by all reviewers.

Test Form	Intra-Class Correlation
Listening Level K-1	0.84
Listening Level 2-3	0.66
Listening Level 4-5	0.73
Listening Level 6-8	0.69
Listening Level 9-12	0.76
Reading Level K-1	0.91
Reading Level 2-3	0.71
Reading Level 4-5	0.60
Reading Level 6-8	0.68
Reading Level 9-12	0.80
Speaking Level K-1	0.78
Speaking Level 2-3	0.83
Speaking Level 4-5	0.83
Speaking Level 6-8	0.71
Speaking Level 9-12	0.65
Writing Level K-1	0.87
Writing Level 2-3	0.67
Writing Level 4-5	0.86
Writing Level 6-8	0.61
Writing Level 9-12	0.76

Table 10: Summary or Reliability among Reviewers

References

§74.4. English Language Proficiency Standards. (n.d.). Retrieved January 07, 2019, from http://ritter.tea.state.tx.us/rules/tac/chapter074/ch074a.html#74.4

Annotated Assessment Peer Review Guidance. (2015). The Center for Assessment for the Council of Chief State School Officers based on the guidance released by the U.S. Department of Education on September 25, 2015.

Board of Regents of the University of Wisconsin System. (2012). WIDA's 2012 Amplification of the English Language Development Standards, Kindergarten–Grade 12.

Cook, H. G. (2005). Aligning English Language Proficiency Tests to English Language Learning Standards. *State Collaborative on Assessment and Student Standards: Council of Chief State School Officers* (pp. 131–54).

Cook, H. G. (2007). Alignment study report: The WIDA Consortium's English language proficiency standards for English language learners in kindergarten through grade 12 to ACCESS for ELLs assessment. Madison, WI: University of Wisconsin.

McFarland, J., Hussar, B., Wang, X., Zhang, J., Wang, K., Rathbun, A., Barmer, A., Forrest Cataldi, E., and Bullock Mann, F. (2018). *The Condition of Education 2018* (NCES 2018-144). U.S. Department of Education. Washington, DC: National Center for Education Statistics. Retrieved March 1, 2019 from https://nces.ed.gov/pubsearch/pubsinfo.asp?pubid=2018144.

Shrout, P. E., & Fleiss, J. L. (1979). Intra-class correlations: Uses in assessing rater reliability. *Psychological Bulletin*, *86*(2), 420–428.

Subkoviak, M. J. (1988). A practitioner's guide to computation and interpretation of reliability indices for mastery tests. *Journal of Educational Measurement*, 25(1), 47–55.

Webb, N. L. (1997). *Criteria for alignment of expectations and assessments in mathematics and science education* (Research Monograph No. 6). Washington, DC: Council of Chief State School Officers.

Webb, N. L. (1999). Alignment of science and mathematics standards and assessments in four states (Research Monograph No. 18). Washington, DC: Council of Chief State School Officers.

Webb, N. L. (2002). Alignment study in language arts, mathematics, science, and social studies of state standards and assessments for four states. Washington, DC: Council of Chief State School Officers.

Webb, N. L. (2005, November). *Depth-of-knowledge levels for four content areas*. Paper presented at the 50th annual meeting of the Florida Education Research Association, Miami, FL.

Webb, N. L. (2007). Issues related to judging the alignment of curriculum standards and assessments. *Applied Measurement in Education*, 20(1), 7–25.

Wisconsin Center for Education Products and Services [WCEPS]. (Producer). (2014, July 1). *Dr. Norman Webb's DOK overview* [Video file]. Retrieved from https://www.youtube.com/watch? v=qFXU6_TYIjc

LAS



Appendix A: Linguistic Difficulty Levels

Linguistic Difficulty Levels by Skill Area

Skill Areas	Areas		Linguistic Difficulty Levels	
		Level 1: Elementary Features	Level 2: Standard Constructions	Level 3: Complex Formulations
Oral/Aural	Listening	Limited to a basic understanding of simple	Basic understanding of and attending to every	Understanding of and attending to more
		formulaic, classroom and social discourse;	day classroom and social discourse, common	complex or specialized grade appropriate
		attending to day-to-day brief and common	idiomatic expressions both in the classroom	discourse and interactions, comprehending
		language interactions	and in social situations and contexts	contextualized and acculturated
				communications (e.g., ellipsis, comedy,
				parody)
	Speaking	Limited to a basic ability to produce	Facility to produce standard classroom and	Facility to produce and interact within
)	formulaic expressions in standardized	social discourse interactions using extended	complex classroom and social discourse
		classroom and social situations	formulaic expressions as well as common	interactions utilizing more contextualized and
			idiomatic expressions	acculturated forms and constructions
Textual	Reading	Little to a basic facility to process and	Basic understanding of and ability to attend to	Understanding of and attending to grade
)	attend to English phonemic and alphabetic	standard everyday grade appropriate texts	appropriate vocabulary and texts; grade-level
		constructions; little to a basic ability to	which include vocabulary and passages most	ability to comprehend classroom and socially
		comprehend high frequency grade	commonly encountered in classroom and every	appropriate texts
		appropriate classroom and survival	day social situations	
		vocabulary items		
	Writing	Limited to a basic ability to copy and/or	Basic ability to produce simple, grade relevant	Facility to produce grade appropriate text
)	produce simple text constructions (e.g.,	classroom-based and/or social-text-utilizing	constructions using grade appropriate
		letters, basic vocabulary items, name)	standard vocabulary and grammatical features	vocabulary and grammatical features and
			and constructions	constructions; ability to produce and express
				grade appropriate ideas and concepts

LAS

Materials reproduced from Cook, H.G. (2006). Aligning English Language Proficiency Tests to English Language Learning Standards. In *Aligning Assessment to Guide the Learning of All Students*. CCSSO: TILSA SCASS. Washington, D.C.

31

ENGLISH LANGUAGE LEARNER ALIGNMENT EXAMPLES OF ASSIGNING LINGUISTIC DIFFICULTY

Level 1 Elementary Features

Example 1:1 Listening

Objective: Follow simple two-step oral directions to complete a task in English

This objective is an example of level 1. Students most commonly meet this expectation by using elementary linguistic features—an elementary understanding of direction would be all the student would need to know to meet this objective.

Example 1:2 Speaking

Objective: Use common social greetings and simple repetitive phrases

This objective is a level 1 since only simple constructions and formulaics are needed to meet this objective.

Example 1:3 Reading

Objective: Recognize some common English morphemes in simple phrases or sentences.

This objective requires students to identify elementary text-based features and represents a linguistic difficulty level of one.

Example 1:4 Writing

Objective: Student will be able to write basic personal information (name, address, phone number)

This objective highlights basic writing expectations and is a level 1 linguistic difficulty.

Example 1:5 Writing

Item: Prompt

Choose the correct word. Triangles always ______ three sides.

- A. have
- B. half
- C. has

This item is a level 1 since it requires a student to use elementary vocabulary and syntactic features.

Level 2 Standard Constructions

Example 2:1 Listening

Objective: Participate in routine classroom discussions.

This objective requires students to move beyond simple directions or responding to common formulaic expressions. It requires students to listen to more sophisticated, albeit routine classroom interactions and is thus a level 2.

Example 2:2 Speaking

Objective: Deliver simple narrative and informative presentations and express with simple, detailed sentences.

Here students must go beyond simple formulaics. This objective requires students to engage in simple by dynamic interactions representing standard classroom interactions and discourse and hence is a level 2.

Example 2:3 Reading

Objective: The student will read, comprehend, and analyze fiction and nonfiction. Answer simple, factual questions about what is read.

Students are required to process more sophisticated everyday texts and understand and respond to simple factual questions; hence this is a level 2.

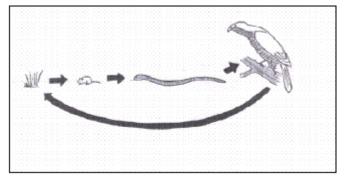
Example 2:4 Writing

Objective: Write a short narrative story that includes the basic elements of setting and characters and that follows a visually supported outline provided by the teacher.

This objective asks students to write short, basic, grade-relevant narratives making this a level 2.

Example 2:5 Listening

Item: Prompt



SCRIPT

Narrator: Listen to a teacher talking to her science class

Teacher: When we study the food chain, we might want to think of a circle. There is not a real beginning or end—it just continues to go round and round. It is a system or circle of soil, plants, animals, and then all of those things dying and becoming "soil" again. So, think about it in concrete terms. You have soil, which provides food for plants to grow. Then you have plants. They are the food of many animals. The animals eat them, and eventually the animals die. The organic material in the animals' bodies then becomes part of the soil, which feeds new plants, which feed new animals. The circle continues around and around. These are the workings of an ecosystem – plants and animals feed each other.

Now take this model of an ecosystem that I've just described for you and work with a partner to describe specific plants and animals that live together exactly the way I've described.

Narrator: What does the teacher compare the food chain to?

- A. A circle
- B. Death
- C. A house
- D. Plants

This item requires students to process the passage and then determine how "circle" fits into the discussion. The passage is relatively long and involved, but the question above taps into standard listening expectations.

Level 3 Complex Formulations

Example 3:1 Listening

Objective: Evaluate use of materials or resources needed to complete tasks based on oral discourse

To meet this objective, students must be involved in complex academic and social interactions and negotiations. Task specific vocabulary and discourse strategies are need and hence this is a level 3.

Example 3:2 Speaking

Objective: In a variety of academic and social contexts, ask for or provide specific information that confirms or denies beliefs

Here students much have complex mastery of a variety of discourse features. Successful mastery of this object exhibits a level 3 linguistic difficulty.

