# Istation's Indicators of Progress Español

# Technical Manual

Computer Adaptive Testing System for Continuous Progress Monitoring of Reading Growth for Students Pre-K through Grade 3

**Istation**

**Supporting Educators. Empowering Kids.**
**Changing Lives.**

2000 Campbell Centre II
8150 North Central Expressway
Dallas, Texas 75206
866.883.7323

**www.istation.com**

# Table of Contents

# Chapter 1: Introduction

ISIP™, Istation's Indicators of Progress, Español (ISIP Español) is a sophisticated, web-delivered Computer Adaptive Testing (CAT) system that provides Continuous Progress Monitoring (CPM) by frequently assessing and reporting student ability in critical domains of Spanish early reading.

The ISIP Español assessment is based on sound standards for educational testing and is guided by the latest publications used internationally through Early Grade Reading Assessment (EGRA) (Sprenger-Charolles et al., 2000). These foundational bases were used to design the framework utilized for item writing and editing.

Designed for students in Pre-Kindergarten through Grade 3, who are receiving language arts reading instruction in Spanish, ISIP Español provides teachers and other school personnel with easy-to-interpret, web-based reports that detail student strengths and deficits and provide links to teaching resources. Use of this data allows teachers to more easily make informed decisions regarding each student's response to targeted reading instruction and intervention strategies.



ISIP Español provides growth information in the five critical domains of early reading: phonemic awareness, alphabetic knowledge and skills, vocabulary, fluency, and comprehension. It is designed to (a) identify children at risk for reading difficulties, (b) provide automatic continuous progress monitoring of skills that are predictors of later reading success, and (c) provide immediate and automatic linkage of assessment data to student-learning needs, which facilitates differentiated instruction.

ISIP Español has been designed to automatically provide continuous measurement of Pre-Kindergarten through Grade 3 student progress throughout the school year in all the critical areas of early reading, including phonemic awareness, alphabetic knowledge and skills, fluency, vocabulary, and comprehension. This is accomplished through short tests, or "probes," administered at least monthly, that sample critical areas that predict later performance. Assessments are computer–based, and teachers can arrange for entire classrooms to take assessments as part of scheduled computer lab time or individually as part of a workstation rotation conducted in the classroom. The entire assessment battery for any assessment period requires 40 minutes or less. It is feasible to administer ISIP Español assessments to an entire classroom, an entire school, and even an entire district in a single day - given adequate computer resources. Classroom and individual student results are immediately available to teachers, illustrating each student's past and present performance and skill growth. Teachers are alerted when a particular student is not making adequate progress so that the instructional program can be modified before a pattern of failure becomes established.

# The Need to Improve Testing Practices in Bilingual Classrooms

Districts implementing special language programs are required to designate students' academic plans based on their unique needs. Consequently, school districts are currently addressing academic issues in both Spanish and English, utilizing tests results in both languages when students are enrolled in bilingual education classrooms. The current testing practices have proven to be unfavorable to teachers' instructional time as well as the district- and school-funding needs.

Current national discussions regarding academic services for students whose first language is Spanish have found that academic programs suffer from a dearth of assessments that prove to be non-biased and appropriate (O'Hanlon, 2005; Escamilla, 2006 and the National Center for Latino Child and Family Research, 2009).

Bilingual education programs, including Dual language models, are in great need of improving current testing practices. Improved practices should (a) allow teachers to re-direct time to instructional purposes, (b) target funding to other academic needs, and (c) adopt testing tools that are culturally, linguistically, and cognitively appropriate for programs that follow Spanish Language Arts and Reading standards.

Monitoring students' literacy ability and academic growth in each language is necessary to attend to bilingual students' academic needs. Both versions of ISIP, Early Reading and Español, can provide tools for monitoring literacy development in two languages. These tests were built individually, using different items, and based on separate field-test data.

Obtaining data results that are relevant, reliable, and valid improve assessment practices. To be relevant, data must be available on a timely basis and target important skills that are influenced by instruction. To be

reliable, there must be a reasonable degree of confidence in the student scores. To be valid, the skills assessed must provide information that is related to student performance expectations. Hispanic students who rely on their mother language to excel academically have been found to be impacted negatively by results of invalid, biased, or inadequate commonly used assessment practices. As a result, there has been an over-identification and/or under-identification of Spanish speaking students in special education programs (Espinosa & López, 2007). Identifying students' progress toward literacy in the language of instruction will produce more effective identification of true intervention needs and special education cases.

Roseberry-McKibbin and O'Hanlon (2005) reviewed surveys completed by public school speech-language pathologists on service delivery for non-native English speakers from 1990 through 2001 and found that there was a dearth of assessments that proved to be both unbiased and appropriate. Test items were generally outside of the students' cultural knowledge; and therefore unfamiliar to speakers of other languages, resulting in students' inability to demonstrate the skill being tested. Additionally, test norms based on native speakers of English should not be used with individuals whose first language is not English; and those individuals' test results should be interpreted as reflecting, in part, their current level of proficiency rather than their ability, potential, aptitude, etc. (AERA, APA, & National Council on Measurement in Education [NCME]; Standards for Educational and Psychological Testing, 1999). In order to assess Spanish speaking students, tests must be adequately tested for cultural relevance and proper Spanish terminology that avoids regionalisms and colloquial terms. At the same time, items must demonstrate internal consistency when tested and scored with the population they intend to evaluate; in this case, Hispanic students in bilingual education classrooms in the US public education system are targeted.

There are many reasons why a student score at a single point in time under one set of conditions may be inaccurate: confusion, shyness, illness, mood or temperament, communication or language barriers between student and examiner, scoring errors, and inconsistencies in examiner scoring. However, by gathering assessments across multiple time points, student performance is more likely to reflect actual ability. By using the computer, inaccuracies related to human administration errors are also reduced. Additionally, opportunities to retest are plausible and efficient.

# The Need to Link Spanish Early Reading Assessment to Instructional Planning

Instructional time is utilized more effectively when assessment is linked to instruction. Early reading assessments of Spanish literacy development need to (a) identify students at risk for reading difficulties, students that may need extra instruction or intensive intervention if they are to progress toward grade-level standards in reading by year end; (b) monitor student progress for skill growth on a frequent and ongoing basis and identify students that are falling behind; (c) provide information about students who will be helpful

in planning instruction to meet their needs; and (d) assess whether students achieved grade level reading standards at the end of the school year.

A teacher needs to be able to identify students at risk of reading failure and/or struggling to meet end-of-year grade level expectations. These individualized student data support differentiated instruction; therefore, teachers must first have information about the specific needs of each child.

Linking teacher instruction to the results of assessment is promoted by using formative assessments. Following progress through formative assessments needs to occur often enough that teachers may discover when instruction has not been effective in order to make modifications in a timely manner (Crooks, T., 2001). According to current research, the best examples that follow a formative assessment structure are called "Online Formative Assessment" (Gomersall, 2005; Nicol, D.J. & Macfarlane-Dick, D., 2006). It is also envisioned that computer-based formative assessments will play an increasingly important role in learning, with the increased use of banks of question items for the construction and delivery of dynamic, on-demand assessments (Guide to Assessment, Scottish Qualifications Authority; June 2008).

Research suggests that children with different levels of language proficiency who are also developing literacy skills (whether in one language or two) respond successfully to frequent formative assessments. These assessments' results pinpoint skills as they are emerging and provide the best information as to which readers require additional support in specific reading skills (Gersten et al., 2007).

The purpose of formative assessment can be defined as assessment "for learning, rather than of learning" (Stiggins & Chappuis, 2006, p. 10). Equal educational opportunities for emergent readers should offer the use of formative assessments as a necessity, regardless of language of instruction. Formative assessments provide detailed pictures of the abilities that are measured, in order to make modifications to the instruction that is relevant and, in many cases, critical to students' progress. The primary goals of formative assessment are to guide curriculum and teaching strategies. Districts, teachers, and curriculum developers use data to differentiate classroom instruction while monitoring academic progress. It is important to engage in an ongoing process rather than a single test when using formative assessment. Consistent measures of student progress that involve students in the process enable opportunities for both teachers and students to work together toward common goals. Assessment tools that support self-monitoring contribute to engaging students in self-driven progress practices (McManus, 2008).

A systematic and collaborative process that involves self-monitoring and feedback benefits both teachers and students, because it promotes engagement in meta-cognitive processing that informs learning and increases student achievement (Stiggins & Chappuis, 2006). This type of assessment is most useful when (a) it is conducted periodically, (b) it provides information immediately, (c) it is easy and systematic in administration, and (d) it helps gather a more complete picture of each student, including a range of ability to perform an academic task that varies constantly (Gersten et al., 2007). Computer-based evaluations support all four strengths of formative assessment and allow students to self-monitor their progress.

# Continuous Progress Monitoring

ISIP Español grows out of the model of Continuous Progress Monitoring (CPM) called Curriculum-Based Measurement (CBM). Teachers who monitor their students' progress and use this data to inform instructional planning and decision-making have higher student outcomes than those who do not (Conte & Hintze, 2000; Fuchs, Fuchs, Hamlett, & Ferguson, 1992; Mathes, Fuchs, & Roberts, 1998). These teachers also have a more realistic conception of the capabilities of their students than teachers who do not regularly use student data to inform their decisions (Fuchs, Deno, & Mirkin, 1984; Fuchs, Fuchs, Hamlett, & Stecker, 1991; Mathes et al., 1998).

The collection of sufficient, reliable assessment data on a continuous basis is a daunting task for schools and teachers. Screening and inventory tools for Spanish literacy such as the *Tejas LEE®* (Brookes Publishing Co.) and IDEL®: *Indicadores Dinámicos del Èxito en la Lectura* (Good & Kaminski, 2002) use a benchmark or screen schema, in which testers administer assessments three times a year. More frequent continuous progress monitoring is recommended for all low-performing students, but administration is at the discretion of already overburdened schools and teachers.

Districts currently use CBM models to index student progress over time, which in turn can facilitate teachers' formative evaluation of their teaching effectiveness. Research indicates that CBM can accurately, meaningfully, and sensitively describe such progress (Marston, 1989). This is accomplished through the frequent administration of short, equivalent tests sampling all the skills in the curriculum. A student's past, present, and probable future growth is tracked. When students are not making adequate progress, teachers modify their instructional programs. The educational value of CBM would greatly benefit the outcomes of bilingual education programs, and research demonstrates that instructional programs designed with CBM can result in greater student achievement, enhanced teacher decision making, and improved student awareness of learning (e.g., Fuchs, Fuchs, Hamlett, & Stecker, 1991). Thus, CBM represents a logical model for helping bilingual teachers to identify those students for whom the standard curriculum in place in the classroom is not having the desired effects. Once identified, teachers can intervene before failure has already occurred.

Although proven to be a great tool for classroom teachers, CBM has not been as widely embraced as would be hoped and has hardly been recognized in the field of bilingual education. These assessments, even in their handheld versions, require a significant amount of work to be administered individually to each child. The examiners who implement these assessments must also receive extensive training in both the administration and scoring procedures to uphold the reliability of the assessments and avoid scoring errors. Because these assessments are so labor-intensive, they are expensive for school districts to implement. Bilingual education classrooms, already pressed for time to evaluate students' academic and proficiency needs in two languages, are unable to easily accommodate the requirements of CBM implementation. Therefore, it is difficult for bilingual teachers to be able to use CBM models for continuous progress monitoring and validation of test results.

The introduction of handheld technology has allowed for graphing of student results. Assessments like Tejas LEE (Brookes) can be recorded using palm-pilot devices, but information in this format is often not available on a timely basis for total class or whole school results. Additionally, the time needed for one-on-one administration and the need for additional staff to support classroom teachers during testing periods make it difficult to implement with fidelity and consistency.

Computer applications have been found to be reliable means by which to deliver CBM models by applying similar equivalent test sampling with students over time, using the computer platform to deliver the assessments and a program to collect the data, both immediately and over time.

Computerized CBM applications are a logical step in increasing the likelihood that continuous progress monitoring occurs more frequently with monthly or even weekly assessments in both the general education and bilingual education classrooms. Computerized CBM applications have been developed and used successfully in upper grades in mathematics and spelling (Fuchs et al., 1995). Computerized applications save time and money. They eliminate burdensome test administrations and scoring errors by calculating, compiling, and reporting scores. They provide immediate access to student results that can be used to affect instruction. They provide information organized in formats that automatically group children according to risk and recommended instructional levels. Student results are instantly plotted on progress charts with trend lines projecting year-end outcomes based upon growth patterns, eliminating the need for teachers to manually create documentation of results.

# Computer Adaptive Testing

With recent advances in Computer Adaptive Testing (CAT) and computer technology, it is now possible to create CPM assessments that adjust to the actual ability of each child. Thus, CAT replaces the need to create parallel forms. Assessments built on CAT are sometimes referred to as "tailored tests" because the computer selects items for students based on their performance, thus tailoring the assessment to match the performance abilities of the students. This also means that students who are achieving significantly above or below grade expectations can be assessed to more accurately reflect their true abilities.

There are many advantages to using a CAT model rather than a more traditional parallel forms model, as is used in many early-reading instruments. For instance, it is virtually impossible to create alternate forms of any truly parallel assessment. The reliability from form to form will always be somewhat compromised. However, when using a CAT model, it is not necessary for each assessment to be identically difficult to the previous and future assessments. Following a CAT model, each item within the testing battery is assessed to determine how well it discriminates ability among students and how difficult it actually is through a process called Item Response Theory (IRT) work. Once item parameters have been determined, the CAT algorithm can be programmed. Then, using this sophisticated computerized algorithm, the computer selects items based on each student's performance, selecting easier items if previous items are missed and harder items if the student answers correctly. Through this process of selecting items based on student

performance, the computer is able to generate "probes" that have higher reliability than those typically associated with alternate formats and that better reflect each student's true ability.
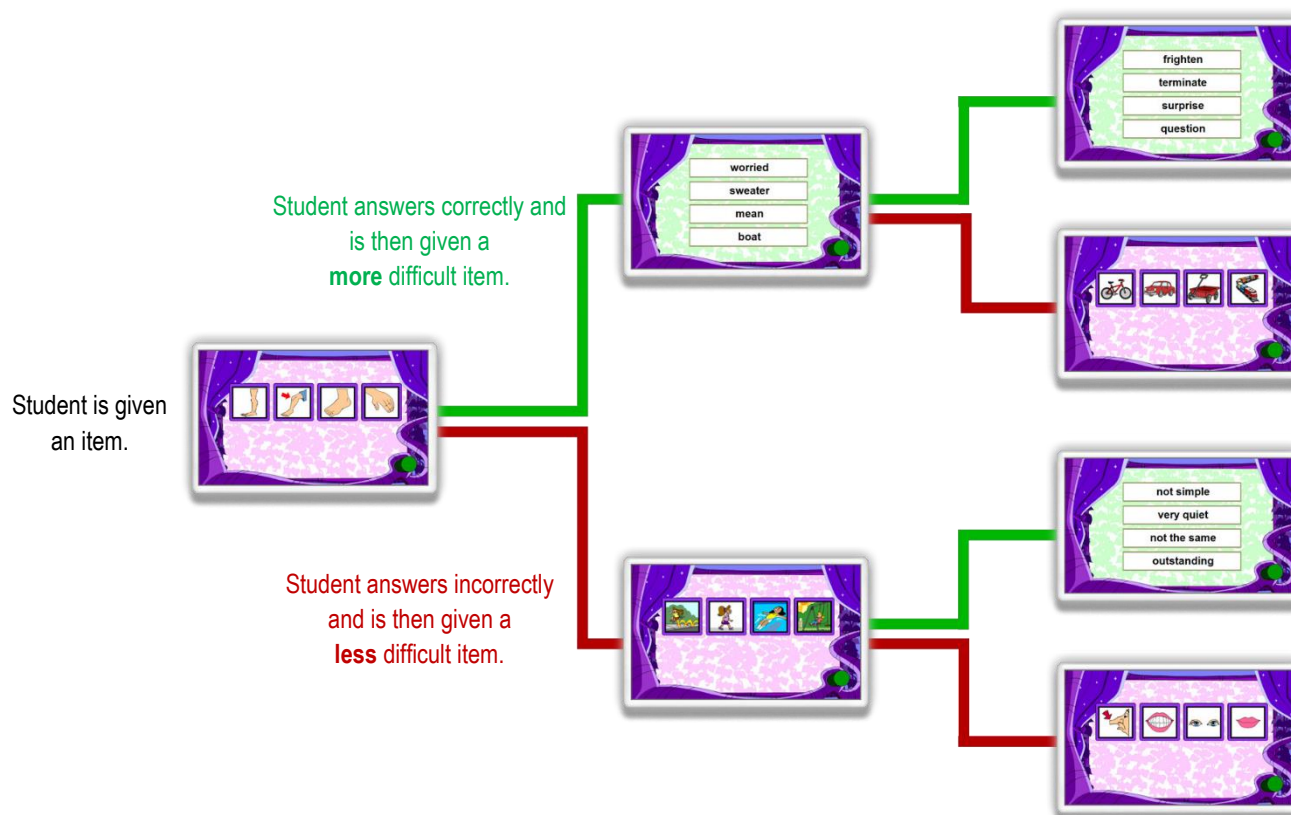


# ISIP Español Domains

ISIP Español uses a CAT algorithm that tailors each assessment to the performance abilities of individual children while measuring progress in the critical early reading skill domains.

The specific domains and the order in which the domains and skills are presented in ISIP Español are based on an analysis of the findings and recommendations of the United States National Reading Panel, European and Latin-American research, including the latest publications from *Marco Común Europeo de Referencia Para Las Lenguas: Aprendizaje, Enseñanza, y Evaluación. [Instituto Cervantes, Ministerio de Educación, Cultura y Deporte, España; 2001].* In addition, the following research findings were considered when developing the assessment blueprint for ISIP Español:

> Es evidente que las prácticas educativas orientadas a exponer al niño a experiencias de comunicación, de intercambio comunicativo, de partir de sus experiencias previas, de tener sentido aquello que se trata de descodificar, etc. es algo que está plenamente justificado y que no importa para ello el contexto idiomático. Sin embargo, los hallazgos

más recientes, desde una perspectiva psicolingüística, ponen de manifiesto que todo ello no sería suficiente ya que el proceso cognitivo de asociación grafía-fonema es un elemento imprescindible cuando se aprende a leer en un sistema alfabético *[Enseñanza de la lectura: de la teoría y la investigación a la práctica educativa. Juan E Jiménez & Isabel O'Shanahan; Universidad de La Laguna, España. Marzo 2008].*

English Translation: It is clear that educational practices designed to expose children to experiences of communication, communicative exchange, use of their prior experiences, making sense of what is referred to as decoding, etc. is something that is fully justified and is not dependent upon the linguistic context. However, recent findings from a psychological perspective indicate that all of this would not be sufficient since the cognitive process of **grapheme/sound correspondence** is an essential element when learning to read in an alphabetic system.

Studies have also demonstrated that the performance of reading words for students learning to read in different linguistic contexts (such as English, French, and Portuguese) is systematically higher in Spanish than in other languages. These studies have also found that knowledge of the complex rules of grapheme/sound correspondence occurs earlier in Spanish than in English, French, or Portuguese. Similarly, Spanish-speaking students reach higher levels of word reading earlier, when compared to students who speak other languages. Such a finding indicates that the appropriate use of the phonological process occurs earlier in Spanish than in English, French, or Portuguese. Findings from these studies have been confirmed most recently with research from Université Paris V - René Descartes in France that compared English-, German-, French-, and Spanish-speaking children learning to read. Conclusions from both studies are disclosed below:

On the one hand, when Grapheme-Phoneme Correspondences (GPC) are almost regular, as in Spanish, reliance on the phonological procedure very often leads to the production of the correct word. Thus, in shallow orthographies, reading skills burst out very rapidly. On the other hand, when the number of inconsistent words is significant, as in English, and to a lesser extent in French, reliance on the GPC procedure sometimes leads to a reading error and reading acquisition is slowed down because of some incoherence between sub-lexical and lexical outputs (From Linguistic Description to Psycholinguistic Processing, Liliane Sprenger-Charolles and Danielle Béchennec, CNRS & Université René Descartes, Paris, 2008).

Cuando se ha comparado el rendimiento en lectura de palabras entre niños que aprenden a leer en distintos contextos idiomáticos (v. gr., inglés, francés y portugués) éste es sistemáticamente más alto en español que en otras lenguas Así, el conocimiento de las reglas de CGF complejas es más temprano en español que en inglés, francés y portugués. Igualmente, los niños españoles alcanzan altos rangos de lectura de palabras muy temprano si los comparamos con otras lenguas, lo que indicaría que la utilización adecuada del procedimiento fonológico ocurre más pronto en español que en inglés,

francés y portugués (Enseñanza de la lectura: de la teoría y la investigación a la práctica educativa. Juan E Jiménez & Isabel O'Shanahan; Universidad de La Laguna, España, 2008).