Example 3:3 Reading

Objective: The student will use strategies to read a variety of materials, fiction and nonfiction. Make connections between previous knowledge and/or experiences and what is read.

Students are required to not only process a variety of reading materials, but they must also connect currently gained knowledge with previous experience or knowledge; thus, this is a level 3.

Example 3:4 Writing

Objective: Write clear and coherent grade appropriate paragraphs with effective transitions and sentence structures.

For students to meaningfully exhibit this objective, they must have grade appropriate fluency in writing. This characteristic exhibits a level 3.

Example 3:4 Reading

Item: Prompt



Read the instructions for making a book cover that the teacher gave to her students: We have finished reading Bridge to Terabithia by Katherine Paterson. Now, you are going to prepare a book jacket for the book. Follow these instructions for designing your book jacket:

- 1. The title and author are placed on the front cover. Don't forget to follow the rules for capitalization of both title and author.
- 2. Put an illustration on the front cover. The illustration should reflect the characters, plot, or setting of the book.
- 3. On the back cover, include at least three positive reviews.

Which chart is made correctly?

- A. Graphic A
- B. Graphic B
- C. Graphic C
- D. Graphic D

This task is a level 3 since students have several contextualized reading tasks, e.g., understand the notion of book jackets, process the three requested tasks and interpret each instruction relative to the tasks context.

Reproduced from Cook, H. G. (2005). Aligning English Language Proficiency Tests to English Language Learning Standards. *State Collaborative on Assessment and Student Standards: Council of Chief State School Officers* (pp. 131–54).



Appendix B: Linguistic Difficulty Level Consensus Values

Texas English Language Proficiency Standards LDL Consensus

Standard	Consensus
Learning Strategies	
1.A Use prior knowledge and experiences to understand meanings in English	2
1.B Monitor oral and written language production and employ self-corrective techniques or other resources	3
1.C Use strategic learning techniques such as concept mapping, drawing, memorizing, comparing, contrasting, and reviewing to acquire basic and grade-level vocabulary	2
1.D Speak using learning strategies such as requesting assistance, employing non-verbal cues, and using synonyms and circumlocution (conveying ideas by defining or describing when exact English words are not known)	2
1.E Internalize new basic and academic language by using and reusing it in meaningful ways in speaking and writing activities that build concept and language attainment	2
1.F Use accessible language and learn new and essential language in the process	2
1.G Demonstrate an increasing ability to distinguish between formal and informal English and an increasing knowledge of when to use each one commensurate with grade-level learning	2
expectations 1.H Develop and expand repertoire of learning strategies such as reasoning inductively or	2
deductively, looking for patterns in language, and analyzing sayings and expressions	
commensurate with grade-level learning expectations	3
Listening	1
2.A Distinguish sounds and intonation patterns of English with increasing ease	1
2.B Recognize elements of the English sound system in newly acquired vocabulary such as long and short vowels, silent letters, and consonant clusters	1
2.C Learn new language structures, expressions, and basic and academic vocabulary heard during classroom instruction and interactions	2
2.D Monitor understanding of spoken language during classroom instruction and interactions and seek clarification as needed	2
2.E Use visual, contextual, and linguistic support to enhance and confirm understanding of increasingly complex and elaborated spoken language	2
2.F Listen to and derive meaning from a variety of media such as audio tape, video, DVD, and CD ROM to build and reinforce concept and language attainment	3
2.G Understand the general meaning, main points, and important details of spoken language ranging from situations in which topics, language, and contexts are familiar to unfamiliar	3
2.H Understand implicit ideas and information in increasingly complex spoken language commensurate with grade-level learning expectations	3
2.I Demonstrate listening comprehension of increasingly complex spoken English by following directions, retelling or summarizing spoken messages, responding to questions and requests, collaborating with peers, and taking notes commensurate with content and grade-level need	3

Speaking	
Speaking 3.A Practice producing sounds of newly acquired vocabulary such as long and short vowels, silent	
letters, and consonant clusters to pronounce English words in a manner that is increasingly	
comprehensible	1
3.B Expand and internalize initial English vocabulary by learning and using high-frequency	
English words necessary for identifying and describing people, places, and objects, by retelling	
simple stories and basic information represented or supported by pictures, and by learning and	
using routine language needed for classroom communication	2
3.C Speak using a variety of grammatical structures, sentence lengths, sentence types, and	
connecting words with increasing accuracy and ease as more English is acquired	3
3.D Speak using grade-level content area vocabulary in context to internalize new English words	
and build academic language proficiency	2
3.E Share information in cooperative learning interactions	2
3.F Ask and give information ranging from using a very limited bank of high-frequency, high-	
need, concrete vocabulary, including key words and expressions needed for basic communication	
in academic and social contexts, to using abstract and content-based vocabulary during extended	
speaking assignments	2
3.G Express opinions, ideas, and feelings ranging from communicating single words and short	
phrases to participating in extended discussions on a variety of social and grade-appropriate	2
academic topics 3.H Narrate, describe, and explain with increasing specificity and detail as more English is	3
acquired	2
3.I Adapt spoken language appropriately for formal and informal purposes	3
	5
3.J Respond orally to information presented in a wide variety of print, electronic, audio, and visual	2
media to build and reinforce concept and language attainment	3
Reading 4.A Learn relationships between sounds and letters of the English language and decode (sound out)	
words using a combination of skills such as recognizing sound-letter relationships and identifying	
cognates, affixes, roots, and base words	1
4.B Recognize directionality of English reading such as left to right and top to bottom	1
	1
4.C Develop basic sight vocabulary, derive meaning of environmental print, and comprehend	2
English vocabulary and language structures used routinely in written classroom materials	Z
4.D Use prereading supports such as graphic organizers, illustrations, and pretaught topic-related	2
vocabulary and other prereading activities to enhance comprehension of written text	2
4.E Read linguistically accommodated content area material with a decreasing need for linguistic	
accommodations as more English is learned	2
4.F Use visual and contextual support and support from peers and teachers to read grade-	
appropriate content area text, enhance and confirm understanding, and develop vocabulary, grasp	
of language structures, and background knowledge needed to comprehend increasingly challenging language	2
4.G Demonstrate comprehension of increasingly complex English by participating in shared	<u> </u>
reading, retelling or summarizing material, responding to questions, and taking notes	
commensurate with content area and grade level needs	3
4.H Read silently with increasing ease and comprehension for longer periods	2

4.I Demonstrate English comprehension and expand reading skills by employing basic reading	
skills such as demonstrating understanding of supporting ideas and details in text and graphic	
sources, summarizing text, and distinguishing main ideas from details commensurate with content	2
area needs	3
4.J Demonstrate English comprehension and expand reading skills by employing inferential skills	
such as predicting, making connections between ideas, drawing inferences and conclusions from	
text and graphic sources, and finding supporting text evidence commensurate with content area needs	3
4.K Demonstrate English comprehension and expand reading skills by employing analytical skills	3
such as evaluating written information and performing critical analyses commensurate with	
content area and grade-level needs	3
Writing	5
5	
5.A Learn relationships between sounds and letters of the English language to represent sounds when writing in English	1
5.B Write using newly acquired basic vocabulary and content-based grade-level vocabulary	2
	۷
5.C Spell familiar English words with increasing accuracy, and employ English spelling patterns	
and rules with increasing accuracy as more English is acquired	2
5.D Edit writing for standard grammar and usage, including subject-verb agreement, pronoun	
agreement, and appropriate verb tenses commensurate with grade-level expectations as more	
English is acquired	2
5 E Employ in an annual su complex company at a later structures in contant and writing commencements	
5.E Employ increasingly complex grammatical structures in content area writing commensurate	
with grade-level expectations, such as:	
(i) using correct verbs, tenses, and pronouns/antecedents(ii) using possessive case (apostrophe s) correctly	
(ii) using negatives and contractions correctly	3
	5
5.F Write using a variety of grade-appropriate sentence lengths, patterns, and connecting words to	_
combine phrases, clauses, and sentences in increasingly accurate ways as more English is acquired	2
5.G Narrate, describe, and explain with increasing specificity and detail to fulfill content area	
writing needs as more English is acquired	3



Listening Grades K-1

Table 1C1L: Summary of Alignment Results

			Standa	urd Alignment o	Standard Alignment of Grades K-1 Listening Ite	tening Items		
		Align	Alignment Statistics	CS		Align	Alignment Findings	
Standards	CAT	LDL	CAT LDL Range	Balance	CAT	LDL	Range	Balance
LRN STRT	8.88	56%	22%	0.88	STRONG	STRONG	LIMITED	STRONG
LISTEN	22.38 50%	50%	54%	0.71	STRONG	STRONG	STRONG	STRONG

Table 1C2L: Summary of LDL Consistency

			Hits	The		Percent of Questions at LDL Level	s at LDL Lev		
				Under	ler	At		Ab	Above
	Standards	Mean	Standard Deviation	Mean Percent Below	Standard Deviation	Mean Percent At	Standard Deviation	Mean Percent Above	Standard Deviation
LRN									
STRT	8	8.88	8.87	13%	0.19	0.56	0.10	0.31	0.21
LISTEN	9	22.38	5.60	46%	0.22	0.50	0.19	0.04	0.09

Listening Grades K-1

Concurrence
Categorica
÷
0
Summary o
••
3
C3L
\leq
5
Table

	Standards	LDL Level	No. of Standards	Percentage of Standards	Coded Mean	Standard Deviation	Categorical Concurrence
		1	0	0%0			
STRT	8	2	6	75%	8.88	8.87	STRONG
		3	2	25%			
		1	2	22%			
LISTEN	6	2	3	33%	22.38	5.60	STRONG
		3	4	44%			

Table 1C4L: Summary of Range-of-Knowledge Correspondence

	Derrar	kange of Knowledge	LIMITED	STRONG
	otal	Standard Deviation	0.24	0.13
tandards	% of Total	Avg. Percent of Standards Hit	22%	54%
Range of Standards	dards Hit	Standard Deviation	1.91	1.13
	No. of Standards Hit	Mean Standards Hits	1.75	4.88
Hits		Standard Deviation	8.87	5.60
		Mean	8.88	22.38
		Standards	8	6
			LRN STRT	LISTEN

LAS

Listening Grades K-1

Table 1C5L: Summary of Balance of Representation

		Percent of Total Hits	Total Hits		Index	
	Standards	Mean	Standard Deviation	Mean	Standard Deviation	Balance of Representation
LRN STRT	8	13.84%	8.87	0.88	0.18	STRONG
LISTEN	6	34.89%	5.60	0.71	0.10	STRONG

Listening Grades 2-3

Table 2C1L: Summary of Alignment Results

			St	andard Alignr	tandard Alignment of Grades 2-3 Listening Items	2-3 Listening It	ems	
		Alignn	dignment Statisti	ics			Alignment Findings	
Standards	\mathbf{CAT}	CAT LDL	Range	Balance	CAT	LDL	Range	Balance
LRN STRT	11.50	55%	42%	0.86	STRONG	STRONG	MODERATE	STRONG
LISTEN	26.25	26.25 55%	57%	0.70	STRONG	STRONG	STRONG	STRONG

Table 2C2L: Summary of LDL Consistency

		Hits			Percent of Questions at LDL Level	ons at LDL I			
			Un	Under	At		Ab	Above	
standards Mean Standard I Deviation P	Standard Deviation	~ ~ 	Mean Percent Below	Standard Deviation	Mean Standard Percent At Deviation	Standard Deviation	Mean Percent Above	Standard Deviation	Consistency
8 11.50 8.98	8.98		19%	0.36	0.55	0.32	26%	0.23	STRONG
9 26.25 6.54	6.54		39%	0.23	0.55	0.16	6%	0.11	STRONG

LAS

Table 2C3L: Summary of Categorical Concurrence

Categorical Concurrence		STRONG			STRONG	
		S			S	
Standard Deviation		8.98			6.54	
Coded Mean		11.50			26.25	
Percentage of Standards	0%0	75%	25%	22%	33%	% 77
No. of Standards	0 6			2	3	4
LDL Level	1	2	3	1	2	3
Standards		8			6	
		STRT			LISTEN	

Listening Grades 2-3

Table 2C4L: Summary of Range-of-Knowledge Correspondence

			Hits	Ray No. of Standards Hit	Range of St		andards	Standards % of Total
	Standards	Mean	Standard Deviation	Mean Standards Hits		Standard Deviation	dard Avg. Pe ation of Standa	dard Avg. Pe ation of Standa
∞		.50	8.98	3.38	2.26		42%	
o ∞	1	11.50	8.98	5 12	2.26		42% 57%	
	9	26.25	6.54	5.13	0.99		57%	57% 0.11

Table 2C5L: Summary of Balance of Representation

		Percent of	Percent of Total Hits	Index	X	
	Standards	Mean	Standard Deviation	Mean	Standard Deviation	Balance of Representation
LRN STRT	8	14.98%	8.98	0.86	0.16	STRONG
LISTEN	6	34.20%	6.54	0.70	0.08	STRONG

Listening Grades 4-5

Table 4C1L: Summary of Alignment Results

			Sti	andard Alignn	nent of Grades 4	tandard Alignment of Grades 4-5 Listening Items		
		Alignn	Alignment Statistics	cs		Alignr	Alignment Findings	
Standards	CAT LDL	LDL	Range	Balance	CAT	LDL	Range	Balance
LRN STRT	14.25	14.25 43%	48%	0.81	STRONG	MODERATE	MODERATE	STRONG
LISTEN	26.50	59%	56%	0.75	STRONG	STRONG	STRONG	STRONG
~								

Table 4C2L: Summary of LDL Consistency

Image: Figure formStandardsMeanUnderAtAtAboveLDLStandardsMeanStandardMeanStandardPercentDeviationMeanStandardConsistencyImage: Figure formDeviationPercentDeviationMeanStandardDeviationDeviationConsistencyImage: Figure formDeviationPercentDeviationMeanStandardDeviationDeviationImage: Figure formFigure formDeviationMeanStandardDeviationDeviationDeviationImage: Figure formFigure formFigure formDeviationMeanStandardDeviationImage: Figure formFigure formFigure formDeviationMeanStandardMeanImage: Figure formFigure formFigure formDeviationMeanMeanMeanImage: Figure formFigure formDeviationDeviationMeanMeanMeanImage: Figure formFigure formDeviationDeviationMeanMeanMeanImage: Figure formFigure formDeviationDeviationMeanMeanMeanImage: Figure formFigure formDeviationDeviationMeanMeanMeanImage: Figure formFigure formDeviationDeviationDeviationMeanMeanImage: Figure formFigure formDeviationDeviationDeviationMeanMeanImage: Figure formFigure f			I	Hits		Pe	rcent of Que	Percent of Questions at LDL Level	evel		
StandardsHeanStandardMeanStandardMeanStandardPercentStandardPercentStandardPercentStandardStandardBelowBelowBelowAtDeviationAtDeviationAboveDeviation814.257.5918%0.340.430.2939%0.31926.507.6332%0.200.590.169%0.07					Un	ıder		At	Ab	0Ve	
8 14.25 7.59 18% 0.34 0.43 0.29 39% 0.31 9 26.50 7.63 32% 0.20 0.59 0.16 9% 0.07		Standards	Mean	Standard Deviation	Mean Percent Below	Standard Deviation	Mean Percent At	Standard Deviation	Mean Percent Above	Standard Deviation	Consistency
8 14.25 7.59 18% 0.34 0.43 0.29 39% 0.31 9 26.50 7.63 32% 0.20 0.59 0.16 9% 0.07	LRN										
9 26.50 7.63 32% 0.20 0.59 0.16 9% 0.07	STRT	8	14.25	7.59	18%	0.34	0.43	0.29	39%	0.31	MODERATE
	LISTEN	6	26.50		32%	0.20	0.59	0.16	9%6	0.07	STRONG

Categoi Ica oniniary or j I ante

	Standards	LDL Level	No. of Standards	Percentage of Standards	Coded Mean	Standard Deviation	Categorical Concurrence	
		1	0	0%0				
LKN STRT	8	2	9	75%	14.25	7.59	STRONG	
		3	2	25%				
		1	2	22%				
LISTEN	6	2	3	33%	26.50	7.63	STRONG	
		3	4	44%				

Listening Grades 4-5

Table 4C4L: Summary of Range-of-Knowledge Correspondence

LISTEN	LRN STRT	70		
9	8	Standards	-	
26.50	14.25	Mean		
7.63	7.59	Standard Deviation	11163	Hite
5	3.88	Mean Standards Hits	No. of Standards Hit	R
0.76	1.64	Standard Deviation	ds Hit	Range of Standards
56%	48%	Avg. Percent of Standards Hit	% of Total	ndards
0.08	0.21	Standard Deviation	Total	
STRONG	MODERATE	Range of Knowledge		

Table 4C5L: Summary of Balance of Representation

		Percent of	Percent of Total Hits		Index	
	Standards	Mean	Standard Deviation	Mean	Standard Deviation	Balance of Representation
LRN STRT	8	17.40%	7.59	0.81	0.12	STRONG
LISTEN	6	32.37%	7.63	0.75	0.08	STRONG

Listening Grades 6-8

Table 6C1L: Summary of Alignment Results

			Stal	ndard Alignme	nt of Grades 6-8	Standard Alignment of Grades 6-8 Listening Items		
		Alignn	lignment Statistics	cs		AI	dignment Findings	
Standards	CAT	CAT LDL	Range	Balance	CAT	LDL	Range	Balance
LRN STRT	17.25	69%	44%	0.72	STRONG	STRONG	MODERATE	STRONG
LISTEN	31.63	61%	57%	0.72	STRONG	STRONG	STRONG	STRONG

Table 6C2L: Summary of LDL Consistency

Hits	Hits	Hits		Perc	Percent of Questions at LDL Level	s at LDL Leve			
			U	Under	A	t	Above	ove	
Standards Mean Standard Deviation	, ,	Standard Deviation	 Mean Percent Below	Standard Deviation	Mean Percent At	Standard Deviation	Mean Percent Above	Standard Deviation	Consistency
8 17.25 7.98		7.98	7%	0.08	0.69	0.20	0.23	0.18	STRONG
9 31.63 9.55		9.55	32%	0.19	0.61	0.18	0.07	0.05	STRONG

AS

Table 6C3L: Summary of Categorical Concurrence

Standards		LKN 8 STRT 8	IW		JISTEN 9	
		2	3	1	2	"
No. of Standards	0	6	2	2	3	~
Percentage of Standards	%0	75%	25%	22%	33%	707 V
Coded Mean		17.25			31.63	
Standard Deviation	PC 414 HOI	7.98			9.55	
Categorical Concurrence		STRONG			STRONG	

					Range o	Range of Standards		
				No. of Sta	No. of Standards Hit	% of Total	ĩ	
	Standards	Mean	Standard Deviation	Mean Standards Hits	Standard Deviation	Avg. Percent of Standards Hit	Standard Deviation	Range of Knowledge
LRN STRT	8	17.25	7.98	3.50	1.85	44%	0.23	MODERATE
LISTEN	6	31.63	9.55	5.13	0.83	57%	0.09	STRONG