Additionally, ISIP Español domains are parallel to those in the Early Grade Reading Assessment (EGRA) conducted in Latin American countries like Nicaragua and Guatemala. These research studies have consistently reported the following critical areas for Spanish early literacy development, shown below:

| Domain |
| --- |
| **CONCIENCIA FONOLÓGICA (CF)** <br> **Phonemic Awareness** <br><br> La instrucción en CF consiste en enseñar a los niños a segmentar el lenguaje oral en fonemas sin apoyo de las letras del alfabeto. <br><br> *Phonemic awareness allows students to segment oral language in phonemes without using the letter names.* |
| **CONVERSIÓN GRAFEMA-FONEMA (CGF)** <br> **Grapheme-phoneme conversion** <br><br> La instrucción de reglas de CGF es una forma de enseñar a leer que enfatiza la adquisición de las correspondencias símbolo-sonido. <br><br> *Grapheme-phoneme conversion comprises the reading rules to acquire symbol-sound correspondence.* |
| **VOCABULARIO** <br> **Vocabulary** <br><br> Hay dos tipos de vocabulario: el oral y el escrito. Cuando un lector encuentra una palabra en el texto puede decodificarla, es decir, convertirla en habla. <br><br> *Vocabulary objectives can be divided in two categories: oral and written. Decoding enables conversion of text into a verbal outcome.* |
| **COMPRENSIÓN** <br> **Comprehension** <br><br> Las investigaciones sugieren que la comprensión mejora cuando los alumnos son capaces de relacionar las ideas que están representadas en el texto con su propio conocimiento y experiencias, al igual que las representaciones mentales construidas en la memoria. <br><br> *Comprehension is improved when the students are able to relate ideas from the text to their own background knowledge.* |
| **LECTURA CON FLUIDEZ** <br> **Fluency** <br><br> La fluidez en la lectura es necesaria para la comprensión. Leer con velocidad, precisión, y entonación respetando los signos de puntuación facilita la comprensión del texto. <br><br> *Fluency is necessary to develop correct pace, observing punctuation and thus enhancing reading comprehension.* |

Taking into consideration the studies conducted on the transparency of languages, there are common elements for assessment such as grapheme-phoneme correspondence, word reading, level of vocabulary, reading comprehension of both narrative and expository texts, listening comprehension, and fluency, but there are also elements that are critical to the support of reading development in each particular language (O'Shanahan, Jimenez, 2008). In the case of the Spanish language, writing development and orthography are closely tied to reading. This is the reason why reading in Spanish is referred to as lecto-escritura (Ferreiro, Chile 2002; Bazán, Acuña, Vega, Mexico 2008). This term posits an intrinsic relationship between writing and reading.

There are differences, as well as similarities, in emergent reading and writing behaviors of Spanish-speaking children (Escamilla & Coady, 1998). English writing rubrics cannot help to guide instruction in Spanish. Differences in writing development can impact outcomes in grade-level and state-standards-based assessments. Issues that emerged from this research highlight Spanish primary students' development, in which vowels emerge before consonants; primary students move from strings of letters to invented spelling in Spanish earlier than English speakers do. A writing component is critical for the assessment of emergent literacy skills in Spanish-speaking children. ISIP Español domains include a writing component that aims to address the specific needs of children developing Spanish literacy skills based on the principles of Spanish lecto-escritura.

The growing enrollment of Spanish-speaking students in the Texas public education system clearly reveals the need to develop Spanish assessments that prove to be not only linguistically and culturally appropriate but also aligned with language arts standards and delivered efficiently. ISIP Español allows for more equitable educational opportunities for students, particularly for English-language learners who are Spanish-speaking, the largest growing number of ELLs in Texas. This student population requires qualified instructors and special language programs that support efficient and appropriate ways to assess bilingual education methodologies and special language programs geared toward improved academic achievement for Hispanic students (Guiding Principles for Dual Language Education, Center for Applied linguistics, CAL, 2007; The National Task Force on Early Childhood Education for Hispanics/La Comisión Nacional para la Educación de la Niñez Hispana, 2007; Miller, L.S. & Garcia, E. 2008).

The domains selected for the assessment measures of ISIP Español were established using the literature review described in the section above. Additionally, a number of revisions and feedback were solicited from nationally known researchers in the field of bilingual education, including Dr. Kathy Escamilla from the University of Colorado at Boulder; Dr. Barbara Flores from the University of California, San Bernardino; and Dr. William Pulte from Southern Methodist University, Dallas, Texas.

# CONCIENCIA FONOLÓGICA (CF)
## Phonemic Awareness

Items in this domain intend to evaluate the early literacy skills associated with the mechanics of reading that link to recent findings in neuropsychology studies emerging from post-modern views that impact current students' educational experience (Serie Didáctica de la Lengua y de la Literatura: Catalá Agrás G, Molina H, Bareche Monclús, R., & Editorial Graó, Barcelona, 2007).

Phonemic and syllabic awareness is "the ability to notice, think about, and work with the individual sounds in spoken words" (Armbruster, Lehr, & Osborn, 2003, p. 2). A broader term for this concept is phonological awareness.

> Los modos de representación pre-alfabética se suceden en cierto orden: primero varios modos de representación ajena a toda búsqueda de correspondencia entre la pauta sonora de una emisión y la escritura. Luego modos de representación silábicos (con o sin valor sonoro convencional) y modos de representación silábico-alfabético que preceden la aparición de la escritura. Estos niveles están caracterizados por formas de conceptualización que actúan en un sistema asimilador, absorbiendo la información dada (Alfabetización: teoría y práctica Emilia Ferreiro, 2002).

> English Translation: Prior to any alphabetic representation, there are identifiable audible emissions of written text that find correspondence, beginning with a single sound, followed by syllabic representations (with or without conventional value), and ending with alphabetical syllabic representations that precede emerging writing. These levels of representation are assimilated in conceptual systems that absorb the information given.

The concepts that describe phonological awareness suggest that before children learn to read print, they must understand that words are made up of speech sounds. The United States' National Reading Panel (NRP, 2000) found that children's ability to read words, to comprehend what they read, and to spell is improved with phonemic awareness. Studies of phonemic awareness conducted with Spanish-speaking children have been used recently to confirm the different levels of phonological awareness that are relevant to the Spanish language (Serrano et al., 2009). These levels comprise identification of phonemes in isolation, beginning and ending sounds and syllables, and intra-syllabic sounds, which impact the development of reading skills in unequal levels of relationship (Serrano et al, 2009).

# CONVERSIÓN GRAFEMA-FONEMA (CGF)

Grapheme-Phoneme Conversion

Grapheme-phoneme correspondence is the ability to pair sounds (phonemes) with the letters (graphemes) that represent them. The term phonics is also widely used to describe methods used for teaching children to read and decode words (Abadzi, 2006). According to US base studies, children begin learning to read using phonics, usually around the age of five or six (NRP, 2000). In the case of some alphabetic languages such as Spanish, the orthographic representations of sounds are even simpler because there is nearly a one-to-one correspondence between letter patterns and the sounds that represent them. Even though studies conducted with Spanish-speaking children have not been completed in large quantities, as studies with English-speaking students have, the transparency of the language has been widely researched in the field of linguistics (Wimmer, Mayringer, 2001; Ziegler, Perry, Ma-Wyatt, Ladner, & Körne, 2003), placing languages such as Spanish and French on the transparent side of the languages scale and English and German on the opaque side (Seymour et al., 2003).

> En el caso del castellano, diversos trabajos han mostrado la relevancia de la sílaba, señalando que la conciencia silábica se puede usar como un buen indicador de las habilidades lectoras importantes en una ortografía transparente como el castellano, debido a la correspondencia directa entre grafemas y fonemas (Carrillo, 1994; Jiménez & Ortiz, 2000).

> English Translation: Studies conducted in Spanish language (terminology using "castellano" refers to Spanish) have demonstrated that syllabic awareness is a good predictor of reading skills, due to the direct influence of a transparent orthography over the grapheme-phoneme correspondence.

# Comunicación Escrita

Written Communication

The subtests in this domain measure orthography development and dictation. Orthography measures comprise spelling and use of accent marks, while dictation measures a student's ability to follow grammatically correct sentence structures and emergent syntactic skills.

Spelling refers to the ability to determine the fully specified orthographic representations of words in the language. Knowing the spelling of a word makes the representation of it sturdy and accessible for fluent reading (Ehri, 2000; Snow et al., 2005).

Dictation refers to receptive and productive syntactic skills that have been found to be related to reading ability (Scarborough, 1990). These studies found that there are evident discrepancies between the sentences produced by preschoolers who became poor readers and sentences by those who did not.

According to research by Snow, Burns, and Griffin (1998) with young learners, there are three components of first language ability that have been shown to correlate with later reading development. These components include story recall, lexical skills, and syntactic skills. If a student is able to find the relationship of words inside a sentence after hearing it, he or she should also be able to demonstrate it productively.

# Vocabulario
## Vocabulary

Current scientific research overwhelmingly supports the idea that a dearth of vocabulary impedes reading comprehension and a broad vocabulary increases comprehension and facilitates further learning (Hirsch Jr., 2003). Adequate reading comprehension has been correlated to the number of words in a text that a reader already knows. Experts consider that a good reader knows between 90 and 95 percent of the words in a text (Nagy & Scott, 2000).

Oral language vocabulary refers in general to "the words we must know to communicate effectively" (Armbruster, Lehr, & Osborn, 2003). On the other hand, reading vocabulary refers to words that a student needs to know in order to be able to understand what is read. The development of oral language proficiency—both productive (speaking) and receptive (listening)— is key to literacy growth. Furthermore, there is a rich vein of literature that suggests that vocabulary is an important precursor to literacy (see Scarborough, 2005 for a summary of this literature).

Reading vocabulary demands knowledge of words and their relationships, as well as the ability to extract meaning from words in context. The percentage of words that a reader understands when reading a text either causes the reader to miss the gist of the reading or allows the reader to get a good idea of what is being said and; therefore, to make correct inferences in order to determine the meaning of any unfamiliar words (Hirsch Jr., 2003). Bilingual education settings need to take advantage of the vocabulary that students acquire in their native language. First language vocabulary ability has been shown to correlate with later Academic Language development. "Academic language refers to the decontextualized, cognitively challenging language used not only in school, but also in business, politics, science, and journalism, and so forth. In the classroom, it means the ability to understand story problems, write book reports, and read complex texts" (Crawford & Krashen, 2007).

# Comprensión
## Comprehension
### Comprensión auditiva | Listening Comprehension

Items in this domain intend to evaluate students' listening comprehension proficiency levels as indicators of foundational early literacy skills.

En todo proceso auditivo, para poder ser asimilada; la información debe ser integrada a un sistema previamente construido (o un sistema en proceso de construcción). No es la información de por sí, la que crea conocimiento, el conocimiento es el producto de la construcción de un sujeto cognoscente (Alfabetización: teoría y práctica Emilia Ferreiro, 2002).

English Translation: Listening proficiency assimilates new information based on existing constructed systems that integrate new information. The information alone does not constitute knowledge. Knowledge is a product constructed through a cognizant subject.

Listening Comprehension refers to the ability to effectively receive auditory input (receptive skills) in order to understand the information that was said. In bilingual education settings, listening skills developed in the native language benefit second language acquisition. In fact, listening often develops before the productive skill of speaking, so students may depend on listening skills while they are silent for an extended period during second language acquisition (Díaz-Rico, 2008; Crawford & Krashen, 2007).

Students developing early reading skills continue to strengthen listening comprehension abilities while they are able to read silently. Reading difficulties have been found to reach the same levels that the listening comprehension disorders do. Even so, these two may not be evident simultaneously or may not be mutually exclusive (Junqué I Plaja et al., 2004).

Latest publications from the United States National Early Reading Panel (NEPL) and the National Institute for Literacy identify listening comprehension as one of the key foundational skills for later reading achievement.

**Comprensión de lectura | Reading Comprehension**

This domain parallels reading comprehension measures, as determined by each state's criterion reference tests, by incorporating the same types of questions. Comprehension questions are aligned to fiction and non-fiction objectives, such as main idea, summarization, drawing conclusions, and predicting as it applies to explicit and implicit cues.

Comprehension is defined as the process through which meaning is extracted from the written language. Comprehension measures can be classified in two types: (a) *literal* comprehension, which focuses on the recognition or retrieval of primary details, main ideas, sequences, or cause-effect patterns from the information that is explicit in the text, and (b) *inferential* comprehension, which requires establishing logical connections and relationships among facts in texts, thus allowing readers to deduce events, make generalizations, interpret facts, and relate previous knowledge or personal experiences to information implicit in the text (Camba, 2006).

# Lectura con fluidez
## Fluency

This domain allows students to read silently, indicating the number of words that they are able to read per minute while, at the same time, demonstrating accuracy.

Fluency is the ability to read text correctly and with appropriate pace. Reading with fluency requires accuracy and speed. Therefore, a fluent reader is able to read aloud effortlessly using a natural expression, as if speaking; whereas a reader who has not yet developed fluency reads slowly, word by word, with inconsistency and constant interruption. Accuracy refers to the percentage of words read correctly, and speed is the number of words read per minute. In order to measure fluency, a calculation of the number of words read correctly in one minute yields a fluency rate. Recent studies of fluency outcomes conclude that attention to connections among ideas and between these ideas, as they relate to background knowledge, are more characteristic of fluent readers than non-fluent readers (Armbruster, Lehr, & Osborn, 2003; NRP, 2000).

# ISIP Español Items

The purpose of the ISIP Español Item Bank is to support teachers' instructional decisions. Specifically, the item bank is designed to serve as a computerized adaptive universal screening and progress monitoring assessment system. By administering this assessment system, teachers and administrators can use the results to answer two questions: (1) are students in grades Pre-K through Grade 3 at risk of failing reading in Spanish, and (2) what is the degree of intensity of instructional support students need to be successful readers? Because the assessment is designed to be administered, these decisions can be applied over the course of the school year.

The United States has not adopted a set of national common Spanish Language Arts and Reading (SLAR) standards. The state of Texas requires the implementation of TEKS for Spanish Language Arts Reading (SLAR) and English as a Second Language (ESL) Elementary Standards under Texas Education Code 128.10. These standards were used to develop a "hybrid" version of Spanish standards combined with selected states and countries (i.e., California, Texas, Puerto Rico, WIDA consortium, Colombia, Mexico and Spain). The combined standards were utilized to determine end of year expectations for Pre-Kindergarten through Grade 3. These standards were revised by national experts who collaborated with istation as members of the ISIP Español Advisory Council. Once the standards and benchmarks were developed, a blueprint of skills to be assessed was determined based on an analysis of existing Spanish literacy assessments. Input and suggestions were then sought from the group of experts in Spanish language content, linguistics, and language acquisition that comprised the Advisory Council. Members of this council are listed below:

**Dr. Iliana Alanis**
Assistant Professor
University of Texas at San Antonio
Department of Interdisciplinary Learning and Teaching

**Dr. Igone Arteagoitia**
Research Associate
Center for Applied Linguistics (CAL)
Washington DC

**Dr. Kathy Escamilla**
Professor of Education
School of Education
University of Colorado at Boulder

**Dr. Gilda Evans**
Retired Assistant Superintendent over the Multi-language Department of Dallas ISD
Current Vice-president of Bilingual Education Association of the Metroplex (BEAM)

**Dr. Eugene E. Garcia**
Professor and Vice President, Education Partnerships
Arizona State University

**Kathleen Leos**
Former Assistant Deputy Secretary and Director to the US Department of Education's Office of English
Language Acquisition (OELA) President and Co-Founder of Global Institute for Language and Literacy
Development (GILD)

**Dr. William Pulte**
Associate Professor, Director of Bilingual Education Programs
Simmons School of Education
Southern Methodist University

**Dr. Luis Rosado**
Professor – College of Education
Director – Center for Bilingual Education
University of Texas at Arlington

**Lisa Saavedra**
Vice President and Co-Founder of Global Institute for Language and Literacy Development (GILD)
Former Bureau Chief for the Bureau of Academic Achievement through Language
Acquisition for the Florida Department of Education

**Dr. Annette Torres-Elias**
Assistant Professor of Education
School of Education
Texas Wesleyan University

A Texas-based editorial firm, Tri-Lin, devoted to assessment, development, and special education services focusing on the Spanish and bilingual educational community, was contracted by Istation to write more than 5,000 items that make up the item bank for forms for ISIP Español.

Items were written by the Spanish test development staff. The items were required to follow specific rules for each domain, based on the ISIP Español assessment blueprint. The multiple-choice answer options were also driven by elimination rules specifications (rules for item creation and answer-choice elimination are available upon request). All items were originally written for use in this assessment; no items were translated or derived from any assessment delivered in English. In addition, all items underwent comprehensive analysis to ensure that no items contained linguistic or cultural bias and that all were age- and grade-level appropriate. Thus, the range of item types was extended to include items with difficulties as low as the end of Pre-K and as high as Grade 5/6. Additionally, items were developed within each domain to represent easy, moderate, and hard items for each grade. This wide range of items make ISIP Español an appropriate measure for the full range of students, including students with special needs or who struggle and students who are high-achieving or gifted. While ultimately the IRT calibration work identified the difficulty of each item, the team was assured of having items representing the full continuum of achievement for each domain.

The use of CAT algorithms also creates efficiencies in test administration. The adaptive item algorithm allows the computer to adjust item difficulty while the child is taking the test, quickly zeroing in on ability level. Thus, the use of CAT algorithms reduces the amount of time necessary to accurately determine student ability.

# Accuracy and Fluency

Within ISIP Español, each subtest has both an accuracy component and a fluency component. Assessments that measure a student's accuracy and speed in performing a skill have long been studied by researchers. Such fluency-based assessments have been proven to be efficient, reliable, and valid indicators of reading success (Fuchs et al. 2001; Good, Simmons, & Kame'enui, 2001). Fluency in cognitive processes is seen as a proxy for learning, such that as students learn a skill, the proficiency with which they perform the skill indicates how well they know or have learned the skill. In order to be fluent at higher-level processes of reading connected text, a student will also need to be fluent with foundational skills.

Because each of the subtests has a fluency component, the tests are brief. This makes it feasible to administer the subtests on a large scale with minimal disruption of instructional time. Numerous items are

available for each subtest, making the subtests repeatable throughout the school year with many alternative forms.

## Teacher Friendly

ISIP Español is teacher friendly. The assessment is computer based, requires little administration effort, and requires no teacher/examiner testing or manual scoring. Teachers monitor student performance during assessment periods to ensure result reliability. In particular, teachers are alerted to observe specific students identified by ISIP Español as experiencing difficulties as they complete ISIP Español. They subsequently review student results to validate outcomes. For students whose skills may be a concern, based upon performance level, teachers may easily validate student results by re-administering the entire ISIP Español battery or individual skill assessments.

## Child Friendly

ISIP Español is also child friendly. Each assessment session feels to a child like he or she is playing a fast-paced computer game called "A ver cuánto sabes" (Show what you know). In the beginning of the session, an animated owl enters the screen (named Don Buhiermo for Búho and Guillermo) that acts as a game show announcer and invites children to participate by saying, "¡Bienvenidos! En este juego vas a demostrar que ¡si puedes!" (It's time to show that you can do it!) The owl helps the children to understand the game rules, and then the assessment begins. At the end of each assessment, children see an animated graph of their progress. Each activity proceeds in a similar fashion.

# ISIP Español Subtests

ISIP Español measures progress in each critical component of reading instruction in a manner appropriate to the underlying domain.  There are a total of six subtests that align to the critical domains of Spanish reading, as shown in the table below. Of these subtests, four are built using a CAT algorithm, while two use parallel forms. Subtests that tailor items using CAT include Destreza fonológica y fonética, Vocabulario, Comprensión de lectura, and Comunicación escrita. Lectura con fluidez and Comprensión auditiva are designed as parallel forms that measure end of grade level expectations.

| Domain | Subtest |
|--------|---------|
| CONCIENCIA FONOLÓGICA<br>*Phonemic Awareness* | Destreza fonológica y fonética<br>*Phonemic and Phonological Awareness* |
| CONVERSIÓN GRAFEMA-FONEMA<br>*Grapheme-phoneme conversion* | Comunicación escrita<br>*Written Communication* |
| VOCABULARIO<br>Vocabulary | Vocabulario<br>*Vocabulary* |
| COMPRENSIÓN<br>*Comprehension* | Comprensión auditiva<br>*Listening Comprehension*<br><br>Comprensión de lectura<br>*Reading Comprehension* |
| LECTURA CON FLUIDEZ<br>*Fluency* | Lectura con fluidez<br>*Text Fluency* |

# ISIP Español Administration Format

ISIP Español is presented to students using a game-like format. Students are never told that they are being given a test. Instead, they are told that they are playing a game called "A ver cuánto sabes" (Show What You Know).