Table 6C4L: Summary of Range-of-Knowledge Correspondence

LAS

		Percent of Total Hits	Fotal Hits		Index	
	Standards	Mean	Standard Deviation	Mean	Standard Deviation	Balance or Representation
LRN						
STRT	8	17.51%	7.98	0.72	0.16	STRONG
LISTEN	9	32.11%	9.55	0.72	0.11	STRONG

Table 9C1L: Summary of Alignment Results

			Stan	idard Alignmen	Standard Alignment of Grades 9-12 Listening Ite	2 Listening Items	18	
		Alignr	Alignment Statistics	ics		Ali	lignment Findings	
Standards	CAT	LDL	CAT LDL Range	Balance	CAT	LDL	Range	Balance
LRN STRT	17.38 55%	55%	47%	0.75	STRONG	STRONG	MODERATE	STRONG
LISTEN	32.00 55%	55%	58%	0.71	STRONG	STRONG	STRONG	STRONG
Listening Grades 6-8	8-9							

Table 9C2L: Summary of LDL Consistency

L

		Under
Mean Percent At	рц	Standard Deviation
0.55		0.16
0.55		0.20

LAS

Listening Grades 9-12

Table 9C3L: Summary of Categorical Concurrence

		LISTEN			STRT			
		9			8		Standards	
	3	2	1	3	2	1	Level	LDL
-	4	3	2	2	6	0	Standards	No. of
-	44%	33%	22%	25%	75%	0%	Standards	Percentage of
		32.00	<u>I</u>		17.38	<u>I</u>	Mean	Coded
-		9.96			8.86		Deviation	Standard
		STRONG			STRONG		Concurrence	Categorical

Table 9C4L: Summary of Range-of-Knowledge Correspondence

LAS

LISTEN	LRN STRT			
9	8	Standards	1	
32.00	17.38	Mean		
9.96	8.86	Standard Deviation		Hite
5.25	3.75	Mean Standards Hits	No. of Standards Hit	
1.49	1.67	Standard Deviation	ards Hit	Range of
58%	47%	Avg. Percent of Standards Hit	% of Total	Range of Standards
0.17	0.21	Standard Deviation	al	
STRONG	MODERATE	Range of Knowledge		

Listening Grades 9-12

Table 9C5L: Summary of Balance of Representation

		Percent of Total	Total Hits		Index	
	Standards	Mean	Standard Deviation	Mean	Standard Deviation	balance of Representation
LRN						
STRT	8	17.27%	8.86	0.75	0.17	STRONG
LISTEN	6	31.80%	9.96	0.71	0.13	STRONG

Reading Grades K-1

Table 1C1R: Summary of Alignment Results

			Star	ıdard Alignmer	Standard Alignment of Grades K-1 Reading Items	Reading Items		
		Aligni	Alignment Statistics	ics		Alignme	Alignment Findings	
Standards	CAT	LDL	CAT LDL Range	Balance	CAT	LDL	Range	Balance
LRN STRT	10.25	10.25 41% 30%	-	0.84	STRONG	MODERATE	LIMITED	STRONG
READ	35.50	35.50 62% 59%	59%	0.80	STRONG	STRONG	STRONG	STRONG

Table 1C2R: Summary of LDL Consistency

READ	LRN STRT			
11	×	Standards		
35.50	10.25	Mean		
35.50 8.04	10.73	Standard Deviation		Hits
33%	35%	Mean Percent Below	U	
0.14	0.29	Standard Deviation	Under	Pe
0.62	0.41	Mean Percent At	A	Percent of Questions
0.13	0.17	Standard Deviation		ons at LDL Level
0.05	0.24	Mean Percent Above	Above	vel
0.04	0.24	Standard Deviation	ove	
0.04 STRONG	0.24 MODERATE	Consistency		

Reading Grades K-1

Concurrence
al
Categoric
f
Summary of
Ż
3R
\mathbf{C}
-
Table

	Standards	LDL Level	No. of Standards	Percentage of Standards	Coded Mean	Standard Deviation	Categorical Concurrence
		1	0	0%0			
		2	6	75%			
STRT	8	3	2	25%	10.25	10.73	STRONG
		2	5	50%			
		3	4	40%			
		1	2	18%			
READ	11	2	5	45%	35.50	8.04	STRONG
		£	£	27%			

Table 1C4R: Summary of Range-of-Knowledge Correspondence

LAS

Standards bandardsMo. of Standards Hit% of TotalStandards bandardsMean StandardsStandards Hit% of TotalMean beviationStandards HitsDeviation DeviationAvg. Percent of Standards Hit810.2510.732.381.5130%1135.508.046.501.3159%				5 7 :11		Range of	Range of Standards		
Standards MeanStandard Standards DeviationMean Standards HitsMean Standards DeviationMean Avg. Percent of Standards Hit of Standards Hit810.2510.732.381.5130%1135.508.046.501.3159%			-		No. of St	andards Hit	% of Total		
8 10.25 10.73 2.38 1.51 30% 11 35.50 8.04 6.50 1.31 59%		Standards	Mean	Standard Deviation	Mean Standards Hits	Standard Deviation	Avg. Percent of Standards Hit	Standard Deviation	Range of Knowledge
8 10.25 10.73 2.38 1.51 30% 11 35.50 8.04 6.50 1.31 59%	LRN								
11 35.50 8.04 6.50 1.31 59%	STRT	8	10.25	10.73	2.38	1.51	30%	0.19	LIMITED
	READ	11	35.50	8.04	6.50	1.31	59%	0.12	STRONG

Reading Grades K-1

Table 1C5R: Summary of Balance of Representation

		Percent of Total Hits	Total Hits		Index	
	Standards	Mean	Standard Deviation	Mean	Standard Deviation	ыаапсе от Representation
LRN	8	0 40%	10 73	0 84	0 1 5	STRONG
READ	11	32.57%	8.04	0.80	0.08	STRONG

Reading Grades 2-3

Table 2C1R: Summary of Alignment Results

				Standard Al	ignment of G	Standard Alignment of Grades 2-3 Reading Items	Items	
		Alignme	lignment Statistics	S			Alignment Findings	SBI
Standards	CAT	LDL	Range	Balance	CAT	LDL	Range	Balance
LRN STRT	16.38	48%	42%	0.77	STRONG	STRONG MODERATE	MODERATE	STRONG
READ	37.13	61%	61%	0.67	STRONG	STRONG	STRONG	MODERATE

Table 2C2R: Summary of LDL Consistency

			Hits		Percei	Percent of Questions at LDL Level	s at LDL Leve	el		
				Under	ler	A	t	AI	Above	
	Standards	Mean	Standard Deviation	Mean Percent Below	Standard Deviation	Mean Percent At	Standard Deviation	Mean Percent Above	Standard Deviation	Consistency
LRN										
STRT	8	16.38	13.24	17%	0.20	0.48	0.30	0.35	0.34	MODERATE
READ	11	37.13	9.96	21%	0.17	0.61	0.17	0.18	0.19	STRONG

LAS

Reading Grades 2-3

READ	LRN STRT			
Ð	ηz			
11	8	Standards		
37.13	16.38	Mean		
9.96	13.24	Standard Deviation	11163	Hite
6.75	3.38	Mean Standards Hits	No. of Standards Hit	R
1.39	2.26	Standard Deviation	ds Hit	Range of Standards
61%	42%	Avg. Percent of Standards Hit	% of Total	dards
0.13	0.28	Standard Deviation	Fotal	
STRONG	MODERATE	Range of Knowledge		

READ	LRN STRT	
11	∞	Standards
2 1	2 3 3	LDL Level
3 5	4 5 2 6 0	No. of Standards
18% 45% 27%	0% 75% 25% 50% 40%	Percentage of Standards
37.13	16.38	Coded Mean
9.96	13.24	Standard Deviation
STRONG	STRONG	Categorical Concurrence

Table 2C4R: Summary of Range-of-Knowledge Correspondence

LAS

Reading Grades 2-3

Table 2C3R: Summary of Categorical Concurrence

	د ۲ ۲	balance of Representation	STRONG	MODERATE
	Index	Standard Deviation	0.15	0.05
		Mean	0.77	0.67
	of Total Hits	Standard Deviation	13.24	9.96
	Percent of Total	Mean	13.95%	31.63%
•		Standards	8	11
-			LRN STRT	READ

Table 2C5R: Summary of Balance of Representation

Reading Grades 4-5

Table 4C1R: Summary of Alignment Results

			Sta	ndard Alignn	nent of Grades	Standard Alignment of Grades 4-5 Reading Items		
		Alignme	Alignment Statistics	-		Alig	Alignment Findings	
Standards CAT		LDL	Range Balance	Balance	CAT	LDL	Range	Balance
LRN STRT	19.00	49%	42%	0.80	STRONG	MODERATE	MODERATE	STRONG
READ	45.00	60%	64%	0.64	STRONG	STRONG	STRONG	MODERATE

Table 4C2R: Summary of LDL Consistency

	Standards		Hits Standard		Under Crandard	nt of Questior	Percent of Questions at LDL Level At Mean Standard		Ve
	Standards	Mean	Standard Deviation	Mean Percent Below	Standard Deviation	Mean Percent At	Standard Deviation	Mean Percent Above	Standard Deviation
LRN	8	19 00	14 47	%00	0 13	0 4 9	0 12	0 42	0 14 MODERATE
READ	11	45.00	14.38	23%	0.15	0.60	0.17	0.18	0.18

Reading Grades 4-5

Concurrence
Ξ
f Categorical
4
0
Summary o
C3R
ŝ
\bigcirc
4
e
le
ab
Ë

	Standards	LDL Level	No. of Standards	Percentage of Standards	Coded Mean	Standard Deviation	Categorical Concurrence
		1	0	%0			
LKN	8	2	9	0%5L	19.00	14.47	STRONG
		3	2	25%			
		1	2	18%			
READ	11	2	5	45%	45.00	14.38	STRONG
		3	3	27%			

Table 4C4R: Summary of Range-of-Knowledge Correspondence

LAS

				Ran	Range of Standards	ards			
			SIII	No. of Standards Hit	Hit	% of Total	Fotal		
	Standards	Mean	Standard Deviation	Mean Standards Hits	Standard Deviation	Avg. Percent of Standards Hit	Standard Deviation	Range of Knowledge	
STRT	8	19.00	14.47	3.375	2.39	42%	0.30	MODERATE	
READ	11	45.00	14.38	7	0.76	64%	0.07	STRONG	

Reading Grades 4-5

Table 4C5R: Summary of Balance of Representation

READ	LRN STRT			
11	8	Standards		
45.00	19.00	Mean		
14.38	14.47	Standard Deviation	LIIIS	
7	3.375	Mean Standards Hits	No. of Standards Hit	Rai
0.76	2.39	Standard Deviation	s Hit	Range of Standards
64%	42%	Avg. Percent of Standards Hit	% of Total	ards
0.07	0.30	Standard Deviation	Total	
STRONG	MODERATE	Range of Knowledge		

Reading Grades 6-8

Table 6C1R: Summary of Alignment Results

			Standard A	Mignment of Gr	tandard Alignment of Grades 6-8 Reading Items	g Items		
		Alignment	nt Statistics			Alignm	dignment Findings	
Standards	CAT	LDL	Range	Balance	CAT	LDL	Range	Balance
LRN STRT	23.88	56%	55%	0.77	STRONG	STRONG	STRONG	STRONG
READ	44.13	57%	59%	0.72	STRONG	STRONG	STRONG	STRONG

Table 6C2R: Summary of LDL Consistency

			Hits		Perce	Percent of Questions at LDL Level	ns at LDL Le	vel		
				Under	ler	A	t	Ab	Above	
	Standards	Mean	Standard Deviation	Mean Percent Below	Standard Deviation	Mean Standard Percent At Deviation	Standard Deviation	Mean Percent Above	Standard Deviation	Consistency
1										
	8	23.88	10.72	7%	0.10	0.56	0.23	0.37	0.20	STRONG
	11	44.13	14.00	22%	0.11	0.57	0.13	0.21	0.11	STRONG

LAS

Reading
Grades
6-8

Table 6C3R: Summary of Categorical Concurrence

	READ			STRT	LRN			
	11	Γ		G	×	Ι	Standards	
3	2	1	3	3	2	1	Level	LDL
3	5	2	4	2	6	0	Standards	No. of
27%	45%	18%	40%	25%	75%	0%	Standards	Percentage of
	44.13			20.00	22 22		Mean	Coded
	14.00			10.72	10 72		Deviation	Standard
	STRONG			O FROM	STRONG		Concurrence	Categorical

Table 6C4R: Summary of Range-of-Knowledge Correspondence

LAS

READ	LRN STRT			
11	8	Standards		
44.13	23.88	Mean		
14.00	10.72	Standard Deviation	IIIts	Hite
6.5	4.375	Mean Standards Hits	No. of Standards Hit	
1.31	2.00	Standard Deviation	rds Hit	Range of Standards
59%	55%	Avg. Percent of Standards Hit	% of Total	indards
0.12	0.25	Standard Deviation	al	
STRONG	STRONG	Range of Knowledge		

Reading Grades 6-8

Table 6C5R: Summary of Balance of Representation

L

		Percent of Total	of Total Hits		Index	
	Standards	Mean	Standard Deviation	Mean	Standard Deviation	Balance of Representation
LRN						
STRT	8	15.50%	10.72	0.77	0.12	STRONG
READ	11	28.65%	14.00	0.72	0.04	STRONG

LAS

Reading Grades 9-12

Table 9C1R: Summary of Alignment Results

			Standar	d Alignment of	Standard Alignment of Grades 9-12 Reading Items	eading Items		
		Alignme	Alignment Statistics			Ali	Alignment Findings	
Standards	CAT	LDL	Range	Balance	CAT	LDL	Range	Balance
LRN STRT	18.88	55%	42%	0.81	STRONG	STRONG	MODERATE	STRONG
READ	46.38	56%	57%	0.72	STRONG	STRONG	STRONG	STRONG

Table 9C2R: Summary of LDL Consistency

		Н	Hits		Per	Percent of Questions at LDL Level	s at LDL Lev	el		
				U	Under	At		Above	ove	
	Standards	Mean	Standard Deviation	Mean Percent Below	Standard Deviation	Mean Percent At	Standard Deviation	Mean Percent Above	Standard Deviation	Consistency
LRN STRT	8	18.88	13.01	12%	0.13	0.55	0.17	0.33	0.22	STRONG
READ	11	46.38	14.83	25%	0.09	0.56	0.18	0.19	0.17	STRONG

Reading Grades 9-12

Concurrence
Categorical
f
Summary o
::
2
3
C3R
×
U 1
Table

	Standards	LDL Level	No. of Standards	Percentage of Standards	Coded Mean	Standard Deviation	Categorical Concurrence
		1	0	0%0			
LRN	0	2	9	75%	10 00	12 01	UNO ALS
STRT	0	3	2	25%	10.00	10.01	DNIONIC
		3	4	40%			
		1	2	18%			
READ	11	2	5	45%	46.38	14.83	STRONG
		3	3	27%			

Table 9C4R: Summary of Range-of-Knowledge Correspondence

LAS

	Hits
No. of Standards Hit	
Mean Standards Hits	Standard Me Deviation
3 38	13.01
n	10.01
6.25	14.83

Reading Grades 9-12

Table 9C5R: Summary of Balance of Representation

		Percent of Total Hits	Total Hits		Index	
	Standards	Mean	Standard Deviation	Mean	Standard Deviation	Balance of Representation
LRN						
STRT	8	12.93%	13.01	0.81	0.11	STRONG
READ	11	31.76%	14.83	0.72	0.05	STRONG

Speaking Grades K-1

Table 1C1S: Summary of Alignment Results

				Standard A	Alignment of	tandard Alignment of Grades K-1 Speaking Items	ing Items	
		Alignmer	lignment Statistics	S			Alignment Findings	dings
Standards	CAT	LDL	Range	Balance	CAT	LDL	Range	Balance
LRN STRT 24.13	24.13	56%	42%	0.73	STRONG	STRONG	MODERATE	STRONG
SPEAK	52.00	46%	966%	0.68	STRONG	STRONG MODERATE	STRONG	MODERATE

Table 1C2S: Summary of LDL Consistency

			Hits		Percent	t of Questio	Percent of Questions at LDL Level	/el		
				Under	ľ	7	٨t	Ab	Above	
	Standards	Mean	Standard Deviation	Mcan Percent Below	Standard Deviation	Mean Percent At	Standard Deviation	Mean Percent Above	Standard Deviation	Consistency
LRN STRT	×	74.13	15 20	1 50%	0.14	0 56	0.00	0.79	0.16	STRONG
SPEAK	10	61.72 62.00	15.20	%6C	0.21	0.20	0.0	0.25	0.10	MODERATE
	01	21	10:00	0/01	11.0	00	01.0	0.10	0.11	

LAS

Speaking Grades K-1

Table 1C3S: Summary of Categorical Concurrence

	SPEAK			STRT	LRN			
	10			G	ø		Standards	
3	2	1	3	3	2	1	Level	LDL
4	5	1	4	2	6	0	Standards	No. of
40%	50%	10%	44%	25%	75%	0%	Standards	Percentage of
52.00				£4.10	2/ 12		Mean	Coded
16.00				10.20	15 20		Deviation	Standard
	STRONG			o I RONG	STRONG		Concurrence	Categorical

Table 1C4S: Summary of Range-of-Knowledge Correspondence

7.0				
SPEAK	LRN STRT			
10	8	Standards		
52.00	24.13	Mean		
16.00	15.20	Standard Deviation	IIII	
6.63	3.38	Mean Standards Hits	No. of Standards Hit	
1.77	1.77	Standard Deviation	rds Hit	Range of Standards
66%	42%	Avg. Percent of Standards Hit	% of Total	ndards
0.18	0.22	Standard Deviation	Total	
STRONG	MODERATE	Range of Knowledge		

Speaking Grades K-1

Table 1C5S: Summary of Balance of Representation

		Percent of	Percent of Total Hits	, _	Index	a
	Standards	Mean	Standard Deviation	Mean	Standard Deviation	balance of Representation
LRN						
STRT	8	14.90%	15.20	0.73	0.18	STRONG
SPEAK	10	32.12%	16.00	0.68	0.06	MODERATE

Speaking Grades 2-3

Table 2C1S: Summary of Alignment Results

			Standa	rd Alignment c	Standard Alignment of Grades 2-3 Speaking Items	eaking Items		
		Alignme	Alignment Statistics			Alignme	nt Findings	
Standards	CAT	LDL	Range	Balance	CAT	LDL	Range	Balance
LRN STRT	28.25	41%	53%	0.80	STRONG	MODERATE	STRONG	STRONG
SPEAK	56.00	46%	70%	0.72	STRONG	MODERATE	STRONG	STRONG