The first time a student takes ISIP Español, the computer will administer items that are defaulted based on the student's grade, unless the default setting is changed intentionally, as may be appropriate in special education settings. From the very first item, however, the CAT engine immediately begins to tailor the test to

the individual student.  As a result, students will only be administered subtests that are appropriate for their performance abilities.  Within a classroom, students may have some variation in the exact subtest they are administered. However, scores reflect these differences (explained below). For example, students whose performance scores indicate that they are not yet reading words will not be asked to read connected text. Likewise, students whose performance scores indicate that they read connected text fluently, and with comprehension, will not be asked to complete letter knowledge and phonemic awareness tasks.

Listening Comprehension is administered only in Pre-K and Kindergarten. In Grade 1, Text Fluency is administered only after students obtain a high enough overall reading score to suggest that they can handle the task. Lectura con fluidez is administered to all students, beginning in Grade 2.

The table below presents the defaults for subtest administration for each grade level.

| Grade | Subtest |
|---|---|
| Pre-Kindergarten | Destreza fonológica y fonética<br>Vocabulario<br>Comprensión auditiva |
| Kindergarten | Destreza fonológica y fonética<br>Vocabulario<br>Comprensión auditiva<br>Comunicación escrita |
| 1st Grade | Destreza fonológica y fonética<br>Vocabulario<br>Comprensión de lectura<br>Lectura con fluidez<br>Comunicación escrita |
| 2nd and 3rd Grade | Destreza fonológica y fonética<br>Vocabulario<br>Comprensión de lectura<br>Lectura con fluidez<br>Comunicación escrita |

## Rationale for Subtest Defaults by Grade

Children acquire the skills that they need to become proficient readers during the first years of school. These skills may be introduced and monitored separately, but teachers need to practice integrating these skills during daily reading routines as often and quickly as possible. Critical early reading skills are emphasized according to the grade level and the developmental stage of the child; however, daily practice of identified critical domains such as phonemic awareness, vocabulary, comprehension, and fluency is highly desirable (Vaughn & Linan-Thompson, 2004).

Based on research findings from the National Reading Panel (NRP, 2000), instruction in phonemic awareness is emphasized in Kindergarten and is recommended for about 15 minutes a day (Vaughn & Linan-Thompson, 2004). The teaching of phonemic awareness has expanded through many countries (particularly Spanish-speaking countries) by introducing instructional methodologies that allow students to manipulate phonemic and syllabic sounds of spoken words. Such methods involve teaching and practicing blending, segmentation of sounds, and identification of sounds that represent sounds in speech with or without the use of letters (NRP, 2000).

Early literacy instruction that incorporates decoding and word study provides a strong foundation for emergent literacy. Therefore, related skills such as grapheme-phoneme correspondence and print awareness can be introduced as early as Kindergarten.

Based on transparency of language studies that place Spanish as a language with shallow orthography (Wimmer & Mayringer, 2001; Ziegler, Perry, Ma-Wyatt, Ladner, & Körne, 2003; Sprenger-Charolles, Béchennec, 2008), children learn these skills rather quickly, and it is important to integrate these emergent reading skills with reading comprehension questioning. Taylor et al., (2002) found that children in first grade grew more in comprehension and fluency when their teachers asked more high-level questions.

Once students acquire a solid foundation in word recognition and decoding, fluency instruction should be emphasized. English literacy studies have shown that this usually begins during the second semester of first grade (Vaughn & Linan-Thompson, 2004). For fluency instruction in Spanish, the transparency of the language must be taken into consideration. Fluency development may begin as early as Kindergarten, based on the fact that Spanish phoneme combinations can be represented in only 45 variations (Sprenger-Charolles, Béchennec, 2008). Spanish grapho-phonemic conversions also make writing development equally accessible and rapidly attained. Tasks that require knowledge of grapheme-phoneme and symbol sound correspondence can be evaluated through students' spelling and dictation performance. Studies which examine the productive and receptive syntactic skills of Kindergarteners also show correlations with success in reading (Ballantyne, Sanderman, D'Emilio, & McLaughlin, 2008). Research has also shown that learning to spell and learning to read rely on much of the same underlying knowledge, such as the relationships between letters, letter units, and sounds. Knowing the spelling of a word makes the representation of it sturdy and accessible for fluent reading (Ehri, 2000 ; Snow et al., 2005).

Teaching vocabulary supports reading comprehension and increases speed, thus improving fluency. First and second grade vocabulary teaching impacts reading development, and the instructional methodology must incorporate familiar language in order to take advantage of the word superiority effect (Cattell, 1986). Students need to be able to identify with and connect to the reading material in order to be interested in it. Linguistic considerations are important, but cultural relevance may be a determining factor in assessing students' reading success (Skiba, Simmons, Ritter, Kohler, & Wu, 2003). When stories are interesting and written in simple language, they are very likely to encourage struggling students to persevere (Abadzi, 2006).

Successful reading development is also associated with small-group instruction. Hierarchical Linear Modeling (HLM) has been used to analyze classroom variables and compare outcomes. These observations confirm that teachers who engage often in small-group instruction have students who demonstrate more gains in fluency, comprehension, and vocabulary.

The subtests and domains of ISIP Español have been developed based on the sequence and frequency of the critical areas of reading identified in Kindergarten through Grade 3. Additionally, ISIP Español is equipped with research-based downloadable teacher resources that support formative assessment and small-group instruction. ISIP Español items were written for Hispanic students in bilingual education programs whose linguistic and cultural aspects were taken into consideration.

Measures for each subtest are described below:

**Beginning Sound**
Beginning Sound is a measure of phonemic awareness that assesses a child's ability to recognize the initial sound in an orally presented word. This skill is tested in Kindergarten and Grade 1, as aligned to the Spanish Language Arts and Reading Texas Essential Knowledge and Skills (SLAR TEKS) standards, and the resulting score is factored in with other skills under the same domain: Destreza fonológica y fonética.

**Blending**
Blending is a measure of phonemic awareness that assesses a student's ability to blend syllables and phonemes that make up spoken words. This skill is tested in Kindergarten and Grade 1, as aligned to SLAR standards, and the resulting score is factored in with other skills under the same domain: *Destreza fonológica y fonética.*

**Letter Sound**
Letter Sound is a measure of alphabetic principle that assesses how many letter sounds a student can correctly identify. Item selections for this portion of the assessment represent a combination of both upper and lower case letters, including vowels and consonants. This skill is assessed in Kindergarten and Grade 1, as aligned to SLAR standards, and the resulting score is factored in with other skills under the same domain: Destreza fonológica y fonética.

**Symbol Sound**
Symbol Sound is a measure of symbol conversion based on auditory input that combines letter units (syllables), as opposed to single phonemes (letters). Item selections for this portion of the assessment include the following syllable types: opened (CV), closed (CVC), consonant combination (CCV), and vowel combination (VV), and items are presented as they apply to each grade level expectation in Kindergarten and Grade 1. The resulting score is factored in with other skills under the same domain: Destreza fonológica y fonética.

## Vocabulary

Vocabulary is a measure of a student's knowledge of two types of vocabulary words: (1) oral vocabulary, or "common" words, which are primarily used in daily social interactions according to each developmental age (grade-level appropriate) and (2) academic vocabulary, or "meaning" words, which are frequently encountered in text but not typically used in daily conversation (Beck, McKeown, & Kucan, 2002 ). In particular, this second evaluation target contains items that were developed to assess students' knowledge of specific Spanish language elements that support understanding of meaning, such as word association, word derivatives, word roots (prefixes/suffixes), and synonyms. These two types of vocabulary words are evaluated separately in all grade levels, Kindergarten through Grade 3. The two scores (oral and academic vocabulary) are combined into a single score and reported as a vocabulary result in this domain: Vocabulario.

## Listening Comprehension

Listening Comprehension is a measure of a student's ability to listen and retain enough information in his or her working memory to be able to recall simple facts. This skill is tested in Kindergarten and Grade 1, as aligned to SLAR standards. It presents one narrative passage and one expository passage, followed by short answer questions that use a similar pattern to the reading comprehension battery. The resulting score is reported independently under the domain with the same name: Comprensión auditiva.

## Comprehension

Comprehension is a measure of a student's ability to read and understand grade level appropriate narrative and expository passages. According to the NRP research (NRP, 2000), text comprehension is a complex cognitive process that incorporates all foundational reading areas, including vocabulary and fluency. In order to assess a student's ability to comprehend the passages, this type of evaluation requires an intentional and thoughtful interaction between the student and the computer screen where the passages are presented. Students in Kindergarten are able to listen to the answer choices associated with a picture before they select their answer. Students in Grades 1 through Grade 3 are able to apply reading comprehension strategies to enhance understanding. The passage appears on the screen, and the student prompts the computer to begin the questions. Once the questions begin, the passage moves to the left side of the screen, and each question changes after 50 seconds to avoid inactivity. The questioning design is similar to the multiple-choice patterns used in state criterion referenced tests that combine explicit and implicit answers as they apply to grade-level requirements aligned to SLAR standards. The resulting score is reported independently under the domain with the same name: Comprensión de lectura.

## Fluency

Fluency is a measure of a student's ability to read fluently with comprehension. This subtest is constructed in a very different manner than others, using grade-level, culturally-appropriate passages. Each of these passages was carefully written to conform to specific word level features, follow linear story grammar structure, and have readability according to a commonly accepted readability formula for end-of-grade level expectations as it applies to Spanish fluency. (It uses the middle to mid-high ranges of the Spanish literacy fluency chart, in Table 3). In order to assess text reading on the computer, a maze task is utilized in which

every seventh word is left blank, with a three-word menu of choices to complete the sentence. This task has been shown to be highly correlated to measures of both fluency and comprehension, and it has high reliability and concurrent validity (Espin, Deno, Maruyama, & Cohen, 1989; Fuchs & Fuchs, 1990; Jenkins, Pious, & Jewell 1990; Shinn, Good, Knurson, Tilly, & Collins, 1992). As opposed to fluency measures for Grade 1, in which the teacher relies on oral/observable measures, students in Grade 2 and 3 would be more accurately assessed using tools that register receptive reading skills. Fluency is tested in Grade 2 and 3, and the resulting score is reported independently under the domain with the same name: Lectura con fluidez.

### Spelling and Dictation

Spelling and Dictation is a measure designed to determine if students are developing fully specified orthographic representations of words. Items were designed following different rules, depending on the grade level assessed. For Grade 1 students, a word is given and an array of syllables appears on the screen, with which the student spells the word. Grade 2 and 3students use individual letters to spell the words. Items for the dictation subtest follow a similar functionality. Students choose, from a word bank, the necessary word to complete a sentence. These items have been carefully constructed to move from easier to harder, using a sequence of difficulty aligned to SLAR standards. Additionally, students in Grade 1 through Grade 3 are required to select correctly spelled words that exemplify commonly used accented patterns for word categories such as palabras llanas, graves, agudas, and esdrújulas. The scores for each subtest are factored together into a single score under the domain Comunicación escrita.

# Description of Each Subtest

## Destreza fonológica y fonética

The Destreza fonológica y fonética subtest is comprised of several types of items:

**Conversión grafema-fonema** items measure the students' ability to identify symbols that correspond to specific sounds of the Spanish language: letras *(letters),* sílabas *(syllables),* combinaciones vocálicas *(vowel combinations),* grupos consonánticos *(consonant clusters),* and palabras *(words).* The computer presents items representing various upper- and lower- case letter combinations. Four boxes appear on the screen, and only one choice contains the correct answer for each item. The narrator asks students to click on a particular grapheme (letter, syllable, etc.) that represents a sound produced orally by the narrator.

Screenshot examples:



**Conciencia fonética y silábica** (*phonemic and syllabic awareness*) items measure the students' ability to identify single sounds (letter or syllable) in grade-level appropriate words. The level of difficulty adapts with the student response. Students identify beginning sounds and use syllables or letters to find words.

Screenshot examples:
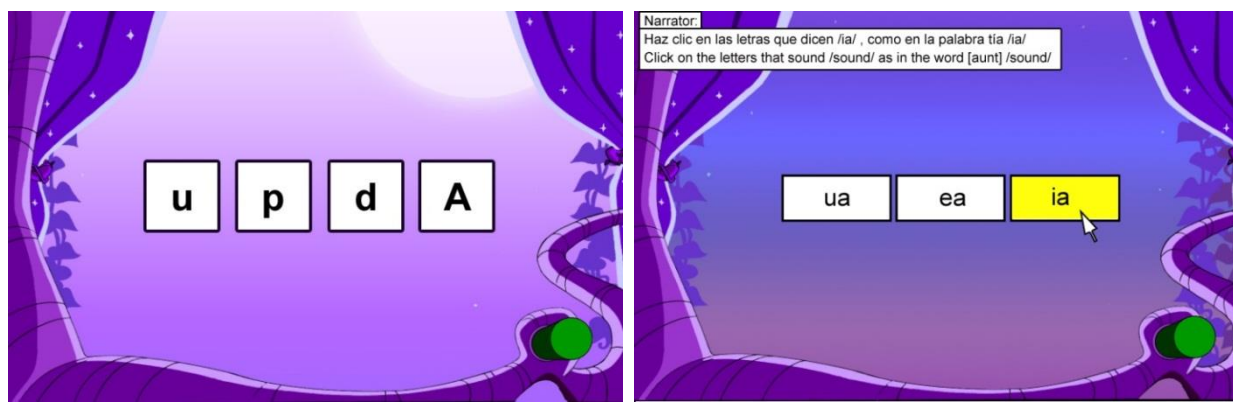


**Unión de sonidos** (*blending)* and **sonido inicial** (*beginning sounds)* are items presented independently. First students find the beginning sound of a word following the narrator instructions. The name of each picture is given as these appear on the screen. Each box is highlighted while students are asked to click on the picture that has the same beginning sound as the sound is produced orally by the narrator. For blending items, a box appears in the middle of the screen containing an animated side view of a head that pronounces the sounds. Once the word is said by pronouncing each phoneme or syllable, the student is asked to click on the picture that shows the word that has been spoken using only sounds.

Screenshot examples:
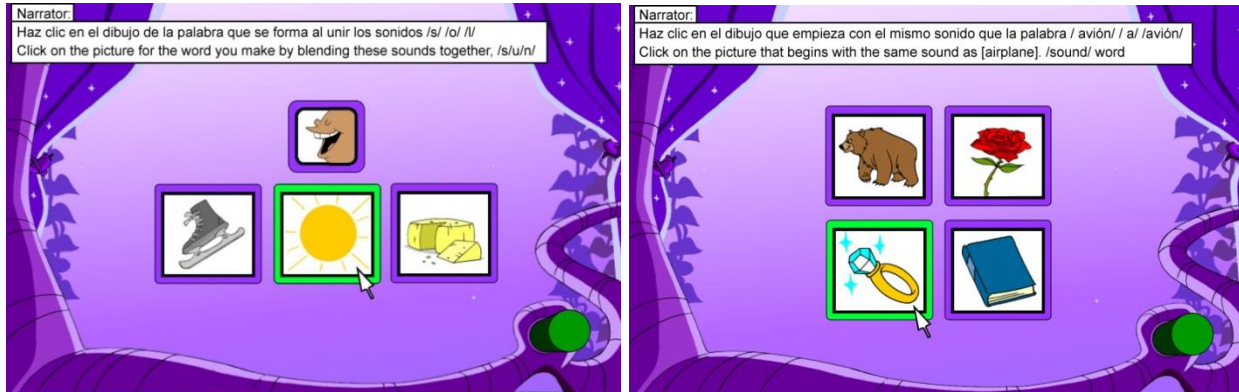


## Comprensión auditiva

**Comprensión auditiva** (*Listening Comprehension*) is a subtest used to evaluate children's ability to listen, understand, and answer questions related to a story that is presented orally. In this activity, a picture related to a short story appears on the screen. The narrator reads aloud with no text present on the screen. The narrator then asks the student a question related to the story. From the four pictures that appear on the screen, the student chooses the one that best answers the question.



## Vocabulario

The Vocabulario subtest is comprised of several types of items:

**Vocabulario de lenguaje oral** (*Oral Vocabulary*) items measure a student's vocabulary knowledge. In this subtest, four pictures appear on the screen. The narrator asks the student to identify the picture that best illustrates the word spoken orally.
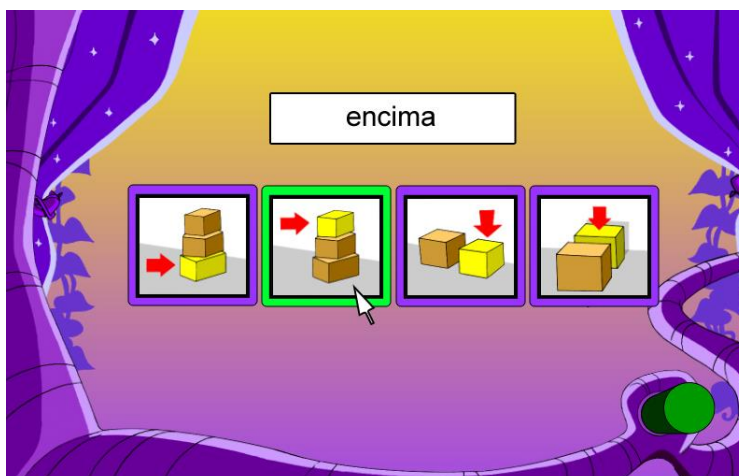
**Vocabulario para lectura y escritura** (*Reading Vocabulary*). For these items, a combination of word strategies (i.e., knowledge of roots, prefixes, and suffixes) is assessed using both pictures and words that appear in sets of four on the screen. The q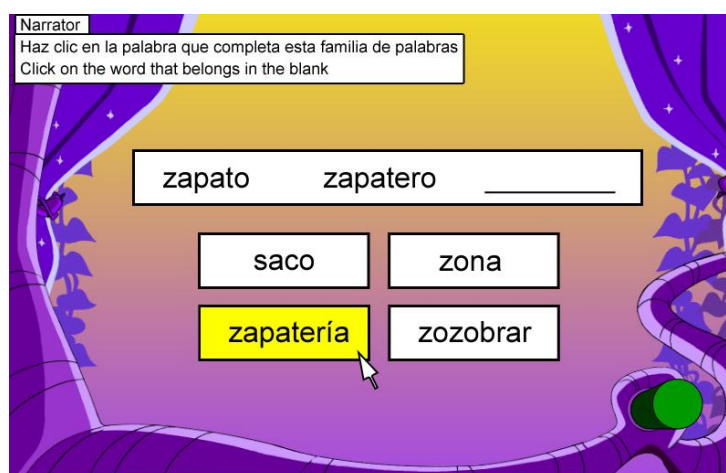uestions spoken by the narrator cover word knowledge such as familias de palabras, clasificación de palabras, sinónimos, etc. (*synonyms, classification of words, derivatives, etc*.) After instructions are given, the student is asked to identify each word accordingly.



# Comprensión de lectura

**Comprensión de lectura** (*Reading Comprehension).* In this subtest, students are assessed on entendimiento de lo leído (*their ability to read and understand sentences and texts*). This is accomplished using evidential/inferential question patterns to evaluate both narrative and expository texts. The item types that measure reading comprehension and thinking skills are linked to criterion-referenced tests. In this task, a passage appears on the screen. The student indicates when he or she is finished reading by clicking on a green button. After this button is clicked, questions populate on the right side of the screen. The student is able to read the text as often as needed while choosing an answer from among four choices. Kindergarten students select from pictures that represent each answer choice, and these are read by the narrator and repeated as needed.

# Lectura con fluidez

**Lectura con fluidez** *(Text Fluency)* is a subtest constructed in a very different manner than the other subtests. Students are assessed on their skill in reading text with meaning in a specified period of time. In order to assess text reading on the computer, a maze task is utilized in which every seventh word of a grade-level story is left blank from the text. The student is given three choices for each blank from which to choose the word that makes the most sense to complete the sentence. It is the student's job to read the text and select the correct maze responses in two and one-half minutes. This task has been shown to be highly correlated to measures of both leer un texto and tener precisión *(fluency and accuracy)*.