Table 2C2S: Summary of LDL Consistency

	<u> </u>		<u> </u>		
SPEAK	STRT	LRN			
10	8		Standards		
52.00	24.13		Mean		
16.00	15.20		Standard Deviation		Hits
29%	15%		Mean Percent Below	Under	
0.21	0.14		Standard Deviation	r	Percen
0.46	0.56		Mean Percent At		t of Questio
0.15	0.09		Standard Deviation	At	Percent of Questions at LDL Level
0.25	0.29		Mean Percent Above	Al	vel
0.12	0.16		Standard Deviation	Above	
0.25 0.12 MODERATE	STRONG		Consistency		

LAS

Speaking Grades 2-3

	Standards	Level	No. of Standards	Percentage of Standards	Coded Mean	Standard Deviation	Categorical Concurrence
		1	0	0%0			
		2	9	75%			
LKN	8	3	2	25%	28.25	15.38	STRONG
		2	3	33%			
		3	4	44%			
		1	1	10%			
SPEAK	10	2	5	50%	56.00	19.99	STRONG
		3	7	40%			

Table 2C3S: Summary of Categorical Concurrence

Table 2C4S: Summary of Range-of-Knowledge Correspondence

LAS

			11:45	Ran	Range of Standards	rds		
				No. of Standards Hit	Hit	% of Total	Total	
	Standards	Mean	Standard Deviation	Mean Standards Hits	Standard Deviation	Avg. Percent of Standards Hit	Standard Deviation	Range of Knowledge
LRN								
STRT	8	28.25	15.38	4.25	2.12	53%	0.27	STRONG
SPEAK	10	56.00	19.99	7.00	1.20	70%	0.12	STRONG

Speaking Grades 2-3

Table 2C5S: Summary of Balance of Representation

		Percent of	Percent of Total Hits		Index	
	Standards	Mean	Standard Deviation	Mean	Standard Deviation	Balance of Representation
LRN						
STRT	8	16.19%	15.38	0.796	0.165	STRONG
SPEAK	10	32.09%	19.99	0.724	0.039	STRONG

Speaking Grades 4-5

Table 4C1S: Summary of Alignment Results

			Standard	Alignment of	Standard Alignment of Grades 4-5 Speaking Items	king Items		
		Alignme	lignment Statistics			Alignment Findings	Findings	
Standards	CAT	LDL	Range	Balance	CAT	LDL	Range	Balance
LRN STRT	33.75	43%	61%	0.76	STRONG	MODERATE	STRONG	STRONG
SPEAK	60.13	40%	75%	0.77	STRONG	MODERATE	STRONG	STRONG

Table 4C2S: Summary of LDL Consistency

L

	F	tency	RATE	RATE
		Consistency	MODERATE	MODERATE
	ve	Standard Deviation	0.26	0.20
evel	Above	Mean Percent Above	0.46	0.38
ins at LDL L		Standard Deviation	0.20	0.11
Percent of Questions at LDL Level	At	Mean Percent At	0.43	0.40
Per	Under	Standard Deviation	0.10	0.10
	Un	Mean Percent Below	11%	22%
Hits		Standard Deviation	11.07	20.18
		Mean	33.75	60.13
		Standards	∞	10
			LRN STRT	SPEAK

LAS

Speaking Grades 4-5

Table 4C3S: Summary of Categorical Concurrence

SPEAK 10				STRT 8			Standards I	
2	1	3	2	3	2	1	Level	LDL
4 5	1	4	3	2	6	0	Standards	No. of
50% 40%	10%	44%	33%	25%	75%	0%	Standards	Percentage of
60.13			<u>.</u>	33.75		<u> </u>	Mean	
20.18				11.07			Deviation	Standard
STRONG				STRONG			Concurrence	Categorical

Table 4C4S: Summary of Range-of-Knowledge Correspondence

LAS

SPEAK	LRN STRT			
10	8	Standards		
60.13	33.75	Mean		
20.18	11.07	Standard Deviation	1103	Hite
7.5	4.875	Mean Standards Hits	No. of Standards Hit	Ra
1.20	2.10	Standard Deviation	s Hit	Range of Standards
75%	61%	Avg. Percent of Standards Hit		lards
0.12	0.26	Standard Deviation	% of Total	
STRONG	STRONG	Range of Knowledge		

Speaking Grades 4-5

Table 4C5S: Summary of Balance of Representation

L

		Percent of Total H	Total Hits		Index	د ۲
	Standards	Mean	Standard Deviation	Mean	Standard Deviation	balance of Representation
LRN						
STRT	8	17.33%	11.07	0.76	0.08	STRONG
SPEAK	10	30.87%	20.18	0.77	0.06	STRONG

Speaking Grades 6-8

Table 6C1S: Summary of Alignment Results

			Standard	d Alignment of	Standard Alignment of Grades 6-8 Speaking	king Items		
		Alignme	Alignment Statistics			Alignment Finding	t Findings	
Standards	CAT	LDL	Range	Balance	CAT	LDL	Range	Balance
LRN STRT	35.50	40%	59%	0.75	STRONG	MODERATE	STRONG	STRONG
SPEAK	60.00	47%	76%	0.74	STRONG	MODERATE	STRONG	STRONG

Table 6C2S: Summary of LDL Consistency

	1	1		
SPEAK	LRN STRT			
10	8	Standards		
60.00	35.50	Mean		
60.00 20.69	8.88	Standard Deviation		Hits
23%	14%	Mean Percent Below	Un	
0.20	0.17	Standard Deviation	Under	Perce
0.47	0.40	Mean Percent At	At	Percent of Questions at LDL Level
0.11	0.30	Standard Deviation		at LDL Level
0.31	0.46	Mean Percent Above	Above	
0.21	0.32	Standard Deviation	ve	
0.21 MODERATE	0.32 MODERATE	Consistency		

Speaking Grades 6-8

Concurrence
Ξ
Categorica
÷
0
Summary of
2
23
\mathbf{i}
9
Fable

	Standards	LDL Level	No. of Standards	Percentage of Standards	Coded Mean	Standard Deviation	Categorical Concurrence
		1	0	0%0			
		2	6	75%			
STRT	8	3	2	25%	35.50	8.88	STRONG
		2	3	33%			
		3	4	44%			
		1	1	10%			
SPEAK	10	2	5	50%	60.00	20.69	STRONG
		3	4	40%			

Table 6C4S: Summary of Range-of-Knowledge Correspondence

LAS

			U :46	F	Range of Standards	Idards		
				No. of Standards Hit	rds Hit	% of	% of Total	
	Standards	Mean	Standard Deviation	Mean Standards Hits	Standard Deviation	Avg. Percent of Standards Hit	Standard Deviation	Range of Knowledge
LRN STRT	~	35.50	8.88	4.75	1.58	59%	0.20	STRONG
SPEAK	10	60.00	20.69	7.625	1.41	%9L	0.14	STRONG

Speaking Grades 6-8

Table 6C5S: Summary of Balance of Representation

		Percent of Total Hits	Total Hits		Index	
	Standards	Mean	Standard Deviation	Mean	Standard Deviation	Balance of Representation
LRN						
STRT	8	17.75%	8.88	0.748	0.075	STRONG
SPEAK	10	30.00%	20.69	0.744	0.100	STRONG

9-12
Grades
Speaking

Table 9C1S: Summary of Alignment Results

StandardsCATLDLRangeBalanceCATLDLRangeBalanceLRN STRT34.0028%55%0.72STRONGLIMITEDSTRONGSTRONGSPEAK65.2543%65%0.69STRONGMODERATESTRONGMODERATE				Š	Standard Alig	nment of Grad	ird Alignment of Grades 9-12 Speaking Items	ems	
CATLDLRangeBalanceCATLDLRange34.0028%55%0.72STRONGLIMITEDSTRONG65.2543%65%0.69STRONGMODERATESTRONG			Alignme	nt Statistics				Alignment Fin	dings
34.00 28% 55% 0.72 STRONG LIMITED STRONG 65.25 43% 65% 0.69 STRONG MODERATE STRONG	Standards	CAT	LDL	Range	Balance	CAT	LDL	Range	Balance
65.25 43% 65% 0.69 STRONG MODERATE STRONG	LRN STRT	34.00	28%	55%	0.72	STRONG	LIMITED	STRONG	STRONG
	SPEAK	65.25	43%	65%	0.69	STRONG	MODERATE	STRONG	MODERATE

Table 9C2S: Summary of LDL Consistency

			Hits		P(Percent of Questions at LDL Level	ions at LDL L	evel		
				Ur	Under	At		Above	ve	I NI
	Standards	Mean	Standard Deviation	Mean Percent Below	Standard Deviation	Mean Percent At	Standard Deviation	Mean Percent Above	Standard Deviation	Consistency
LRN STRT	~	34.00	34.00 13.55	19%	0.28	0.28	0.23	0.53	0.32	LIMITED
SPEAK	10	65.25	20.87	22%	0.26	0.43	0.18	0.35	0.23	MODERATE

LAS

Speaking
Grades
9-12

Table 9C3S: Summary of Categorical Concurrence

Table 9C4S: Summary of Range-of-Knowledge Correspondence

LAS

SPEAK	LRN STRT			
10	×	Standards		
65.25	34.00	Mean	-	
20.87	13.55	Standard Deviation	III	Hite
6.5	4.375	Mean Standards Hits	No. of Standards Hit	
0.93	1.69	Standard Deviation	rds Hit	Range of Stand
65%	55%	Avg. Percent of Standards Hit	% of Total	idards
0.09	0.21	Standard Deviation	Total	
STRONG	STRONG	Range of Knowledge		

Speaking Grades 9-12

Table 9C5S: Summary of Balance of Representation

		Percent (Percent of Total Hits	ļ	Index	J
	Standards	Mean	Standard Deviation	Mean	Standard Deviation	balance of Representation
LRN						
STRT	8	16.06%	13.55	0.72	0.15	STRONG
SPEAK	10	30.81%	20.87	0.69	0.15	MODERATE

Writing Grades K-1

Table 1C1W: Summary of Alignment Results

			Ste	andard Alignn	Standard Alignment of Grades K-1 Writing	K-1 Writing Items		
		Alignme	Alignment Statistics			Ali	Mignment Finding	8
Standards	CAT	LDL	Range	Balance	CAT	LDL	Range	Balance
LRN STRT	18.88	44%	33%	0.82	STRONG	MODERATE	LIMITED	STRONG
WRITE	34.38	59%	75%	0.76	STRONG	STRONG	STRONG	STRONG

Table 1C2W: Summary of LDL Consistency

		Hits		Inder Po	Percent of Questions at	estions at LDI	t LDL Level	Ahove	
Standards	ırds Mean	Standar Deviatio	d Mean on Percent J Below	Standard Deviation	Mean Percent At	Standard Deviation	Mean Percent Above	Standarc Deviatior	dard ation
LRN									
STRT 8	18.8	8 14.33	39%	0.15	0.44	0.10	0.16	0.	0.12
WRITE 7	34.3	34.38 5.13 27% 0.17	27%	0.17	0.59	0.15	0.14	0	0.15

Writing Grades K-1

Concurrence
cal
ić,
tegorica
50
ate
\mathbf{O}
G
Summary
3
C3W
-
-
able
ał
Η

	Standards	LDL Level	No. of Standards	Percentage of Standards	Coded Mean	Standard Deviation	Categorical Concurrence
		1	0	0%0			
		2	6	75%			
STRT	8	3	2	25%	18.88	14.33	STRONG
		2	5	45%			
		3	3	27%			
		1	1	14%			
WRITE	7	2	4	57%	34.38	5.13	STRONG
		3	2	29%			

Table 1C4W: Summary of Range-of-Knowledge Correspondence

LAS

			л;+с	Í	Range of Standards	Idards		
			5111	No. of Standards Hit	ds Hit	% of	% of Total	
	Standards	Mean	Standard Deviation	Mean Standards Hits	Standard Deviation	Avg. Percent of Standards Hit	Standard Deviation	Range of Knowledge
LRN STRT	∞	18.88	14.33	2.625	1.77	33%	0.22	LIMITED
WRITE	7	34.38	5.13	5.25	0.89	75%	0.13	STRONG

Writing Grades K-1

Table 1C5W: Summary of Balance of Representation

		Percent o	Percent of Total Hits		Index	
	Standards	Mean	Standard Deviation	Mean	Standard Deviation	Balance or Representation
LRN	8	15.70%	14.33	0.82	0.18	STRONG
WRITE	7	28.59%	5.13	0.76	0.05	STRONG

Table 2C1W: Summary of Alignment Results

Writing Grades 2-3

	Alignment Findings	Range Balance	MODERATE STRONG	STRONG STRONG	
ing Items	Ali	R			
des 2-3 Writ		LDL	STRONG STRONG	STRONG STRONG	
ndard Alignment of Grades 2-3 Writing Items		CAT	STRONG	STRONG	
Standard Ali		Balance	0.77	0.76	
•	Alignment Statistics	Range	41%	71%	
	Alignmer	LDL	55%	54%	
		CAT	22.50	38.50	
		Standards	LRN STRT	WRITE	

Table 2C2W: Summary of LDL Consistency

			Hits		Percent	of Questions	Percent of Questions at LDL Level	- F		
				Under		7	At	Ab	Above	
	Standards	Mean	Standard Deviation	Mean Percent Below	Standard Deviation	Mean Percent At	Standard Deviation	Mean Percent Above	Standard Deviation	Consistency
LRN STRT	~	22.50	14.76	23%	0.12	0.55	0.17	0.22	0.14	STRONG
WRITE	7	38.50	10.42	28%	0.21	0.54	0.10	0.17	0.14	STRONG

LAS Links

Writi
ing G
irade
es 2-3

Table 2C3W: Summary of Categorical Concurrence

3 2	WRITE 7 2 4 57% 38.50	3 3 27%	2 5 45%	STRT 8 3 2 25% 22.50	2 6 75%	LDL No. of Percentage of Coded Standards Level Standards Standards Mean
	10.42 STRONG			14.76 STRONG		Standard Categorical Deviation Concurrence

Table 2C4W: Summary of Range-of-Knowledge Correspondence

LAS

WRITE	LRN STRT			
7	×	Standards		
38.50	22.50	Mean		
10.42	14.76	Standard Deviation	IIIts	Hite
5.00	3.25	Mean Standards Hits	No. of Standards Hit	-
1.51	2.31	Standard Deviation	ds Hit	Range of Standards
71%	41%	Avg. Percent of Standards Hit	% of Total	ndards
0.22	0.29	Standard Deviation	Total	
STRONG	MODERATE	Range of Knowledge		

Writing Grades 2-3

Table 2C5W: Summary of Balance of Representation

StandardsMeanStandardStandardIMeanDeviationDeviationDeviationI817.65%14.760.770.16730.20%10.420.760.07			Percent o	f Total Hits		Index	a
8 17.65% 14.76 0.77 0.16 7 30.20% 10.42 0.76 0.07		Standards	Mean	Standard Deviation	Mean	Standard Deviation	balance of Representation
8 17.65% 14.76 0.77 0.16 7 30.20% 10.42 0.76 0.07	LRN						
7 30.20% 10.42 0.76 0.07	STRT	8	17.65%	14.76	0.77	0.16	STRONG
	WRITE	7	30.20%	10.42	0.76	0.07	STRONG

Writing Grades 4-5

Table 4C1W: Summary of Alignment Results

			Stan	dard Alignmei	Standard Alignment of Grades 4-5 Writing	Writing Items		
		Alignme	Alignment Statistics			7	Alignment Findings	
Standards	CAT	LDL	Range	Balance	CAT	LDL	Range	Balance
LRN STRT	19.88	64%	41%	0.82	STRONG	STRONG	MODERATE	STRONG
WRITE	40.88	58%	63%	0.78	STRONG	STRONG	STRONG	STRONG

Table 4C2W: Summary of LDL Consistency

			Hits		Per	Percent of Questions at LDL Level	is at LDL Lev	vel		
				U	Under	At		Above	ve	
	Standards	Mean	Standard Deviation	Mean Percent Below	Standard Deviation	Mean Percent At	Standard Deviation	Mean Percent Above	Standard Deviation	Consistency
LRN										
STRT	8	19.88	13.88	7%	0.12	0.64	0.25	0.30	0.24	STRONG
WRITE	7	40.88	10.34	9%	0.10	0.58	0.15	0.33	0.13	STRONG

Writing Grades 4-5

Concurrence
f Categorical
Summary of
Table 4C3W:

	Standards	LDL Level	No. of Standards	Percentage of Standards	Coded Mean	Standard Deviation	Categorical Concurrence
		1	0	0%0			
		2	6	75%			
LKN	8	3	2	25%	19.88	13.88	STRONG
		2	5	45%			
		3	3	27%			
		1	1	14%			
WRITE	7	2	4	57%	40.88	10.34	STRONG
		ю	2	29%			

Table 4C4W: Summary of Range-of-Knowledge Correspondence

LAS

			U :46	R	Range of Standards	dards		
			SIII	No. of Standards Hit	ds Hit	J0 %	% of Total	
	Standards	Mean	Standard Deviation	Mean Standards Hits	Standard Deviation	Avg. Percent of Standards Hit	Standard Deviation	Range of Knowledge
LRN								
STRT	8	19.88	13.88	3.25	1.83	41%	0.23	MODERATE
WRITE	7	40.88	10.34	4.375	1.19	63%	0.17	STRONG

Writing Grades 4-5

STRT 8	LRN	Standards		
40.88	19.88	Mean		
10.34	13.88	Standard Deviation		Hits
4.375	3.25	Mean Standards Hits	No. of Standards Hit	H
1.19	1.83	Standard Deviation	ds Hit	Range of Sta
63%	41%	Avg. Percent of Standards Hit	% of	Standards
0.17	0.23	Standard Deviation	% of Total	
STRONG	MODERATE	Range of Knowledge		

LAS

Table 4C5W: Summary of Balance of Representation

6-8
Grades
Writing

Table 6C1W: Summary of Alignment Results

				Standard Ali	ignment of Gra	lard Alignment of Grades 6-8 Writing Items	ems	
		Alignmeı	Mignment Statistics				Alignment Findings	
Standards	CAT	LDL	Range	Balance	CAT	LDL	Range	Balance
LRN STRT	22.63	49%	42%	0.82	STRONG	MODERATE	MODERATE	STRONG
WRITE	38.88	47%	71%	0.76	STRONG	MODERATE	STRONG	STRONG

Table 6C2W: Summary of LDL Consistency

			Hits		Per	Percent of Questions at LDL Level	ins at LDL L	evel		
				Un	Under	At		Above	ve	
	Standards	Mean	Standard Deviation	Mean Percent Below	Standard Deviation	Mean Standard Percent At Deviation	Standard Deviation	Mean Percent Above	Standard Deviation	Consistency
LRN										
STRT	8	22.63	11.77	15%	0.19	0.49	0.14	0.36	0.21	MODERATE
WRITE	7	38.88	9.19	25%	0.22	0.47	0.18	0.28	0.15	MODERATE