# Comunicación escrita

**Ortografía y acentuación de palabras** *(Spelling and accent marks)* is a subtest that determines if students are developing fully specified orthographic representations of words. For each item, an array of letters appears on the screen and the computer asks the student to spell a specific word using those letters and their proper tildes *(accent marks)*. In Grade 1, the same objective is achieved using syllables to write the word said by the narrator. The student then spells the word by clicking on each letter/ syllable. As each letter/syllable is selected, the word is formed on a line that is directly above the letter array.
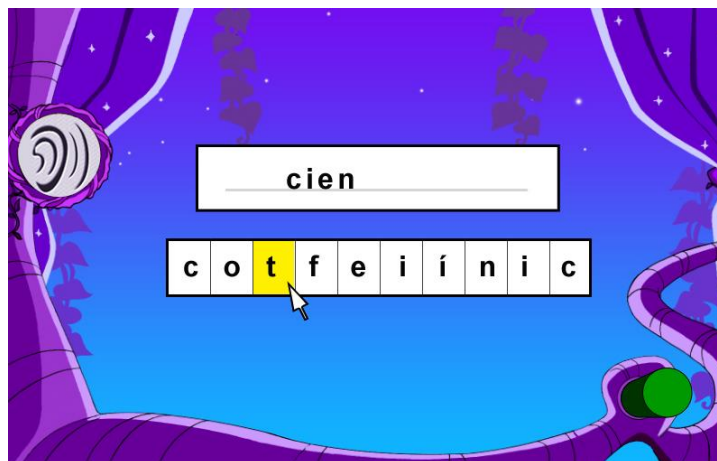


**Dictado** (*Dictation*) the items in this subtest are designed to determine if students are using correct sentence structure: sujeto + verbo + predicado (*subject + verb + predicate*). For each item, an array of words appears on the screen and the computer asks the student to put together a specific sentence using the words available.
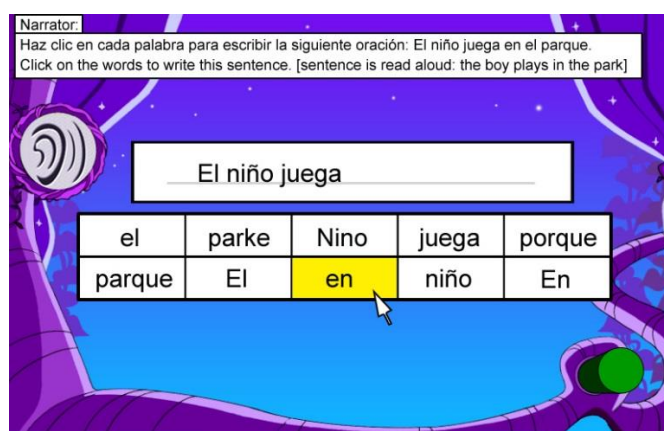
# The ISIP Espanol Link to Instructional Planning

ISIP Español provides continuous assessment results that can be used in recursive assessment-instructional decision loops. Initially ISIP Español identifies students in need of support. If validation of student results is needed re-administering the assessments can increase the reliability of the scores. The technology underlying ISIP Español delivers real-time reports on student progress immediately upon assessment completion. This data facilitates the evaluation of curriculum and instructional plans. Assessment reports automatically group students according to level of support needed as well as skill needs. Data are provided in both graphical and detailed numerical formats on every measure and at every level of a district's reporting hierarchy. Reports provide summary and skill information for the current and prior assessment periods that can be used to evaluate curriculum, plan instruction and support, and manage resources.

At each assessment period, ISIP Español automatically alerts teachers to children in need of instructional support through the "Priority Report." Students are grouped according to instructional level and skill need. Links are provided to teacher-directed plans for each instructional level and skill category. There are downloadable lessons and materials appropriate for each level of instruction.

A complete history of Priority Report notifications, including those from the current year and all prior years, is maintained for each child. On the report, teachers may acknowledge that suggested interventions have been provided. A record of these interventions is maintained with the student history as an Intervention Audit Trail. This history can be used for special education Individual Education Plans (IEPs) and in Response to Intervention (RTI) or other models of instruction that require modifications of a student's instructional plan.

In addition to the recommended activities, reading coaches, and teachers have access to an entire library of teacher-directed lessons and support materials at www.istation.com. These downloadable, printable lessons support small-roup instruction through scripted lessons. These teacher-directed lessons are based on student individualized needs per the Priority Report. As the lessons are taught, teachers document intervention delivery on the Priority Report. This provides a visual reference of teacher intervention and its effectiveness. The ease of identification of skill needs and readily available lessons facilitates intervention and puts instructional time back in the classroom.

All student information is automatically available by demographic classification as well as specially designated subgroups of students who need to be monitored.

A year-to-year history of ISIP Español results is available. Administrators, principals, and teachers may use their reports to evaluate and modify curriculum. Interventions, AYP progress, effectiveness of professional development, and personnel performance may also be correlated to the growth model depicted from the reports.

# Chapter 2: ISIP Español Administration

The specific directions for administering each of the subtests are presented in this section. These directions represent standardized procedures that, when followed, will help to ensure both test reliability and test validity from classroom to classroom, teacher to teacher, and school to school. Information that describes the students' experience in each subtest, as well as information available to Administrators, Principals, and Teachers after completion of the assessments, is also included.

## Teacher and Lab Manager Preparation

Prior to the Initial Administration of ISIP Español:

1.  Enter students' names and their unique District ID numbers at www.istation.com. Student ID numbers are encrypted to prevent interception or identification of student information.

2.  After creating and processing your student accounts, print the student login cards. Place the login cards in a file box near the computers in the lab and/or classroom. Login cards should be easily accessible to students.

3.  Inspect all equipment to be used (computers and headphones) to ensure they are operational. Check the audio volume on all computers prior to test administration. Check computers to ensure access to ISIP Español assessments. The assessment program can easily be downloaded from the Istation website at www.istation.com.

4.  Prior to testing, become familiar with the tests to be administered and the test formats.

5.  Make sure the physical conditions in the testing location are satisfactory. There should be adequate lighting for all students, and students should be able to be seated with enough space between them. Consider posting a "Testing – Do Not Disturb" sign on the classroom or lab door if the testing location is in a high-traffic area or prone to interruption by other students. If the test group will exceed 10 students, it is recommended that arrangements be made for a proctor (e.g. a lab manager) to assist in the test administration.

6.  For first-time users, ensure that students have sufficient proficiency in this medium. Students must be able to move a mouse pointer to an object on screen and click with the left mouse button. Early elementary students should have no difficulty with this task. ISIP Español does provide, prior to the first assessment, a practice activity that is unrelated to the assessments and that allows students to practice point-and-click skills. Although only point-and-click computer skills are necessary to complete the assessments, some users may find it appropriate to provide students without prior access to computers some instruction in basic computer terms, components (keyboard and mouse), and computer-use skills prior to assessment administration.

Once the initial administration of ISIP Español is complete, subsequent administration of tests should require minimal preparation, including the inspection of computers and headphones to ensure that they are operational.

## Materials

Only student login cards, operational headphones, and computers with Internet access are required for test administration. There are no CD-ROMs to install or school-based servers to maintain. Administration for schools is virtually non-existent. ISIP Español is downloaded from the Istation website at www.istation.com. After installation, any number of simultaneous students can be supported in ISIP Español, generally using the bandwidth of a single web surfer. In the event that the school's Internet connection is lost, ISIP Español will continue to function normally and will synchronize with Istation servers when the Internet connection is restored. Since ISIP Español is delivered through the Internet, enhancements and modifications are provided to users transparently, without a service call.

## Test Delivery

A summary of subtests is included under the section entitled Description of ISIP Español. ISIP Español provides for monthly assessment of early reading skills. Assessments can be run more frequently by teacher assignment on the Istation website at www.istation.com.

Upon student login to ISIP Español during each assessment period, ISIP Español will automatically deliver all assessments appropriate for that student, for that time of year. The entire battery of subtests runs seamlessly, back to back, without user or teacher manipulation. Tests are automatically scored by the program, and student results are immediately available to the teacher on the Istation website at www.istation.com.

# Administration Guidelines

1. Explain the assessment process and the setting. Encourage a positive attitude toward the test.

   **SAY** Hoy vamos a usar el computador para completar unos juegos de lectura. Estos juegos son muy importantes porque nos van a ayudar a saber cómo estas aprendiendo a leer. Los personajes en estos juegos son Don Buhiermo y su amiga Vampiresa. Ellos dos te van a ayudar para que logres MUCHO EXITO. Escucha con mucha atención las instrucciones y toma el tiempo necesario para leer. ¡Debes demostrar cuánto sabes!

2. Instruct the students to work independently and to quietly raise their hands if they need assistance.

**SAY**  Recuerda que para poder demostrar cómo estas aprendiendo a leer, debes trabajar tu sólo con el computador, de la misma manera que cuando contestas una prueba o examen. No debes tratar de adivinar o mirar el computador de otros niños. Cuando termines*, debes levantar la mano.

*Monitor to ensure students spend enough time during each subtest and DO NOT try to rush through it.

3.    Pass out login cards and assist the students as they login to Istation.  For first-time users, consider modeling the login steps on a computer or a projection screen. You may want to consider entering student logins and passwords and pressing PAUSE until the students are ready to begin. The test will begin as soon as the students press OK on the login screen.

**SAY**  Vamos a comenzar. ¿Recuerdan como entrar en istation? En la primera caja en blanco, escribe tu nombre como aparece en la tarjeta de istation. En la segunda caja en blanco escribe tu clave. Ponte tus audífonos sobre las orejas y haz clic en donde dice OK.

4.    Monitor student performance to ensure validity and reliability of test results. If a student needs assistance or must take a break, FIRST press the PAUSE key on the keyboard. This will interrupt the assessment currently being given without penalty to the student. The assessments are timed activities. Failure to PAUSE will result in the assessment continuing to run while assistance is being provided. When the student is ready to return to the assessment, press the PAUSE key again. The assessment will automatically return to the question where the student left off.

Be aware of fatigue and other behavioral issues, such as students losing interest, students who are easily distracted, students exhibiting frustration, and students who are not attempting to answer questions or are not trying. All of these behaviors often invalidate results. If any of these behaviors are noted, interrupt the student activity.

To assist a student:

a.    Press the Pause key.

b.    Ask the student to remove the headphones.

c.    Sit with the student at the computer.

d.    Do NOT provide answers or suggestions on how to respond to questions.

e.    If students appear to have lost interest or are not trying, remind him or her that it is important to follow the instructions and to do his or her best.

f.    If the student appears to be frustrated or asks for assistance, ask the student to repeat the instructions for the assessment.  If the student responds correctly, encourage the student to continue following directions and doing his or her best.  If the student responds

incorrectly, provide guidance and have the student demonstrate understanding of the directions before he or she restarts the assessment.

5.  Disruptive behavior should not be tolerated. A student who is disrupting others and whose behavior is not corrected by intervention should be removed from the testing area. Computer time should be rescheduled so that the student has an opportunity to complete the assessment.

6.  It is preferable, but not required, that the assessments be completed in a single session. Allow students to continue working in the assessment as long as they are being productive. The time allotment recommended for each assessment period is at least thirty minutes.

    When students begin the Reading Comprehension section, they should spend three to four minutes reading the passage before they press the "FIN" button to begin the questions. The passage will appear on the left side of the screen, and students MAY go back and review the passage while answering the questions.

7.  Some students will finish earlier than others. When they are finished, give them a book to read.

8.  Document any absent students, and schedule time for make-up assessments.

9.  Adhere to any accommodations for special education or limited English proficiency students. Accommodations should be made on an individual student basis and should take into consideration the needs of the student and whether the student normally receives accommodations.

Some accommodations to consider:

- For students with hearing difficulties, adjust the computer volume.

- For students with sight difficulties, arrange for use of a larger computer monitor.

- Oral instructions may be provided for the activities if necessary, including instructions in sign language.

**NOTE:** Using the PAUSE key to allow for more response time during the assessment is not advised.  The response time given to each item was built in at the time psychometric data was collected in order to determine the difficultly of each item. If the PAUSE key is used to lengthen item response time, the psychometric data collected on the items become invalid and ability scores may not be an accurate measure of student performance. The objective of computer adaptive testing is to adapt the assessment based on student response.  If students are unable to answer questions in the response time given, they will be given less difficult items.  An ability score obtained from modifying the test is not a score of the student's ability according to psychometric data collected.

10. Review test reports. If student results do not match teacher expectations or understanding of current skill knowledge, the entire assessment or individual skill probes may be repeated on a different day with different probes. Go to /.com and assign On-Demand assessments to the student in question. On the next student login, On-Demand assessments will run. The last of the two scores will be used as the current period indicator of the child's skill level.

Student results may require validation in the following situations:

- Session is interrupted. (ie.fire drill, class disturbance)

- Student answers randomly without listening to directions or reading questions.

- Student refuses to complete the assessment

- Student becomes ill

- Results aren't typical of student performance

# ISIP Español Protocols

This section describes subtests for ISIP Español. Samples of some of the assessments in each grade level—Kindergarten through Grade 3—are provided (in no particular order), followed by an explanation of what students are asked to do in each subtest. The explanations include specific directions spoken by the online game show host, Don Buhiermo, and the off-screen Narrator.

Every time a new assessment begins for a student, ISIP Español automatically provides a test warm-up. The test warm-up includes all directions for the assessment, models completion of one or more items, and allows the student to complete practice items. Narrator correction and feedback are provided in student interactions in all practice items.

In both warm-up activities and the assessment, students are also able to self-monitor progress in a fun and engaging manner. Audio prompts are used to distinguish correct and incorrect answers. For incorrect answers, a "boing" is used. For correct answers, a "ping" is used. At the conclusion of each subtest, the student result is presented in a graphical format along with prior results. Efforts are praised, and students are encouraged to "beat" their high score.

During each assessment, student progress is monitored and prompts encouraging student efforts are provided. Prompts vary based on the level of performance observed.

**After multiple incorrect responses:**

Narrator:             Pon mucha atención y contesta correctamente.

*Pay attention and answer correctly!*

# Beginning Sound

**Ages:** Kindergarten, Grade 1

In the **Beginning Sound** subtest, four pictures appear on the screen at once. The narrator says the name of each picture as the box around it highlights. The student is asked to click on the picture that has the same beginning sound as a sound produced orally by the narrator. The narrator then says one of the initial sounds. If the words used are unfamiliar vocabulary for the student, the student may move the mouse pointer over each picture and the narrator will repeat the word associated with it.

## Student Directions

### Modeled Instruction

| | |
|---|---|
| Don Buhiermo: | En este juego vas a buscar el dibujo que va con el sonido que te diga el narrador. Escucha atentamente las instrucciones. |
| | *In this game, you are going to find the picture that begins with a sound. Listen carefully to the instructions.* |
| Narrator: | Vas a ver unos dibujos y yo te diré lo que cada uno de ellos es. PERRO, LUNA, LLAVES, ALFOMBRA. Haz clic en el dibujo que comienza con /lu/, como la palabra /luz/. /lu/…ahora mueves la flecha hasta el dibujo que corresponde y haces clic para contestar. |
| | *You are going to see some pictures. I will say their names. DOG, MOON, KEYS, RUG. You are going to hear something like this: "Click on the picture that begins with the /m/ sound." You'll move your mouse until the arrow is on the picture that begins with the sound the narrator says. Then you'll click the mouse button.* |

**NOTE:** As instruction is provided, an arrow on screen models student behavior.

### Student Practice

| | |
|---|---|
| Narrator: | Vamos a practicar. RATA, SOL, COCHINITO, DADOS. Haz clic en el dibujo que comienza con /co/, como la palabra /cosa/. /co/. |
| | Now let's practice. RAT, SUN, PIG, DICE. Click on the picture that begins with the /p/ sound. /p/ |

**If student gives incorrect answer:**

| | |
|---|---|
| Narrator: | (boing) El dibujo que comienza con /co/ es <cochinito>. Inténtalo otra vez . |

No. Pig begins with the /p/ sound. Try again. (Last instructions given by Narrator are repeated.) **NOTE:** The student must answer correctly to move on.

**If student gives correct answer:**

Narrator: (ping) Cochinito, correcto. El dibujo que comienza con /co/ es <cochinito>.

Yes. Pig begins with the /p/ sound.

**If student does not respond in five seconds:**

Narrator: (boing)  El dibujo que comienza con /co/ es <cochinito>. Inténtalo otra vez.

The picture that starts with /co/ is <pig>.Try again. (Last instructions given by Narrator are repeated.) **NOTE:** The student must answer correctly to move on.

Don Buhiermo: ¡Bien hecho! ¡Uuu uuu uuu!

Good job!  Woo hoo hoo!

## For Assessment

Don Buhiermo: En este juego vas a buscar el dibujo que va con el sonido que te diga el narrador. Escucha con cuidado lo que pregunta el narrador y haz clic para contestar.

Here are some pictures. (The Narrator will say some words for these pictures). Click on the pictures that the narrator says the beginning sound for. Pay attention. Listen carefully to the narrator and click to answer.

Narrator: [pic name 1] - [pic name 2] - [pic name 3] - [pic name 4]
Haz clic en el dibujo que comienza con [1], como la palabra [2]. [1].

[pic name 1] - [pic name 2] - [pic name 3] - [pic name 4]
Click on a picture that starts with [1], like the word [2]. [1].

**If student gives correct answer:**

Narrator: There is no narrator response. Student response is scored as correct, and the next item is presented.

**If student gives incorrect answer:**

Narrator: There is no narrator response. Student response is scored as incorrect, and the next item is presented.

**If student does not respond in18 seconds:**

Narrator:                    There is no narrator response. Student response is scored as incorrect, and the next item is presented.


# Blending

**AGES:** Kindergarten, Grade 1

In the *Blending* subtest*,* four pictures appear on the screen with a box in the middle that contains an animated side view of a head. The narrator says the name of each picture as the box around it highlights. The narrator says a word, syllable by syllable, as the animated head produces each sound. The child is asked to click on the picture showing the word that has been said syllable by syllable. If the words used are unfamiliar vocabulary for the student, the student may move the mouse pointer over each picture and the narrator will repeat the word associated with it.

## Student Directions

### Modeled Instruction

Don Buhiermo:              En este juego vas a combinar unos sonidos para formar la palabra que va con el dibujo. Por lo tanto, escucha cada sonido que el personaje va a decir.

                           In this game, you will combine sounds to form a word that goes with the picture. Therefore, listen to each sound the character says.

Narrator:                  Vas a ver unos dibujos y yo te diré lo que cada uno de ellos es.
                           ALFOMBRA, PERRO, LLAVES, LUNA. Vas a escuchar algo así  /lu/ /na/.
                           Encuentra  la palabra que se forma al unir los sonidos y usa la flecha para señalar el dibujo que corresponde.

                           You are going to see some pictures. I will say their names. RUG, DOG, KEYS, MOON. I will tell you something like this: /m/ /oo/  /n/. You will put the sounds together and decide which picture I named.  You will use the mouse to move the pointer until it is on the correct picture. Then you will click the mouse button.

**NOTE:** As instruction is provided, an arrow on screen models student behavior.


### Student Practice

Narrator:                  Vamos a practicar. Pon mucha atención y contesta correctamente. SOL, COCHINITO, RATA, LLAVES. Haz clic en el dibujo de la palabra que se forma al unir: /so/ / l/.

                           Now let's practice. You have five seconds to answer. SUN, PIG, RAT, KEYS.

Click on the picture for the word you make by blending the sounds together. /s/ /u/ /n/.

**If student gives incorrect answer:**

Narrator:                    (boing) Al unir: /so/ /l/ se forma la palabra <sol>. Inténtalo otra vez .

The sounds /s/ /u/ /n/, blended together, make the word SUN. Try again.

**NOTE:** Student must answer correctly in order to move on.

**If student gives correct answer:**

Narrator:                    (ping) Sol, correcto. Al unir: /so/ /l/ se forma la palabra <sol>.

That's right. The sounds /s/ /u/ /n/, blended together, make the word SUN.

**If student does not respond in five seconds:**

Narrator:                    (boing) Inténtalo otra vez. Try again. (Last instructions given by Narrator are repeated, and student has an additional five seconds to respond.)

**NOTE:** Student must answer correctly in order to move on.

Narrator:                    Vamos a practicar otra vez. PATO, PERRO, GATO, POSTRE. Haz clic en el dibujo de la palabra que se forma al unir: /pe/ /rro/.

Let's do another one. DUCK, DOG, CAT, PIE. Click on the picture for the word you make by blending these sounds together. /d/ /o/ /g/.

**If student gives incorrect answer:**

Narrator:                    (boing) Al unir: /pe/ /rro/ se forma la palabra <perro>. Inténtalo otra vez.

The sounds /d/ /o/ /g/, blended together, make the word DOG. Try again.

(Last instructions given by Narrator are repeated.) **NOTE:** Student must answer correctly in order to move on.

**If student gives correct answer:**

Narrator:                    (ping) Perro, correcto. Al unir: /pe/ /rro/ se forma la palabra <perro>.

That's right. The sounds /d/ /o/ /g/, blended together, make the word DOG.

**If student does not respond in five seconds:**

Narrator: (boing) Inténtalo otra vez. Try again. (Last instructions given by Narrator are repeated, and student has an additional five seconds to respond.) **NOTE:** Student must answer correctly in order to move on.

Don Buhiermo: ¡Bien hecho! ¡Uuu uuu uuu!

Good job! Woo hoo hoo!

## For Assessment

Don Buhiermo: En este juego vas a combinar unos sonidos para formar la palabra que va con el dibujo. Haz clic solamente en el dibujo que contesta la pregunta después de escuchar el personaje que dice los sonidos. Pon mucha atención y contesta correctamente. ¡Uuu uuu uuu!

In this game, you will click on the picture of the word made by blending letter sounds together. Click on as many of the correct pictures as you can. Pay attention and do your best. Hoo hoo hoo!

Narrator: Haz clic en el dibujo de la palabra que se forma al unir: [1] [2].

Click on the picture of the word that is formed by blending: [1] [2].

**If student gives correct answer:**

Narrator: (ping) There is no narrator response. Student response is scored as correct, and the next item is presented.

**If student gives incorrect answer:**

Narrator: (boing) Pon mucha atención y contesta correctamente.

Pay attention and answer correctly.

(This applies to first and second incorrect responses only. Subsequent errors result in no dialogue or sound effect.) Next item is presented.

**If student does not respond in five seconds:**

Narrator: There is no narrator response. After 12 seconds, move on to the audio for the next set of pictures and question. Note: A non-response is scored as incorrect. After a second non-response, the next item is presented. The activity timer is stopped during the re-try.

# Sound Symbol

**AGES:** Kindergarten and Grade 1

In the Sound Symbol subtest, there are two question types:

1. Click on the letter that makes the sound _____.

   Four items, a combination of both upper- and lower-case letters, appear on-screen at once. The narrator asks the student to identify the symbol for the letter sound produced orally by the narrator.

2. Click on the letters that make the sound [syllable], as in the word [word].

   Four syllables appear on-screen at once. The narrator asks the student to identify the syllable sound produced orally by the narrator, as in the word given.

## Student Directions

No warm-up is provided. This is the same venue as Letter Recognition and operates in the same manner.

**Question type: Click on the letter that makes the sound _____.**

### Introduction

| | |
|---|---|
| Don Buhiermo: | Veamos cuántos sonidos de letras puedes encontrar. |
| | Let's see how many letter sounds you can find. |

### Modeling

| | |
|---|---|
| Narrator: | Voy a decir el sonido de una letra y tú haces clic en la letra que corresponde. Observa y escucha este ejemplo. Haz clic en la letra que suena "b." "b." |
| | I will say a letter sound, and you will click on the corresponding letter. Observe and listen to this example. Click on the letter that makes the sound "b." (The letters R c b A appear. The cursor moves to the letter "b" and a "correct" sound plays.) |

### Practice

| | |
|---|---|
| Narrator: | Vamos a practicar. Pon mucha atención y contesta correctamente. Haz clic en la letra que suena "A." |
| | Let's practice. Pay attention and answer correctly. Click on the letter that makes the sound "A." (The letters R c b A appear.) |

**If student gives incorrect answer:**

Narrator:
(boing) No. Ésta es la letra que suena "A". Inténtalo otra vez. Haz clic en la letra que  suena "A".

No, this letter makes the /"target sound"/ sound. Try again. Click on the letter that makes the sound "A."

**If student gives correct answer:**

Narrator:
(ping) There is no narrator response. Student response is scored as correct, and the next item is presented.

**If student does not respond in four seconds:**

Narrator:
(boing) Ésta es la letra que suena "A." Inténtalo otra vez.  Haz clic en la letra que suena "A."

This is the letter that makes the sound "A." Try again. Click on the letter that makes the sound "A."

## For Assessment

Don Buhiermo:
Veamos cuántos sonidos de letras puedes encontrar. Pon mucha atención y contesta correctamente.

Let's see how many letter sounds you can find.  Pay attention and answer correctly.

Narrator:
Haz clic en la letra que suena [1].

Click on the letter that makes the sound [1].

**If student gives incorrect answer:**

Narrator:
(After first and fourth incorrect responses) Pon mucha atención y contesta correctamente.

Pay close attention and answer correctly. (After all other incorrect responses) There is no narrator response. Student response is scored as incorrect, and the next item is presented.

**If student gives correct answer:**

Narrator:
There is no narrator response. Student response is recorded as correct, and the next item is presented.

**If student does not respond in 20 seconds:**

Narrator:               A non-response is scored as incorrect. The next item is presented.

**Question type: Click on the letters that make the sound [syllable], as in the word [word].**

## Introduction

Narrator:               Vamos a practicar otra vez. Haz clic en las letras que suenan "pa" como en la palabra "pato." "pa."

Let's practice again. Click on the letters that make the sound "pa," like in the word "pato." "pa."

**If student gives incorrect answer:**

Narrator:               (boing) No. Éstas son las letras que suenan "pa." Inténtalo otra vez.  Haz clic en las letras que suenan "pa" como en la palabra "pato." "pa."

No. These are the letters that make the sound "pa." Try again. Click on the letters that make the sound "pa," like in the word "pato," "pa."

**If student gives correct answer:**

Narrator:               (ping) There is no narrator response. Student response is recorded as correct, and the next item is presented.

**If student does not respond in four seconds:**

Narrator:               (boing) Estas son las letras que suenan "pa." Inténtalo otra vez.  Haz clic en las letras que suenan "pa" como en la palabra "pato." "pa."

These are the letters that make the sound "pa." Try again. Click on the letters that make the sound "pa," like in the word "pato." "pa."

## For Assessment

Don Buhiermo:               Pon mucha atención y contesta correctamente.

Pay attention and answer correctly.

Narrator:               Haz clic en las letras que suenan [1] como en la palabra [2]. [1].

Click on the letters that make the sound [1], like in the word [2]. [1].

**If student gives incorrect answer:**

Narrator:    (After first and fourth incorrect responses) Pon mucha atención y contesta correctamente.

Pay close attention and answer correctly. (After all other incorrect responses) There is no narrator response. Student response is scored as incorrect, and the next item is presented.

**If student gives correct answer:**

Narrator:    There is no narrator response. Student response is recorded as correct, and the next item is presented.

**If student does not respond in 20 seconds:**

Narrator:    A non-response is scored as incorrect. The next item is presented.

# Vocabulary

**AGES:** All students

Kindergarten students will identify pictures of words spoken orally. Grade 1, 2, and 3 students will both identify pictures of words spoken orally and indentify words that are synonyms or antonyms, complete a group, or are an example of a spoken word.

There are two types of questions used to measure a student's *Vocabulary* knowledge and to evaluate both the upper and lower bounds of knowledge:

1.  Haz clic en el dibujo que muestra un _____.

    In the first question type, four pictures appear on the screen. The narrator asks the student to identify the picture that best illustrates the word spoken orally.

2.  Haz clic en la caja donde dice _____. (example of, used for, completes this group, etc.)

    In the second question type, four words appear on the screen. Each of the four words is spoken by the narrator. The student is asked to identify the word that is a synonym or antonym, completes a group, or is an example of a spoken word.

## Student Directions

**Question type: Haz clic en el dibujo que muestra un _____.**

## Modeled Instruction

Don Buhiermo:    En este juego veremos cuántos dibujos puedes encontrar. Para jugar, primero escucha con atención al narrador y luego haz clic en el dibujo que contesta cada pregunta.

In this game, we will see how many pictures you can find. To play, first listen carefully to the narrator, and then click on the picture to answer each question.

Narrator:    Haz clic en el dibujo que muestra un león.

Click on the picture that goes with the word <lion>.

(Four pictures appear on the screen—a kite, a lion, a desk, and a monkey.)

Lleva la flecha al dibujo que corresponde y haz clic para contestar.

Click on the picture for the word lion. Move your mouse pointer to the picture for the word spoken, lion, and click on it.

**NOTE:** As instruction is provided, an arrow on screen models student behavior.

## For Assessment

Don Buhiermo:    En este juego veremos cuántos dibujos puedes encontrar. Para jugar, primero escucha con atención al narrador y luego haz clic en el dibujo que contesta cada pregunta.

In this game, you will see how many pictures you can find. To play, first listen carefully to the narrator and then click on the picture to answer each question. Pay close attention and answer correctly.

Narrator:    (pictures appear on screen) Haz clic en el dibujo de _____.

Click on the picture for the word "target word."

**If student gives incorrect answer:**

Narrator:    (After first and fourth incorrect responses) Pon mucha atención y contesta correctamente.

Pay close attention and answer correctly.

(After all other incorrect responses) There is no narrator response. Student response is scored as incorrect, and the next item is presented.

**If student gives correct answer:**

Narrator:                    There is no narrator response. Student response is scored as correct, and the next item is presented.

**If student does not respond in 20 seconds:**

Narrator:                    There is no narrator response. Student response is scored as incorrect, and the next item is presented.

## Scoring

This is a 1-minute activity. The activity timer is off during the warm-up, during all instructions, and during non-response retries.  The total number of correct items and accuracy rate are taken into consideration in the student score. The score is then normalized to an accuracy rate per minute. If the accuracy rate is in the range of chance (25% for this activity), a score of 0 is given.

## Student Directions

**Question type: Haz clic en la caja donde dice _____.** (example of, used for, completes this group, etc.)

## Modeled Instruction

Don Buhiermo:          Veamos cuántas palabras sabes. Para jugar, primero escucha lo que dice el narrador y luego haz clic en la mejor respuesta.

Let's see how many words you know. To play, first listen to what the narrator says, and then click on the best answer.

Narrator:                    Escucha las instrucciones y luego escoge la mejor respuesta de la lista. Vas a escuchar algo así.

Listen to the instructions and then choose the best answer from the list. You will hear something like this:

Haz clic en la caja que dice "un color." Si mueves la flecha para señalar cada palabra, podrás escucharla.  ROJO, LÁPIZ, REGLA, LIBRO. Cuando hayas encontrado la respuesta correcta, haz clic en el botón que le corresponde. La palabra "rojo" es un color, en cambio, las otras palabras no lo son.

Click on the box that says a color. If you move the arrow over each word, you will hear it. RED, PENCIL, RULER, BOOK. When you find the correct answer, click on it. The word "RED" is a color, and the other words are not.

**NOTE:** As instruction is provided, an arrow on screen models student behavior.

## For Assessment

| | |
|---|---|
| Don Buhiermo: | Veamos cuántas palabras sabes. Para jugar, primero escucha lo que dice el narrador y luego haz clic en la mejor respuesta. Pon mucha atención y contesta correctamente. |
| | Let's see how many words you know. To play, first listen to what the narrator says, and then click on the best answer. Pay close attention and answers correctly. |
| Narrator: | Haz clic en la caja donde dice una palabra que es lo mismo que_____. |
| | Click on the word that has the same or similar meaning as "target word." |

**If student gives incorrect answer:**

| | |
|---|---|
| Narrator: | (After first and fourth incorrect responses) Pon mucha atención y contesta correctamente. |
| | Pay close attention and answer correctly. (After all other incorrect responses) There is no narrator response. Student response is scored as incorrect, and the next item is presented. |

**If student gives correct answer:**

| | |
|---|---|
| Narrator: | There is no narrator response. Student response is scored as correct, and the next item is presented. |

**If student does not respond in 20 seconds:**

| | |
|---|---|
| Narrator: | There is no narrator response. Student response is scored as incorrect, and the next item is presented. |

# Listening Comprehension

**AGES:** Kindergarten only.  Assessments include audio passages and comprehension questions. Kindergarten answer choices are pictures.

The objective of the *Listening Comprehension* subtest is for children to demonstrate listening comprehension by answering evidential and inferential questions about a reading selection.

## Student Directions

### Modeled Instruction

| | |
|---|---|
| Don Buhiermo: | En este juego vas a escuchar un cuento. Pon mucha atención a lo que dice el narrador. |
| | In this game, you will listen to a story. Pay attention to what the narrator says. |
| Narrator: | Voy a leer un cuento y tú debes escuchar con atención. Después de leer el cuento, vas a contestar unas preguntas. Vas a escoger tu respuesta haciendo clic en uno de los cuatro dibujos. A Daniel le gusta ir de vacaciones. A él le encanta visitar las montañas. ¿Cuál es el lugar que a Daniel le gusta visitar más? |
| | Listen carefully as I read a story. After I read the story, you will answer some questions. You will choose your answer by clicking on one of the four pictures. "Dan likes to go on vacation. His favorite place to visit is the mountains. What is Dan's favorite place to visit?" (The 4 pictures appear and highlight as they are spoken.) |
| Narrator: | la playa |
| | el río |
| | el desierto |
| | las montañas |
| | the beach |
| | the river |
| | the desert |
| | the mountains |
| | Mueve la flecha al dibujo que tiene la respuesta correcta y haz clic. Si quieres escuchar cada respuesta otra vez, usa la flecha para señalar cada dibujo. |
| | Move the pointer to the picture that has the correct answer and click on it. If you want to hear the options again, move the pointer over each picture. |
| PRACTICE | Vamos a practicar. ¿De quién se trata este cuento? (de un perro, de un niño, de una jirafa, de un ratón.) Haz clic en el dibujo que muestra la mejor respuesta. |
| | Let's practice. Who is this story about? (The 4 pictures appear and highlight as they are spoken -- dog, Dan, giraffe, mouse.)      Click on the picture that best answers the question. (for practice only) |

**If student gives incorrect answer:**

Narrator:         No, esa no es la respuesta correcta. Inténtalo de nuevo.

No, that's not the best option. Try again.

(Last instructions given by Narrator are repeated.)

**NOTE:** Student must answer correctly in order to move on.

**If student gives correct answer:**

Narrator:         ¡Bien hecho!  Empecemos de una vez.

Good job. Let's get started.

**If student does not respond in 20 seconds:**

Narrator:         Escucha con atención mientras te leo un cuento. A Daniel le gusta ir de vacaciones. A él le encanta  visitar las montañas. ¿De quién se trata este cuento? Si quieres escuchar cada respuesta otra vez, usa la flecha para señalar cada dibujo. Haz clic en el dibujo que muestre la mejor respuesta.

Listen carefully as I read the story. Dan likes to go on vacation. His favorite place to visit is the mountains. Who is this story about? (The 4 pictures appear and highlight as they are spoken.) If you want to hear the options again, move the pointer over each picture. Click on the picture that best answers the question.

## For Assessment

Don Buhiermo:         En este juego vas a escuchar un cuento. Pon mucha atención a lo que dice el narrador.

In this game, you will listen to a story. Pay attention to what the narrator says.

Narrator:         Voy a leer un cuento y tú debes escuchar con atención. Después de leer el cuento, vas a contestar unas preguntas. Vas a escoger tu respuesta haciendo clic en uno de los cuatro dibujos.

Listen carefully as I read you a story. After I read the story, you will answer some questions. You will choose your answer by clicking on one of the four pictures.

(Passage picture appears. Story is read. Question is asked. The 4 pictures appear and highlight as they are spoken.)

**If student gives incorrect answer:**

Narrator:                       There is no narrator response. Student response is scored as incorrect, and the next item is presented.


**If student gives correct answer:**

Narrator:                       There is no narrator response. Student response is recorded as correct, and the next item is presented.


**If student does not respond in 20 seconds:**

Narrator:                       A non-response is scored as incorrect. The next item is presented.


# Reading Comprehension

**AGES:** Kindergarten, Grade 1, Grade 2, and Grade 3.  Assessments include reading selections and comprehension questions. Kindergarten answer options are pictures. Grade 1, 2, and 3 answer choices are not pictures.

The objective of the *Reading Comprehension* subtest is for children to demonstrate reading comprehension by answering evidential and inferential questions about a reading selection.

## Student Directions

Narrator:                       Para jugar leerás un texto. Cuando termines de leer, haz clic en el botón que dice FIN  y vas a pasar a la primera pregunta. Después de leer la pregunta y las opciones para contestar, haz clic en la mejor respuesta.

In this game, you will read a story. When you finish reading, click on the END button and you will then see the questions. Read the question and answer choices, and then click on the best answer.

**NOTE:** As instruction is provided, an arrow models student behavior. Sample passage from Introduction is already visible. Narrator reads the passage as words highlight, one at a time.

## Practice

Narrator:                       Observa y escucha este ejemplo.

Watch and listen as I do an example.

Me gusta ir a la escuela porque me gusta leer. Me gusta leer cuentos de animales. El libro que más me gusta es el libro del oso. El oso come peces y se va con otros ositos.

[Story about reading a book about bears at school.]

Después de leer, haz clic en el botón que dice FIN.

After reading, click on the END button.

Ahora vamos a leer la primera pregunta juntos.

Now, we will read the first question together.

¿Cuál es el libro favorito del niño?

[Narrator reads the question.] The question highlights as a whole (not wordbyword) while it is read.

La Escuela. El Niño. El Oso. La Casa.

School, Child, Bear, House.

[Narrator reads answer choices] Each answer choice highlights (as a whole) as it is read.

Haz clic en la mejor respuesta. Mueve la flecha  sobre cada respuesta o sobre la pregunta y podrás escucharla otra vez.

Click on the best answer. To hear the question or answer choices, use your mouse to scroll over them.

¿Cuál es el libro favorito del niño?

What is the child's favorite book?

La Escuela. El Niño. El Oso. La Casa.

School, Child, Bear, House.

Al leer el cuento, descubres que el libro que más le gusta al niño es el libro del oso.

According to the story, the child likes to read the bear book.

Haz clic para contestar.

Click on the right answer.


**If student gives incorrect answer:**

Narrator:              No, esa no es la respuesta correcta. Inténtalo de nuevo.

                       No, that's not the best option. Try again.

(Last instructions given by Narrator are repeated.)

**NOTE:** Student must answer correctly in order to move on.

**If student gives correct answer:**

Narrator:              ¡Correcto! El libro que más le gusta al niño es el libro del oso.

                       Yes. The child likes to read the bear book.

**If student does not respond in 50 seconds:**

Narrator:              Haz clic en la mejor respuesta.

                       Click on the best answer.

                       (The last instructions given by the Narrator are repeated, and the student has an additional 50 seconds to respond.)

                       **NOTE:** Student must answer correctly in order to move on.

## For Assessment

Don Buhiermo:          Veamos cuántas preguntas puedes contestar sobre la lectura. Escucha atentamente las instrucciones.

                       It's time to show what you know by reading passages and then answering questions. Listen carefully to the instructions.

Narrator:              Comienza a leer. Cuando termines haz clic en el botón que dice FIN.

                       Read the passage. When you finish reading, click on the END button.

**If student gives incorrect answer:**

Narrator:              (After first and fourth incorrect responses) Pon mucha atención y contesta correctamente.

                       Pay close attention and answer correctly.

                       (After all other incorrect responses) There is no narrator response. Student response is scored as incorrect, and the next item is presented.

**If student gives correct answer:**

Narrator:              There is no narrator response. Student response is recorded as correct, and the next item is presented.

**If student does not respond in 50 seconds:**

Narrator:          A non-response is scored as incorrect. The next item is presented.

# Fluency

**AGES:** Grade 2 and Grade 3

*Fluency* is constructed in a very different manner than the other subtests. Children are assessed on their skills in reading text with meaning in a specified period of time. In order to assess text reading on the computer, a maze task is utilized in which words of grade-leveled stories are left blank from the text. The child is given 3 choices for each blank from which to choose the word that works in the sentence. This task has been shown to be highly correlated to measures of both fluency and comprehension.

## Student Directions

## Modeled Instruction

Don Buhiermo:     Ahora veamos cuanto sabes sobre los cuentos. Pon mucha atención a lo que dice el narrador.

                   It's time to show what you know by reading a story. Listen carefully to the instructions.

Narrator:          Para jugar, leerás un cuento. Al leer, encontrarás que faltan algunas palabras para completar los párrafos del cuento. Cada espacio en blanco tiene 3 palabras posibles para llenarlo pero solo una de ellas completa la oración correctamente.

                   In this game, you will read a story. As you read, you will get to places where a word is missing. Your job will be to fill in the blank with the word that makes the most sense in the sentence.

                   Observa y escucha este ejemplo.   El año pasado, Arturo entró a primer grado. El primer día de clases tenía _____ miedo. Falta una palabra ¿Verdad? Al señalar el espacio en blanco aparecen tres posibles palabras. Haz clic en la palabra que corresponde o que queda mejor en cada oración.

                   Observe and listen to this example. Last year, Arturo entered first grade. The first day of class, he was _____scared. A word is missing, right? When pointing to the blank space, three possible words appear. Click on the word that best completes the sentence.