LAS

Writing Grades 6-8

Table 6C3W: Summary of Categorical Concurrence

Table 6C4W: Summary of Range-of-Knowledge Correspondence

LAS

WRITE	LRN STRT			
7	8	Standards		
38.88	22.63	Mean		
9.19	11.77	Standard Deviation	1115	Hite
5	3.375	Mean Standards Hits	No. of Standards Hit	
0.76	1.92	Standard Deviation	rds Hit	Range of Standards
71%	42%	Avg. Percent of Standards Hit	% of	dards
0.11	0.24	Standard Deviation	% of Total	
STRONG	MODERATE	Range of Knowledge		

Writing Grades 6-8

Table 6C5W: Summary of Balance of Representation

StandardsMeanStandardStandardBalance ofLRNMeanDeviationDeviationRepresentationLRN817.50%11.770.820.11STRT830.08%9.190.760.06STRONG			Percent of	Percent of Total Hits	,	Index	
8 17.50% 11.77 0.82 0.11 7 30.08% 9.19 0.76 0.06		Standards	Mean	Standard Deviation	Mean	Standard Deviation	balance of Representation
8 17.50% 11.77 0.82 0.11 7 30.08% 9.19 0.76 0.06	LRN						
7 30.08% 9.19 0.76 0.06	STRT	8	17.50%	11.77	0.82	0.11	STRONG
	WRITE	7	30.08%	9.19	0.76	0.06	STRONG

Writing Grades 9-12

Table 9C1W: Summary of Alignment Results

				Standard Al	lignment of Gr	Standard Alignment of Grades 9-12 Writing	Items	
		Alignmer	Alignment Statistics	S			Alignment Findings	gs
Standards	CAT	LDL	Range	Range Balance	CAT	LDL	Range	Balance
LRN STRT	21.63	48%	41%	0.82	STRONG	MODERATE	MODERATE	STRONG
WRITE	40.50	51%	71%	0.75	STRONG	STRONG	STRONG	STRONG

Table 9C2W: Summary of LDL Consistency

	Standards		Hits		Under Perce	Percent of Questions at LDL Level	s at LDL Le	Above		
	Standards	Mean	Standard Deviation	Mean Percent Below	Standard Deviation	Mean Percent At	Standard Deviation	Mean Percent Above	an ent we	an Standard cent Deviation
LRN	þ	2	5	110/		0	012		-	2
STRT	8	21.63	12.42	11%	0.14	0.48	0.13	0.	0.41	41 0.23 MODERATE
WRITE	7	40.50	9.35	18%	0.21	0.51	0.12	~	0.31	0.31 0.14

Writing Grades 9-12

ζ	Concurrence
-	a
	Categoric
¢	Ē
7	: Summary 0
	₹.
۶	5
ć	5
è	5
E	I able

	Standards	LDL Level	No. of Standards	Percentage of Standards	Coded Mean	Standard Deviation	Categorical Concurrence
		1	0	0%			
		2	9	75%			
LKN	8	3	2	25%	21.63	12.42	STRONG
		2	5	45%			
		3	3	27%			
		1	1	14%			
WRITE	7	2	4	57%	40.50	9.35	STRONG
		3	2	29%			

Table 9C4W: Summary of Range-of-Knowledge Correspondence

LAS

			11:4~	I	Range of Standards	Idards			
			211	No. of Standards Hit	rds Hit	% of	% of Total		
	Standards	Mean	Standard Deviation	Mean Standards Hits	Standard Deviation	Avg. Percent of Standards Hit	Standard Deviation	Range of Knowledge	
LRN STRT	8	21.63	12.42	3.25	1.83	41%	0.23	MODERATE	
WRITE	7	40.50	9.35	5	0.53	71%	0.08	STRONG	

Writing Grades 9-12

Table 9C5W: Summary of Balance of Representation

	T	Percent of	Percent of Total Hits		Index	
	Standards	Mean	Standard Deviation	Mean	Standard Deviation	Balance of Representation
LRN						
STRT	8	16.44%	12.42	0.82	0.14	STRONG
WRITE	7	30.80%	9.35	0.75	0.05	STRONG



Appendix D: Results of Intra-class Correlation

Notes: The minimum number of judges calculation is based on the Spearman Browne Prophecy formula, $m = \begin{cases} \frac{\rho^*}{1-\rho^*} \\ \frac{\rho_L}{1-\rho_L} \end{cases} = \frac{\rho^* \langle 1-\rho_L \rangle}{\rho_L \langle 1-\rho^* \rangle}$

where ρ^* is the reliability aspired to and ρ_L is the reliability estimate for a single judge.

The two-way analysis assuming both random items and fixed judges gives a result for the mean correlation identical to Cronbach's

alpha, i.e., $\alpha = \frac{\sigma_{Bet}^e - \sigma_e^2}{\sigma_{Bet}^2}$. While SPSS allows the user to select between the random and mixed models, the calculations come out the

other measure of reliability). then the variability among judges must be taken into account, which will result in a lower value for the intra-class correlation (or any component of variance associated with them. Random judges assume the judges used are one of many possible selections of judges; same with either model. Assuming the judges are fixed would imply these are the only judges that would ever be used, so there is no

For the mixed model (i.e., fixed judges), the intra-class correlation would be calculated identically to Alpha

$$ICC_{Fixed, Judgs} = \frac{ItemMS - EMS}{ItemMS}$$

For the random model, the correct calculation is:

$$ICC_{RandomJudgs} = \frac{ItemMS - EMS}{ItemMS + \frac{\langle JudgeMS - EMS \rangle}{}}$$

п

Calculation for two-way model with both questions and judges random.

Listening Gr	ade K-1	
	dF	MS
Questions	19	1.46
Judges	7	1.34
Error	133	0.19
Intra-Class Correlation		0.84
Cronbach's Alpha		0.87

Listening G	rade 2-3	
	dF	MS
Questions	19	0.85
Judges	7	2.43
Error	133	0.21
Intra-Class Correlation		0.66
Cronbach's Alpha		0.75

Listening G	rade 4-5	
	dF	MS
Questions	19	0.81
Judges	7	0.97
Error	133	0.19
Intra-Class Correlation		0.73
Cronbach's Alpha		0.76

Calculation for two-way model with both questions and judges random.

Listening Gi	rade 6-8	
	dF	MS
Questions	22	0.70
Judges	7	1.19
Error	154	0.19
Intra-Class Correlation		0.69
Cronbach's Alpha		0.73

Listening Gr	ade 9-12	
	dF	MS
Questions	22	1.01
Judges	7	1.34
Error	154	0.21
Intra-Class Correlation		0.76
Cronbach's Alpha		0.79

Reading Gra	ade K-1	
	dF	MS
Questions	29	2.74
Judges	7	2.25
Error	203	0.17
Intra-Class Correlation		0.91
Cronbach's Alpha		0.94

Calculation for two-way model with both questions and judges random.

Reading Gr	ade 2-3	
	dF	MS
Questions	29	1.21
Judges	7	3.08
Error	203	0.28
Intra-Class Correlation		0.71
Cronbach's Alpha		0.77

Reading Gr	ade 4-5	
	dF	MS
Questions	29	0.51
Judges	7	0.51
Error	203	0.20
Intra-Class Correlation		0.60
Cronbach's Alpha		0.62

Reading Grade 6-8		
	dF	MS
Questions	29	0.69
Judges	7	1.06
Error	203	0.20
Intra-Class Correlation		0.68
Cronbach's Alpha		0.71

Calculation for two-way model with both questions and judges random.

Reading Grade 9-12		
	dF	MS
Questions	29	1.01
Judges	7	0.85
Error	203	0.19
Intra-Class Correlation		0.80
Cronbach's Alpha		0.81

Speaking Grade K-1		
	dF	MS
Questions	40	1.52
Judges	7	4.88
Error	280	0.24
Intra-Class Correlation		0.78
Cronbach's Alpha		0.84

Speaking Grade 2-3		
	dF	MS
Questions	40	2.01
Judges	7	4.91
Error	280	0.25
Intra-Class Correlation		0.83
Cronbach's Alpha		0.87

Calculation for two-way model with both questions and judges random.

Speaking Grade 4-5		
	dF	MS
Questions	40	1.49
Judges	7	3.53
Error	280	0.19
Intra-Class Correlation		0.83
Cronbach's Alpha		0.87

Speaking Grade 6-8		
	dF	MS
Questions	40	1.20
Judges	7	7.37
Error	280	0.22
Intra-Class Correlation		0.71
Cronbach's Alpha		0.81

Speaking Grade 9-12		
	dF	MS
Questions	40	1.09
Judges	7	9.19
Error	280	0.23
Intra-Class Correlation		0.65
Cronbach's Alpha		0.79

Calculation for two-way model with both questions and judges random.

Writing Grade K-1		
	dF	MS
Questions	31	2.07
Judges	7	2.59
Error	217	0.21
Intra-Class Correlation		0.87
Cronbach's Alpha		0.90

Writing Grade 2-3		
	dF	MS
Questions	31	1.02
Judges	7	2.64
Error	217	0.29
Intra-Class Correlation		0.67
Cronbach's Alpha		0.71

Writing Grade 4-5		
	dF	MS
Questions	31	1.26
Judges	7	1.09
Error	217	0.15
Intra-Class Correlation		0.86
Cronbach's Alpha		0.88

Calculation for two-way model with both questions and judges random.

Writing Grade 6-8		
	dF	MS
Questions	31	0.82
Judges	7	4.87
Error	217	0.23
Intra-Class Correlation		0.61
Cronbach's Alpha		0.71

Writing Grade 9-12		
	dF	MS
Questions	31	1.23
Judges	7	3.45
Error	217	0.21
Intra-Class Correlation		0.76
Cronbach's Alpha		0.83