**NOTE:** As instruction is provided, an arrow models student behavior on the screen.

Leamos juntos. El año pasado, Arturo entró a primer grado. El primer día de clases tenía_____ miedo.  [mucho bonito viajes] ¿Cuál de las tres palabras queda mejor en la oración?

Let's read together.  Last year, Arturo entered first grade. The first day of class, he was_____ scared. [very, beautiful, trips] Which of the three words best completes the sentence?

El primer día de clases tenía [mucho] miedo es la respuesta correcta. Cuando lees la oración con cada una de las palabras, notarás que [bonito] y [viajes] no quedan bien en la oración.

"The first day of class, he was [very] scared" is the correct answer. When you read the sentence with each one of the words, you will notice that [beautiful] and [trips] do not fit well in the sentence.

## Student Practice

Narrator:                       Es tu turno para contestar. Recuerda, cada espacio en blanco tiene 3 palabras posibles para llenarlo, pero solo una de ellas queda bien. Haz clic en la palabra que corresponde o que queda mejor en cada oración.

Now, you try. Remember, each blank has 3 words, but only one makes sense. Choose the word that makes the most sense by moving the pointer over the blank and clicking on the word that makes the most sense in the sentence.

**If student gives incorrect answer:**

Narrator:                       (boing) La palabra [XXX] no queda bien en la oración.

No. The word XXX does not make sense in the sentence. Try again.

Note: Student must answer correctly in order to move on.

**If student gives correct answer:**

Narrator:                       (ping) BIG. That's right. "Jan said it was the best pig. The pig was big." BIG makes the most sense in the sentence.

**If student does not respond in 20 seconds:**

Narrator:                       Mueve la flecha al espacio en blanco para ver las tres palabras posibles. Haz clic en la palabra que corresponde o que queda bien.

Move the pointer to the blank to see the three possible words. Click on the word that makes sense in the sentence.

(Last instructions given by Narrator are repeated.)

Note: Student must answer correctly in order to move on.

Inténtalo otra vez.

Try again.

Para pasar a la próxima página, haz clic en la flecha verde que está en la parte de abajo de la página. Haz clic en la flecha verde que se prende y se apaga cuando estés listo para empezar.

When you get to the end of a page, you will need to turn to the next page. To turn the page, click on the green arrow at the bottom of the page. Click on the flashing green arrow to turn the page now.

## For Assessment

Don Buhiermo:    Ahora veamos cuanto sabes sobre los cuentos. Pon mucha atención y contesta correctamente.

It's time to show what you know by reading a story. Pay attention and answer correctly.

Narrator:    Lee atentamente este cuento. Al leer, encontrarás que faltan algunas palabras para completar los párrafos del cuento. Cada espacio en blanco tiene tres palabras posibles para llenarlo pero solo una de ellas completa la oración correctamente. Mueve la flecha al espacio en blanco para ver las tres palabras posibles. Haz clic en la palabra que corresponde o que queda bien. Haz clic en la flecha verde que se prende y se apaga cuando estés listo para empezar.

Read this story carefully. As you read, you'll find that there are some words missing from the story. Each blank has three possible words to fill it, but only one of them completes the sentence correctly. Move the pointer to the blank to see the three possible words. Click on the word that fits well.  Click on the flashing green arrow when you are ready to begin.

**If student gives incorrect answer:**

Narrator:    There is no narrator response. The correct word is placed in the blank. The item is scored as incorrect.

**If student gives correct answer:**

Narrator:    There is no narrator response. The item is scored as correct.

**If student shows repeating patterns of incorrect answers:**

Don Buhiermo leans in with a series of prompts to encourage the student to focus and to do his or her best without guessing. Examples of the prompts are provided in the opening of the ISIP Español Protocols section.

**If student does not respond in 20 seconds:**

Narrator:    Mueve la flecha al espacio en blanco para ver las tres palabras posibles. Haz clic en la palabra que corresponde o que queda bien.

Move the pointer to the blank to see the three possible words. Click on the word that fits well.

# Spelling

**AGES:**  Grade 1, Grade 2, and Grade 3

The objective of the *Spelling* subtest is to determine if children are developing fully-specified orthographic representations of words.  There are four types of questions in the spelling assessment:

1.  **Click on the word that is written correctly.**

    For each item, a group of words appears on the screen, and the computer asks the child to choose the word that is spelled correctly. The child then chooses the word by clicking on it.

2.  **Click on the correct *letters* to write the word.**

    For each item, an array of *letters* appears on the screen, and the computer asks the child to spell a specific word using those letters.  The child then spells the word by clicking on each letter.  As each letter is selected, the word is formed on a line that appears directly below the letter array.

3.  **Click on the correct *syllables* to write the word.**

    For each item, an array of *syllables* appears on the screen, and the computer asks the child to spell a specific word using those syllables. The child then spells the word by clicking on each syllable. As each syllable is selected, the word is formed on a line that appears directly below the *syllable* array.

4.  **Click on the correct *words* to write the sentence.**

    For each item, an array of *words* appears on the screen. After the computer dictates a sentence, it asks the child to write the sentence using the words on the screen. The child then chooses each word by clicking on it. As each word is selected, the sentence is formed on a line that appears directly below the word array.

# Student Directions

**Question type: Click on the word that is written correctly.**

## Modeled Instruction and Practice

| | |
|---|---|
| Don Buhiermo: | Veamos cuántas palabras sabes escribir correctamente. Para jugar, primero escucha lo que dice el narrador y luego haz clic en la mejor respuesta. |
| | Let's see how many words you can write correctly. To play, first listen to what the narrator says, and then click on the best answer. |
| Narrator: | Escucha las instrucciones y luego escoge la mejor respuesta de la lista de palabras. Vas a escuchar algo así. [lapis - lapiz - lápiz – ládiz] Haz clic en la palabra que está escrita correctamente. Busca la palabra que está escrita correctamente. Las demás opciones están mal escritas.  Cuando la encuentres, mueve la flecha sobre la palabra que está bien escrita y haz clic para contestar. ¿Estás listo para buscar palabras? |
| | Listen to the instructions, and then choose the best answer from the word list. You will listen to something like this: [pensil - penzil - pencil – pemcil]. Click on the word that is written correctly. The rest of the options are written incorrectly. When you find it, move the arrow to the word that is written correctly and click on it to answer. Are you ready to find words? |

**NOTE:** As instruction is provided, an arrow models student behavior by spelling *nest* and clicking OK.

## For Assessment

| | |
|---|---|
| Don Buhiermo: | Para jugar, primero escucha lo que dice el narrador y luego haz clic en la mejor respuesta. Pon mucha atención y contesta correctamente. |
| | To play, first listen to what the narrator says, and then click on the best answer. Pay attention and answer correctly. |
| Narrator: | Haz clic en la palabra que está escrita correctamente. |
| | Click on the word that is written correctly. |

**If student gives incorrect answer:**

| | |
|---|---|
| Narrator: | (After first and second incorrect responses) Pon mucha atención y contesta correctamente. |
| | Pay close attention and answer correctly. |

(After all other incorrect responses) There is no narrator response. Student response is scored as incorrect, and the next item is presented.

**If student gives correct answer:**

Narrator:                    There is no narrator response. Student response is recorded as correct, and the next item is presented.

**If student does not respond in 15 seconds:**

Narrator:                    A non-response is scored as incorrect. The next item is presented.

## Student Directions

**Question type: Click on the correct *letters* to write the word.**

### Modeled Instruction and Practice

Don Buhiermo:        ¿Estás listo para jugar a deletrear? Para jugar, primero escucha la palabra que el narrador dice y luego haz clic en las letras necesarias para escribir esa palabra.

*Are you ready for a spelling game? To play, first listen to each word the narrator says, and then click on the letters you need to spell this word.*

Narrator:                    Yo diré una palabra, luego usaré la palabra en una oración y después diré la palabra otra vez. Ordena las letras que necesitas para deletrear la palabra, haciendo clic en cada una de ellas.

I will say a word, use the word in a sentence, and then repeat the word.

Organize the letters you need to spell the word by clicking on each one of them.

Observa y escucha este ejemplo. Haz clic en las letras correctas para escribir la palabra <ciencias>. <Se van a la clase de ciencias>. <ciencias>. Voy a tener que cambiar la letra que escogí mal. Primero voy al botón de borrar y hago clic. Luego hago clic en las demás letras que necesito y termino de escribir la palabra, de esta manera.

Observe and listen to this example. Click on the correct letters to write the word <science>. <They're going to science class>. <science>. I'm going to have to change the incorrect letter selected. First, I'll click on the eraser button. Then, I'll click on the rest of the letters I need and finish writing the word, like this.

Don Buhiermo:        c - i - e - n- c - i - a - s. ¡Muy bien! Así se escribe correctamente.

s - c - i - e - n - c - e.  Very good! That's how it's written correctly.

| Narrator: | Comencemos. |
|---|---|
| | Let's start. |
| Don Buhiermo: | Pon mucha atención y contesta correctamente. |
| | Pay attention and answer correctly. |
| Narrator: | Haz clic en las letras correctas para escribir la palabra. (for first question only) |
| | Click on the correct letters to write the word. |
| | Escribe la palabra. (for the second question and onward) |
| | Write the word. |

**If student gives incorrect answer:**

| Narrator: | (After first and second incorrect responses) Pon mucha atención y contesta correctamente. |
|---|---|
| | Pay close attention and answer correctly. |
| | (After all other incorrect responses) There is no narrator response. Student response is scored as incorrect, and the next item is presented. |

**If student gives correct answer:**

| Narrator: | There is no narrator response. Student response is recorded as correct, and the next item is presented. |
|---|---|

**If student does not respond in 30 seconds:**

| Narrator: | A non-response is scored as incorrect. The next item is presented. |
|---|---|

## Student Directions

**Question type: Click on the correct *syllables* to write the word.**

## Modeled Instruction and Practice

| Don Buhiermo: | ¿Estás listo para jugar con sílabas? Para jugar, primero escucha la palabra que el narrador dice y luego haz clic en las sílabas o los sonidos necesarios para escribir la palabra. |
|---|---|
| | Are you ready to play with syllables? To play, first listen to each word that the narrator says, and then click on the syllables or sounds to write the word. |

| Narrator: | Yo diré una palabra, luego la voy a repetir mientras tú [pausa] piensas en los sonidos que la forman. Ordena los sonidos o las sílabas que necesitas para formar la palabra, haciendo clic en cada una de ellas. |
|---|---|
| | I'll say a word, and then I'll repeat the word slowly while you [pause] think about the sounds that make the word. Organize the syllables or sounds you need to spell the word by clicking on each one of them. |

## Student Directions

**Question type: Click on the correct *words* to write the sentence.**

### Modeled Instruction

| Don Buhiermo: | Veamos cuántas oraciones puedes escribir. Para jugar, primero escucha la oración que el narrador dice y luego haz clic en las palabras que necesitas para escribir la oración. |
|---|---|
| | Let's see how many sentences you can write. To play, first listen to each sentence the narrator says, and then click on the words you need to write the sentence. |
| Narrator: | Vamos a escribir la oración <Mi mamá está en casa>. Ordena las palabras que necesitas para escribir la oración, haciendo clic en cada una de ellas. Observa y escucha este ejemplo. Cielos…. Hice clic en una palabra que no va en la oración <Mi mamá está en casa>. La palabra que sigue es <en>. Voy a tener que cambiar la palabra que escogí mal. Primero voy al botón de borrar y hago clic. Luego hago clic en las demás palabras que necesito y termino de escribir la oración, de esta manera. |
| | Let's write the sentence <My mother is home>. Organize the words you need to write the sentence by clicking on each one of them. Observe and listen to this example: "Heavens… I clicked on a word that doesn't belong to the sentence <My mother is home>." The word that follows is <en>. I'm going to have to change the incorrect word I selected. First, I'll click on the eraser button. Then, I'll click on the rest of the words I need and finish writing the sentence, like this. |
| Don Buhiermo: | Mi mamá está en casa. ¡Muy bien! Así se escribe correctamente. |
| | My mother is home. Very good! That's how it's written correctly. |
| | Ahora te toca a ti. Escucha cada palabra que dice el narrador y sigue jugando. |
| | Now, it is your turn. Listen to each word the narrator says and continue playing. |
| Narrator: | Yo diré una oración, luego la voy a repetir mientras tú [pausa] piensas cómo vas a escribirla. Ordena las palabras que necesitas para escribir la oración, haciendo |

clic en cada una de ellas. Cuando termines la oración, haz clic en el botón verde. Comencemos.

I'll say a sentence, and then I'll repeat it while you [pause] think of how you're going to write it. Organize the words you need to write the sentence by clicking on each one of them. When you are done writing, click on the green button. Let's start.

## Practice

Narrator:            Haz clic en las palabras correctas para escribir la oración <Mi maestra es buena>.

Click on the correct words to write the sentence <My teacher is good>.

**If student gives incorrect answer:**

Narrator:            Así no se escribe <Mi maestra es buena>. Las palabras para escribir esta oración son: <Mi - maestra - es - buena>. Haz clic en las palabras correctas para escribir la oración  <Mi maestra es buena>. Cuando termines la oración, haz clic en el botón verde.

That is not how you write <My teacher is good>. The words to write this sentence are <My - teacher - is - good>. Click on the correct words and write the sentence <My teacher is good>. When you are done writing, click on the green button.

**If student gives correct answer:**

Narrator:            ¡Bien hecho!

Good job!

**If student does not respond in 30 seconds:**

Narrator:            Las palabras para escribir esta oración  son: <Mi - maestra - es - buena>. Haz clic en las palabras correctas para escribir la oración  <Mi maestra es buena>. Cuando termines la oración, haz clic en el botón verde.

The words to write this sentence are <My teacher is good>. Click on the correct words and write the sentence <My teacher is good>. When you are done writing, click on the green button.

## For Assessment

Don Buhiermo:        Pon mucha atención y contesta correctamente.

|  | Pay attention and answer correctly. |
|---|---|
| Narrator: | Haz clic en las palabras correctas para escribir la oración [1]. [pausa] [1]. |
|  | Click on the correct words and write the sentence [1]. [pause] [1]. |
|  | Note: This for the first question only. |
|  | Escribe la oración [1]. [pausa] [1]. |
|  | Write the sentence [1]. [pause] [1]. |
|  | Note: This is for the second question and onward. |

**If student gives incorrect answer:**

| Narrator: | (After first and second incorrect responses) Pon mucha atención y contesta correctamente. |
|---|---|
|  | Pay close attention and answer correctly. |
|  | (After all other incorrect responses) There is no narrator response. Student response is scored as incorrect, and the next item is presented. |

**If student gives correct answer:**

| Narrator: | There is no narrator response. Student response is recorded as correct, and the next item is presented. |
|---|---|

**If student does not respond in 30 seconds:**

| Narrator: | A non-response is scored as incorrect. The next item is presented. |
|---|---|

# Chapter 3: Using and Interpreting ISIP Español Reports

Providing administrators, teachers, and parents with timely student data is the key ingredient to linking ISIP Español assessment results to instructional planning. In any data-driven or results-oriented model of instruction, the needs are the same:

- Information that will assist in the identification of students who need additional support or different forms of support in order to achieve reading goals.

- Ongoing information on student performance against goals that will assist in evaluating the effectiveness of instruction and in developing and modifying instructional plans that can change reading outcomes for students at risk of failure.

- Information that will assist in the evaluation of instruction and instructional supports at all levels—district, area, school, and classroom—and from year to year, which can inform decisions about allocating resources and efforts.

What is lacking in existing models is the availability of data early enough in assessment–instruction decision loops. When learning builds on prior concepts, the teacher must know quickly who is struggling and whether existing instructional methods are effective in preventing students from falling further and further behind. Only when data results in timely remedial actions can it significantly affect outcomes.

## Understanding ISIP Español Scores

ISIP integrates computerized adaptive testing that accurately reflects the reading ability level of each student and measures growth over time. When administered regularly over time, it is possible to observe whether a student, or an entire classroom, district, or school, is making adequate progress in the critical reading areas.

Adaptive assessments use interactive content to measure a student's reading ability and skill development. Test questions range from easy to hard for each reading domain for students in Pre-Kindergarten through Grade 5. To identify the student's overall reading ability and individual skill ability, the difficulty of the test questions presented changes with every response. If a student answers questions correctly, ISIP presents more challenging questions until the student shows mastery or responds with an incorrect answer. When a student answers a question incorrectly, ISIP presents less difficult questions until the student begins answering correctly again. The ability score is an estimate of the student's reading ability. It shows how a student is doing compared to his or her previous performance and to other students at the same grade level.

## Ability Index

ISIP assessments use a measurement scale that aligns student performance levels with test question levels of difficulty on the same scale. The scale is divided into equal parts. These parts are called ability indices. All test questions are placed on the ability index scale according to their difficulty. Each increasing ability index is assigned a numeric value that indicates a higher level of difficulty. As a student takes an ISIP assessment, he or she is presented with test questions of varying ability indices or levels of difficulty. Once ISIP determines the difficulty level at which the student is able to perform, the test ends and the student is assigned an overall reading ability index, as well as ability indices for individual subtests.

Since ISIP is adaptive and the test questions are displayed based on student performance, not age or grade, identical ability indices across grades mean the same thing. For example, a first grader who receives a score of 215 and a third grader who receives a score of 215 are performing at the same level. Like measuring a child's height, measurements are added together to get class, school, and district averages. Ability indices make it possible to track a student's growth from year to year.

This ability index can be used by teachers to inform instruction around their students' strengths and weaknesses. Targeted instruction leads to better performance and maximum growth.

## Normative Data

National norms for ISIP Español are provided for students in Pre-Kindergarten through Grade 3. These norms enable teachers and parents to know how their students' scores compare with a nationally representative sample of children in their particular grade. The norming samples were obtained as part of Istation's ongoing research in assessing reading ability.

The samples were drawn from enrolled ISIP Español users during the 2010-2011 school year. Student percentile ranks were established using the monthly overall reading ability index, as well as the ability index for each ISIP Español subtest.

If a student scores at the 75th percentile; for example, it would mean that the student performed better than 65 percent of the students in the norm group. This allows for student performance to be compared to a reasonable control group, and provide a fair assessment of their reading abilities.

## Instructional Tier Goals

Consistent with other reading assessments, Istation has defined a three-tier normative grouping based on indices associated with the 20th and 40th percentiles. Students with an index above the 40th percentile for their grade are placed into Tier 1. Students with an index at or below the 20th percentile are placed into Tier 3. These tiers are used to guide educators in determining the level of instruction for each student. That is, students classified as:

A year-to-year history of ISIP Español results is available. Administrators, principals, and teachers may use their reports to evaluate and modify curriculum, interventions, AYP progress, the effectiveness of professional development, and personnel performance.

- Tier 1 (above the 40th percentile) are on track and performing at grade level.

- Tier 2 (between 21st and 40th percentile) are at some risk, are performing moderately below grade level, and are in need of intervention.

- Tier 3 (20th percentile and below) are at risk, are performing seriously below grade level, and are in need of intensive intervention.

Students who are classified as Tier 2 across all subtests should be considered to be having comprehensive reading difficulties and should receive Tier 3 instruction.

## Grade Level Equivalencies

Grade Level Equivalencies are scores based on the performance of students in the 2010–2011 norming group. The grade level equivalent (GE) represents the grade level and month of the typical score for students taking ISIP Español. If a student receives a GE of 2.4, this means that the student earned a score similar to the 50th percentile students in the test's norming group who were in their fourth month of Grade 2.

The grade level equivalent does not represent the appropriate level of instructional material with which a student should be placed. Grade level equivalencies should never be interpreted literally, but rather as a rough estimate of a student's grade level performance.

## Difference Between Ability Index Scores and Grade Level Equivalencies

There are basic differences between Ability Index Scores and Grade Level Equivalencies.  The Ability Indices represent a student's performance on a measurement scale of skill and reading ability. In contrast, the grade level equivalent represents a student's performance in comparison to students who were in the norming group.

## Growth

Growth within ISIP Español can be defined as an increased change in the student's score and improvement in ability over time. District, school, and student growth can be viewed on various ISIP Español reports.

# Using and Interpreting ISIP Español Reading Reports

The technology underlying ISIP Español delivers computer-based assessments, real-time evaluation of results, and immediate availability of reports on student progress. Assessment reports automatically group students according to the level of skill and support needed. Teachers are provided links to teacher-directed plans of instruction, downloadable lessons, and materials appropriate for each group.

Data is provided in both graphical and detailed numerical formats on every measure and at every level of a district's reporting hierarchy. Data is seamlessly and securely shared by users within the district, based upon authorization levels. Data may be shared with state information systems if requested by a school district. Individual student information can be provided to parents or guardians of students tested.

Istation provides the following ISIP Español Reports:

| Report Title | Description | Target Users |
|---|---|---|
| Executive Summary | The Executive Summary Report provides a brief overview of the current ISIP assessment. This report is available only to manager accounts and provides information only for the school or district level. | • Managers (at campus, district, or area) |
| Distribution | The Distribution Report shows the number of students performing in ranges of ability. | • Managers (at campus, district, or area) |
| Summary | The ISIP Summary Report shows the number and percentage of students at each instructional tier for the current month. | • Teachers<br>• Managers (at campus, district, or area) |
| Tier Movement | The Tier Movement Report shows a comparison of the number and percentage of students who were categorized at each instructional tier of Tier I, Tier II, Tier III through the current month. | • Teachers<br>• Managers (at campus, district, or area) |
| Skill Growth | The Skill Growth Reports show each skill assessed and the progress made by the students through the current month as measured against performance goals. | • Teachers<br>• Managers (at campus, district, or area) |

| Skill Growth by Tier | The Skill Growth by Tier Reports show each skill assessed and the progress made by the students through the current month as measured against performance goals within tier groups. | • Teachers<br>• Managers (at campus, district, or area) |
|---|---|---|
| Priority | The Priority Report alerts teachers of students needing additional support, and provides lessons based on demonstrated weaknesses. | • Teachers |
| Priority Summary | The Priority Summary Report, available to manager level users only, summarizes the use of the Priority Report by averaging how many days it has taken to acknowledge student alerts on the Priority Report. | • Managers (at campus, district, or area) |
| Priority Report – Student Intervention History | The Priority Report-Student Intervention History is a history of Priority Report alerts for a student, including those from current and prior school years. | • Teachers<br>• Managers (at campus, district, or area) |
| Student Summary Handout | The Student Summary Handout provides student performance data from the most recently completed ISIP assessment. | • Teachers<br>• Parents |

# District, Area, School, and Classroom Level Reports

## Executive Summary Report

The Executive Summary Report provides a brief overview of the current ISIP assessment. This report is only available to manager accounts and only provides information for the school or district level.



Instructional Tier Growth displayed for each Grade

Student Performance displayed by Grade and Tier
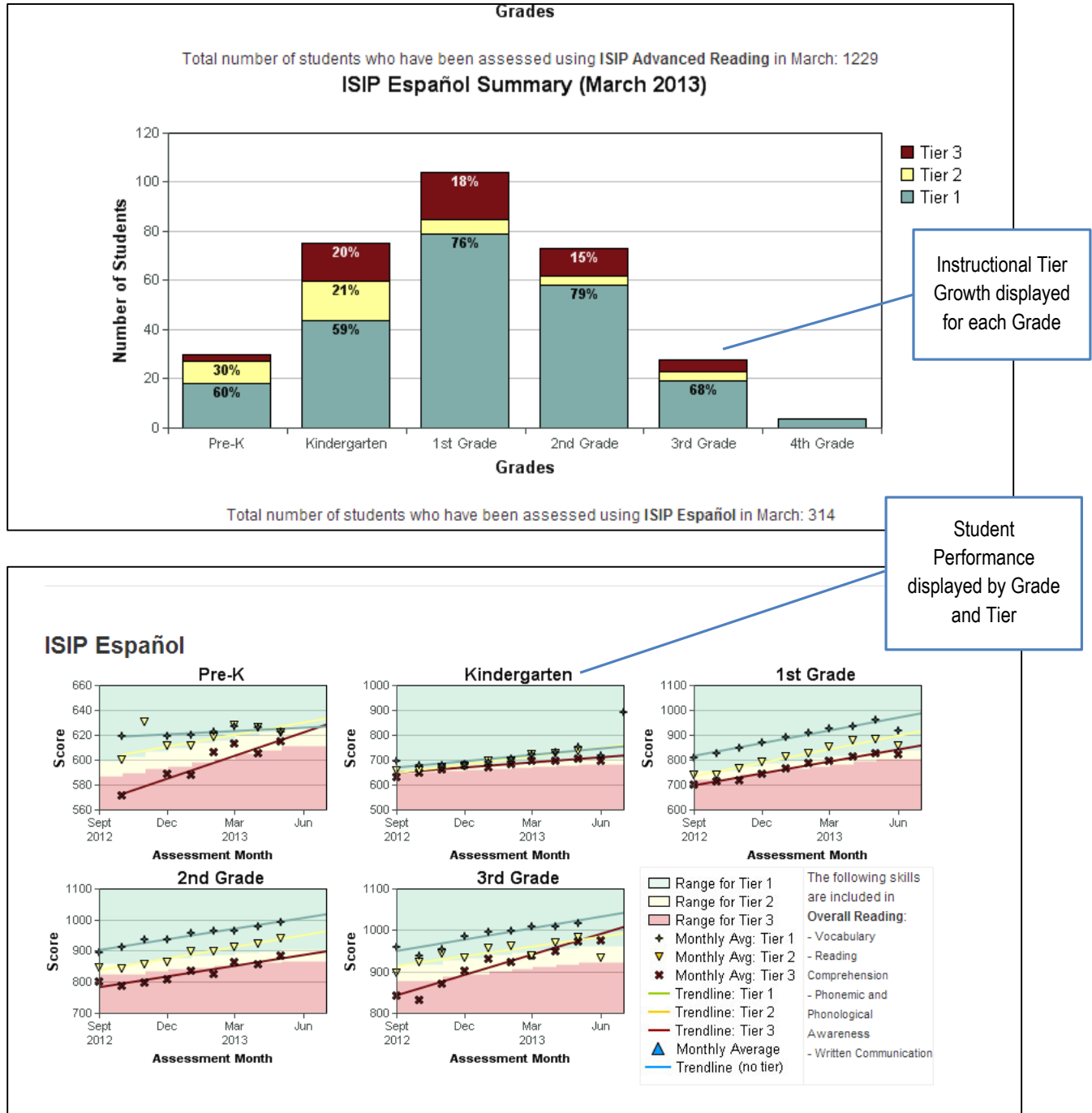
# Distribution Report

The Distribution Report shows the number of students performing by ranges of ability scores. This report can be viewed by overall ability and individual subtests. Ability indices, instructional tiers, and percentile ranks are listed in a table below the graph. This report can be used to observe the shape of the distribution and to identify groups of students in need of additional support.
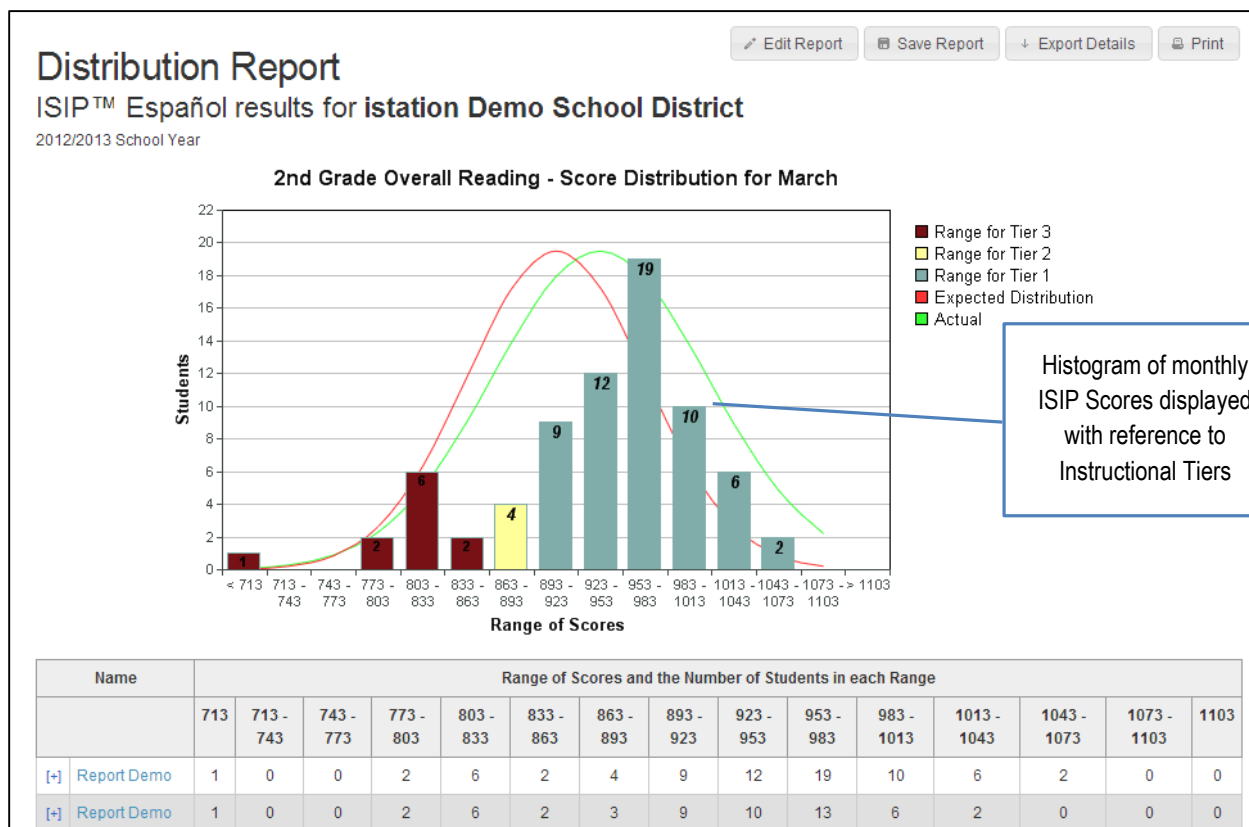
## Summary Report

The Summary Report shows the number and percentage of students at each of three instructional tiers: Tier 1 – no risk (above the 40th percentile), Tier 2 – some risk (between the 21-40th percentile), and Tier 3 – at risk (20th percentile and below). This report may be used by district administrators, principals, or teachers to project year-end outcomes and to judge the effectiveness of instruction. The Summary Report can also be used by administrators to determine which principals and teachers face the greatest challenges. This information can aid in making important decisions about the best use of resources, including the need for professional development.
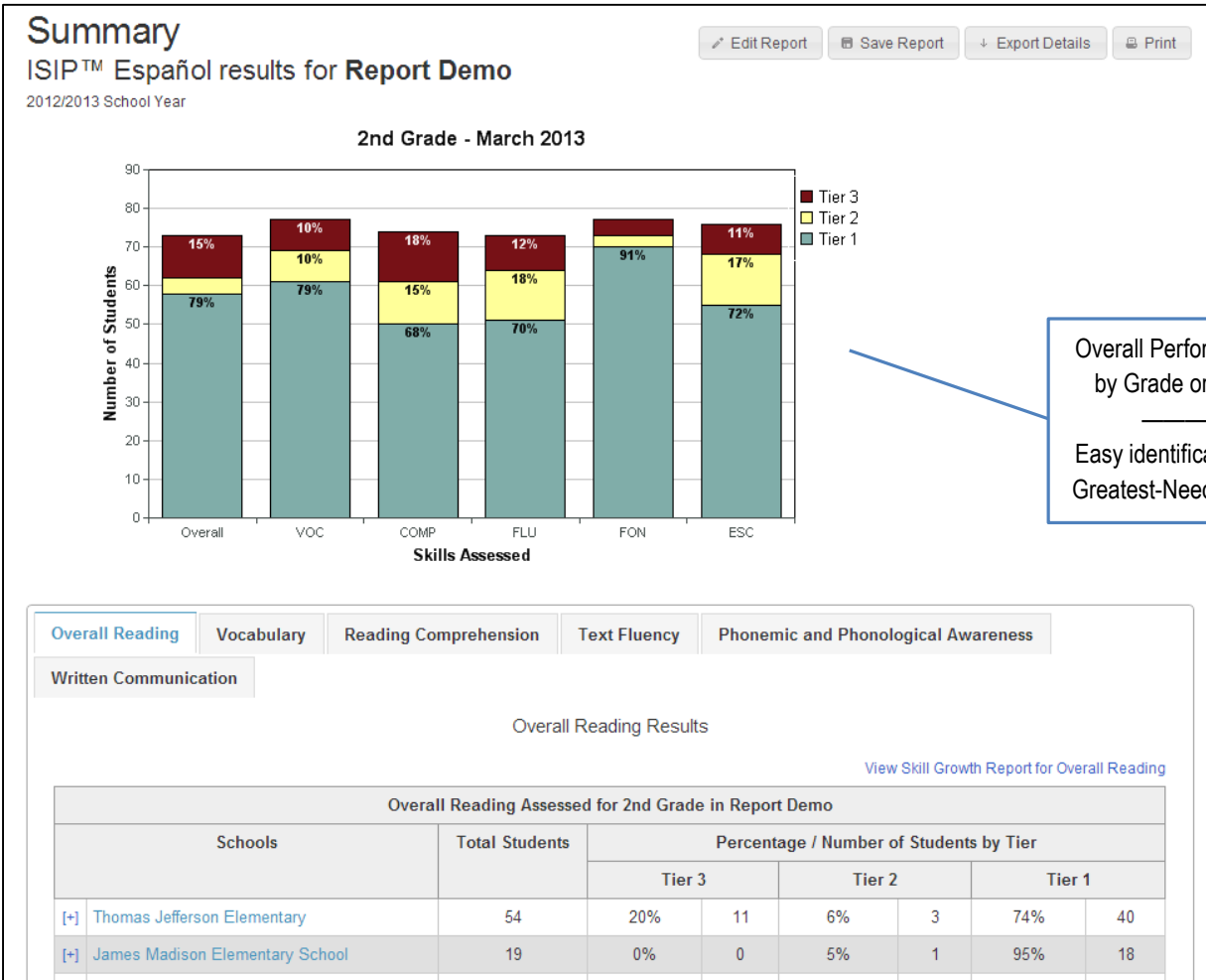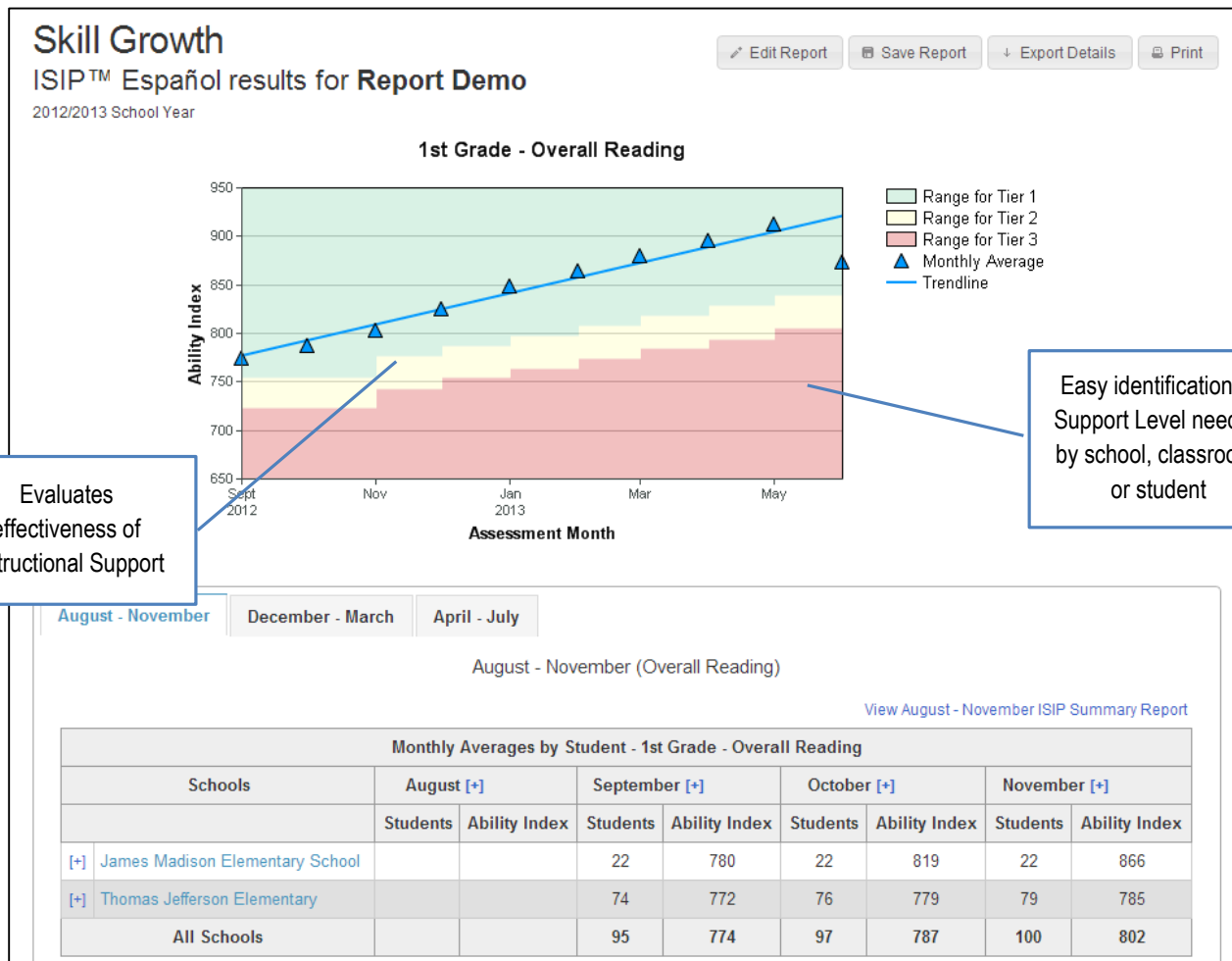
## Tier Movement Report

This report shows a comparison of the number and percentage of students who were categorized at each instructional tier of Tier 1, Tier 2, and Tier 3 through the current month. Assessments are given each month to monitor growth in critical skills.  This report is used to evaluate student growth over the school year.
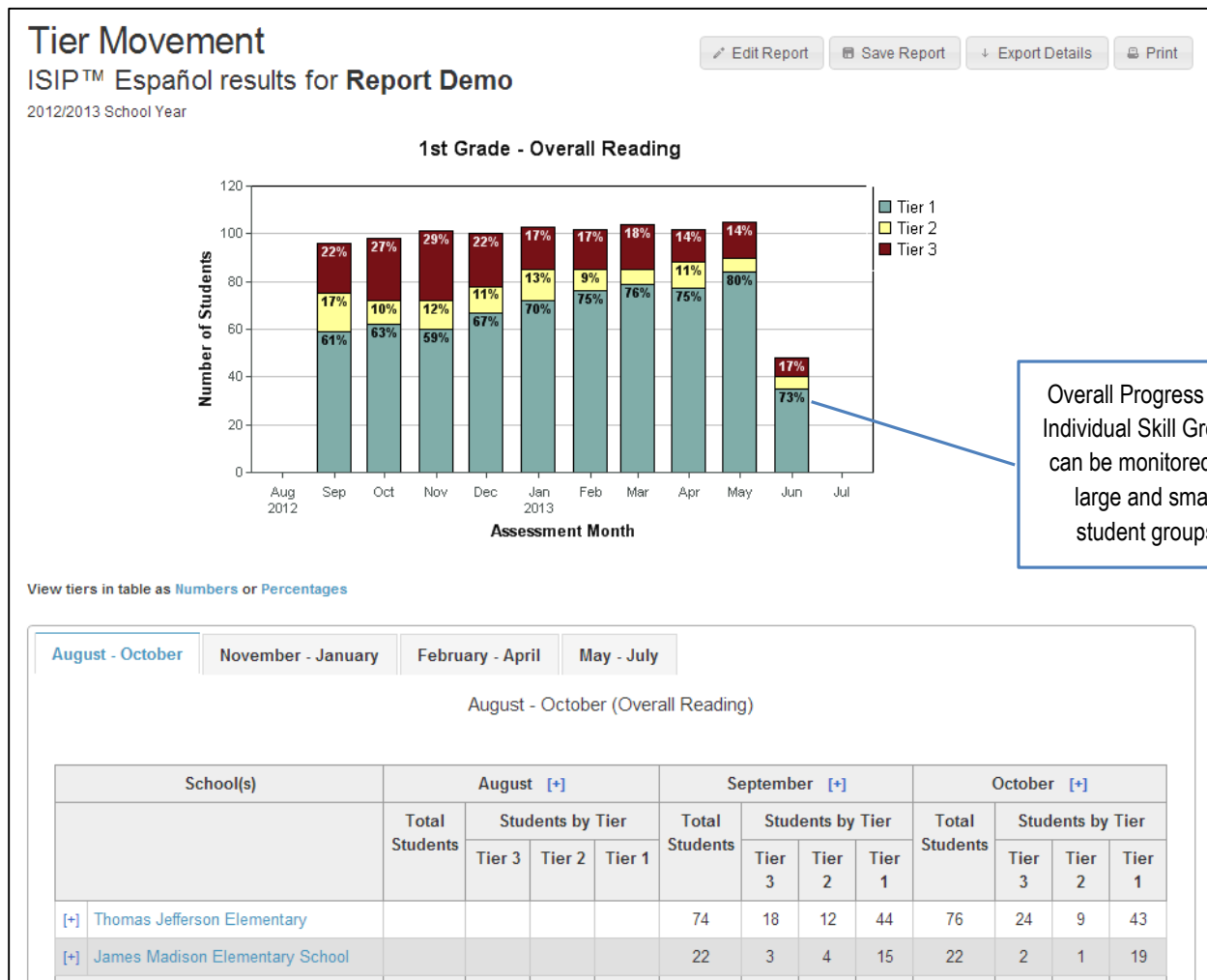
# Skill Growth Report

This report shows the progress made in each skill for all assessment periods to date. Progress is measured against performance goals. This report provides an excellent visual representation of the level of support needed.

This report may be used by district administrators, principals, and teachers to evaluate instructional supports and determine if modifications to the instructional plan should be considered. If progress is below goal for several consecutive assessments, the instructional plan should be re-evaluated. Only when progress exceeds goal are the instructional supports considered sufficient. This report is used to monitor the classroom's progress in skill acquisition, determine the need for whole-group instruction, identify the level of student support needed, evaluate the effectiveness of instructional support, and discuss student performance in Parent/Teacher conferences.



## Skill Growth
### ISIP™ Español results for **Report Demo**
2012/2013 School Year

Edit Report | Save Report | Export Details | Print

**1st Grade - Overall Reading**

Legend:
- Range for Tier 1
- Range for Tier 2
- Range for Tier 3
- ▲ Monthly Average
- — Trendline

Evaluates effectiveness of Instructional Support

Easy identification of Support Level needed by school, classroom, or student

August - November | December - March | April - July

August - November (Overall Reading)

View August - November ISIP Summary Report

| Schools | August [+] | | September [+] | | October [+] | | November [+] | |
|---|---|---|---|---|---|---|---|---|
| | Students | Ability Index | Students | Ability Index | Students | Ability Index | Students | Ability Index |
| [+] James Madison Elementary School | | | 22 | 780 | 22 | 819 | 22 | 866 |
| [+] Thomas Jefferson Elementary | | | 74 | 772 | 76 | 779 | 79 | 785 |
| **All Schools** | | | 95 | 774 | 97 | 787 | 100 | 802 |

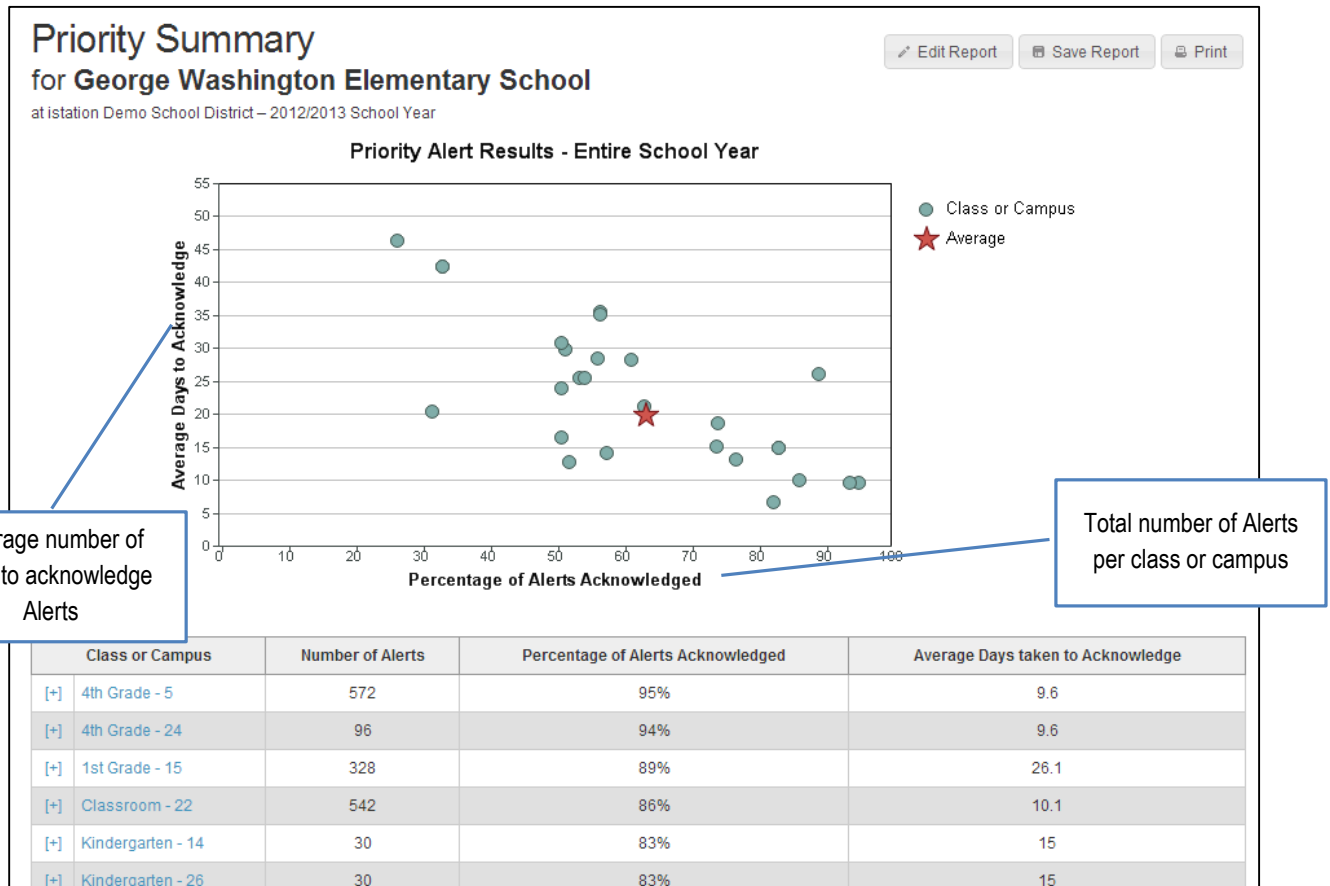Monthly Averages by Student - 1st Grade - Overall Reading

## Skill Growth by Tier Report

The Skill Growth by Tier Report shows how students identified in each tier at the beginning of the year progress in each skill assessed as a group. Even if students change tier classification individually, their group designation for this report is based on their first assessment so that this report accurately reflects the progress of each tier group based on who was in that group at the beginning of the year.  The values plotted on the graph are the average student performance for Tier 1, Tier 2, and Tier 3 students. This report is used to monitor the classroom's tier movement by skill and overall reading ability, monitor the classroom's progress in skill acquisition, identify the level of student support needed, and evaluate the effectiveness of instructional support.
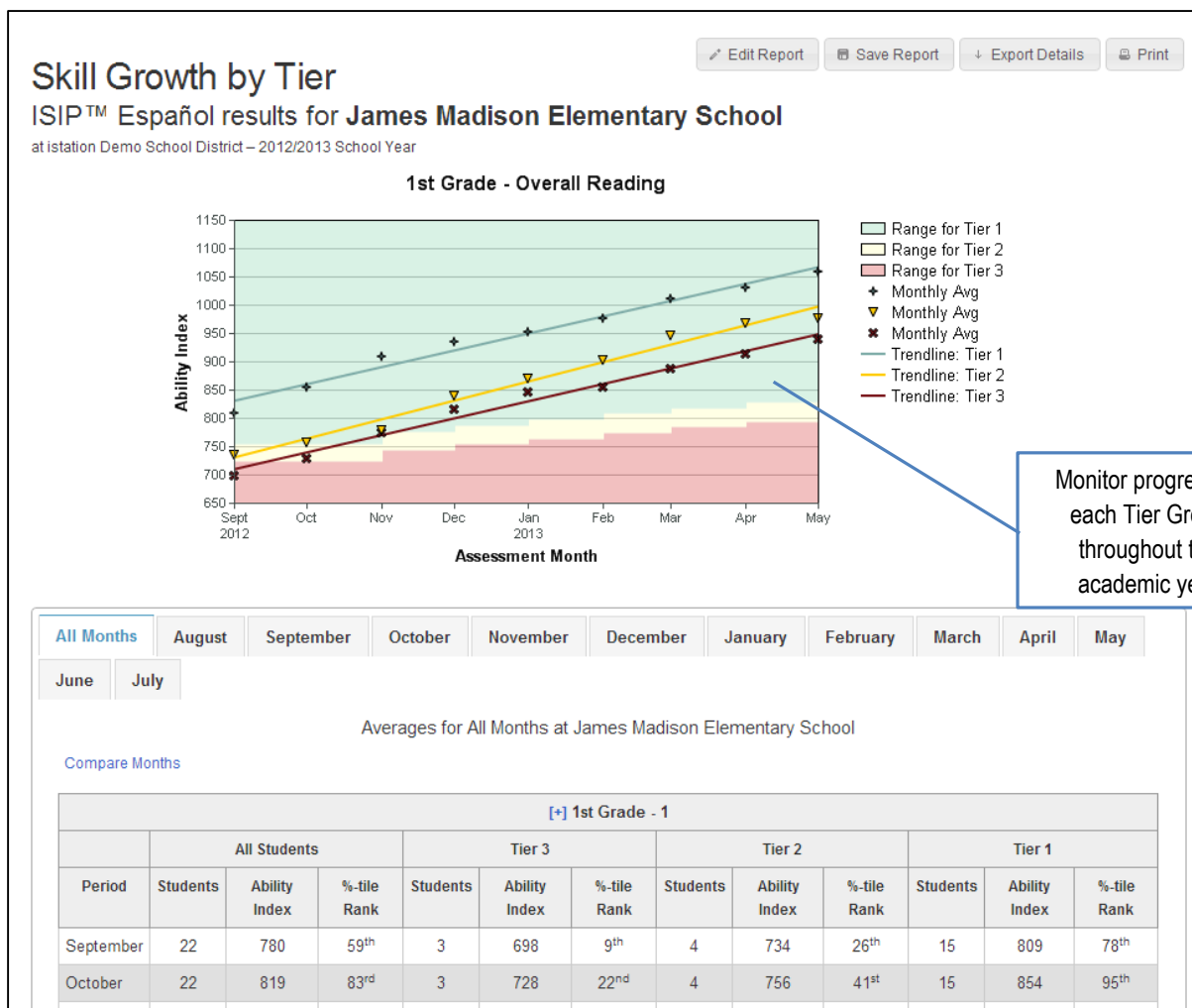
# Priority Summary Report

The Priority Summary Report, available to manager level users only, summarizes the use of the Priority Report (see description below) by averaging the number of Priority Report alerts and how many days it has taken to acknowledge student alerts on the Priority Report.



| Class or Campus | Number of Alerts | Percentage of Alerts Acknowledged | Average Days taken to Acknowledge |
|---|---|---|---|
| [+] 4th Grade - 5 | 572 | 95% | 9.6 |
| [+] 4th Grade - 24 | 96 | 94% | 9.6 |
| [+] 1st Grade - 15 | 328 | 89% | 26.1 |
| [+] Classroom - 22 | 542 | 86% | 10.1 |
| [+] Kindergarten - 14 | 30 | 83% | 15 |
| [+] Kindergarten - 26 | 30 | 83% | 15 |

# Classroom and Student Level Reports

## Priority Report

This report automatically alerts teachers to students in need of instructional support. Students are grouped according to risk level and skill need. Links are provided to teacher-directed plans of instruction and downloadable lessons and materials appropriate for each group. When student performance on assessments is below goal for several consecutive assessment periods, teachers are further notified. This is done to raise teacher concern and signal the need to consider additional or different forms of instruction. Where students have not participated fully in the assessment plan or are non-responsive to intervention and continue to show weakness, recommendations may be made to consider the use of diagnostic tests.

A complete history of Priority Report notifications, including those from the current year and all prior years, is maintained for each student. This report has a feature with which teachers may acknowledge that suggested interventions have been provided. A record of these interventions is maintained with the student history as an Intervention Audit Trail. This history can be used for special education Individual Education Plans (IEPs) and in Response to Intervention (RTI) models of instruction. The combination of progress monitoring data and a record of specific interventions proves to be a practical, clear picture of how a student is responding to intervention.

## Priority Report—Student Intervention History

This report is a history of identified skill weaknesses for a student, including those from the current and prior school years. The recommended teacher-directed lessons for intervention are listed, along with the level of difficulty the student had with the identified skill or skills.

If a recommended teacher-directed lesson was delivered as an intervention and the teacher clicked the Intervention Lesson Delivered button on the Priority Report, the date will be listed. Teachers also have the option of adding an intervention note. This optional note is an opportunity for teachers to give additional information about student progress and interventions delivered for RTI purposes. This type of anecdotal record can be beneficial to those evaluating a student's overall instructional plan.

## Student Summary Handout

This report provides a summary of student performance for the current school year. All completed ISIP assessments, all cycle-based curriculum assessments and practice activities, current Priority Report alerts, Lexile Reader measure, and usage information are all provided on this report.

This report is used to evaluate the student intervention plan, identify student skill weaknesses, discuss student performance with administrators, and plan for Parent/Teacher conferences.

# Navigating ISIP Español Reports

ISIP Español reports are immediately accessible online at www.istation.com to administrators and teachers by logging in with their unique username and password.

Upon login, administrators and teachers have the option to view the ISIP Español Reports Homepage. This page provides an overview and easy access to all reports available on the Istation Reports website. Descriptions and thumbnail images are available to help direct users to the desired report.
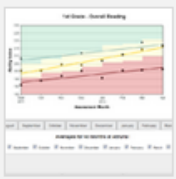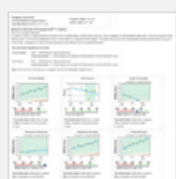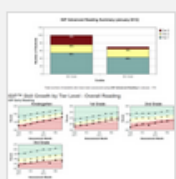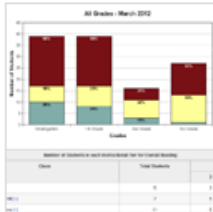
# Accessing Downloadable Lessons

Teachers can access recommended teacher-directed lessons by clicking links to lessons under the Recommended Teacher-Directed Lessons headings on the Priority Report. Additional teacher-directed plans of instruction and downloadable lessons and materials are available in the Teacher Resources section of the Istation Reports website.

## Priority
### for 1st Grade - 94
at Thomas Jefferson Elementary – 2012/2013 School Year

View Legend

*See an overview of current groups for this class...*

Edit Report | Save Report | Print

### ISIP Español: Destreza fonológica y fonética (FON) Group
Recommended Teacher Directed Lesson:

- Destreza fonológica y fonética

| | Students in this Group | Tier | FON Status | Overall Reading Tier | Date Listed | Usage Since this Alert (hh:mm) | Current Cycle |
|---|---|---|---|---|---|---|---|
| ☐ | Bryan [+] | 2 | ▼ | 2 | Tue Jun 4 2013 | 05:25 | 2 |
| ☐ | Alfredo [+] | 2 | ▼ | 2 | Mon Jun 3 2013 | | |

Intervention Lesson Delivered

Recommended Instructional Plan available for Download and Immediate Use

### ISIP Español: Vocabulario (VOC) Group
Recommended Teacher Directed Lesson:

- Vocabulario

| | Students in this Group | Tier | VOC Status | Overall Reading Tier | Date Listed | Usage Since this Alert (hh:mm) | Current Cycle |
|---|---|---|---|---|---|---|---|
| ☐ | Alan [+] | 3 | ☒ | 2 | Tue Jun 4 2013 | 00:00 | 4 |
| ☐ | Christopher [+] | 2 | ▼ | 1 | Mon Jun 3 2013 | 06:41 | 2 |
| ☐ | Alfredo [+] | 2 | ▼ | 2 | Mon Jun 3 2013 | 03:54 | 3 |

Intervention Lesson Delivered

# Chapter 4: IRT Calibration and the CAT Algorithm of ISIP Español

The goals of this study are to determine the appropriate Item Response Theory (IRT) model, estimate item-level parameters, and tailor the Computer Adaptive Testing (CAT) algorithms, such as the exit criteria.

During the 2010-2011 school year, data were collected from Kingdergarten to Grade 3 students in six states. However, most of the testing was in Texas elementary schools. The testing was conducted in 37 school districts, covering 228 schools. Among those, 30 schools districts and 217 schools were in Texas. Table 4-1 shows number of studetns and the demographics of participating students.

Table 4-1: Demographics for Participating Students

| Demographic | n | % |
|---|---|---|
| Total Number of Students | 3,895 | |
| Gender | | |
| Male | 1,818 | 46.48 |
| Female | 1,729 | 44.39 |
| Missing/Unidentified | 348 | 9.13 |
| Enrolled in Special Ed. | | |
| Yes | 241 | 6.19 |
| No | 3,210 | 82.41 |
| Missing/Unidentified | 444 | 11.40 |
| Economic Disadvantage | | |
| Yes | 2,841 | 72.94 |
| No | 610 | 15.66 |
| Missing/Unidentified | 444 | 11.40 |
| English Proficiency | | |
| Non-English Speaker | 1,421 | 36.48 |
| Fluent English Speaker | 2 | 0.05 |
| Limited English Speaker | 2,034 | 63.47 |

Students were escorted by trained SMU data collectors, typically graduate students, project coordinators and/or research associates, in convenience groupings to the school's computer lab for 30-minutes sessions on the ISIP Español program.

It was unrealistic to administer all the items to each student participating in the study. Therefore, items were divided into grade-specific subpools. Each participant was administered all of the items in the subpool for their grade level. Originally, 2,751 items were tried out. Table 4-2 shows the numbers of items in each grade subpool, not including the 10% overlap items.

Table 4-2: Items Used in Study

| | Grade | | | |
|---|---|---|---|---|
| | K | 1 | 2 | 3 |
| Comprensión de lectura | 155 | 135 | 136 | 201 |
| Escritura | - | 104 | 104 | 169 |
| Fonología y fonética | 419 | 471 | - | - |
| Vocabulario | 164 | 170 | 295 | 228 |

To control for order main effects, participating students were assigned items from their grade subpool in random order until they had answered all of the items in the subpool. The total number of sessions required to answer all items varied by participant.

# Data Analysis and Results

Due to the sample size for each item, a 2-parameter logistic item response model (2PL-IRT) was posited. We define the binary response data, $x_{ij}$, with index i=1,...$n$ for persons, and index $j$=1,...$J$ for items. The binary variable $x_{ij}$ = 1 if the response from student $i$ to item $j$ was correct and $x_{ij}$ = 0 if the response was wrong. In the 2PL-IRT model, the probability of a correct response from examinee $i$ to item $j$ is defined as

$$P(x=1) = \frac{1}{1+e^{-a_j\,(\theta_i-b_j)}}$$

where $\theta_i$ is examinee i's ability parameter, $b_j$ is item $j$'s difficulty parameter, and $a_j$ is item $j$'s discrimination parameter.

To estimate the item parameters, BILOG-MG (Zimowski, Muraki, Mislevy, & Bock, 2003) was used. BILOG-MG uses marginal maximum likelihood estimation (MMLE) to maximize the person response vector across both the item difficulty and discriminability dimensions. For example, Equation 2 represents the probability of a response vector of dichotomous items, $X$, in an instrument of length $L$,

$$P(X \mid \theta, J) = \prod_{j=1}^{L} p_j^{x_j} (1-p_j)^{1-x_j}$$

where the probability of a set of responses is conditioned on the person's ability ($\theta$) and the matrix of item parameters, $J$ (i.e., the collection of a s and b s for each item, $j$). In MMLE, an unconditional, or marginalized, probability of a randomly selected person from the population with a continuous latent distribution is specified as an integral function over the population distribution (Bock & Aitken, 1981). Subsequently, the

resulting marginal likelihood function underwent maximum likelihood estimation (MLE) by BILOG-MG to generate item parameters.

Among 2,751 items, 2,419 items were within the desired range of the item difficulty (-3.50, 3.50). 331 items fell below the desired range of the item discrimination (greater than 0.50). Therefore, 2,088 items were used for the ISIP Español item pool.

Overall, most items are in good quality in terms of item discriminations and item difficulties. The reliability was computed from IRT perspective by using this formula; $\rho^2 = 1 - [\text{SE}(\theta)]^2$, where $\theta$ is the student ability. It is 0.850, indicating that ISIP Español is very reliable. The standard error of measurement (SEM) was also computed from IRT point of view. Since the ISIP Español scale score is $(20 * \theta) + 200$, $SEM(\theta) = 20 * \text{SE}(\theta)$. It is 7.748.

# CAT Algorithm

The Computerized Adaptive Testing (CAT) algorithm is an iterative approach to test taking. Instead of giving a large, general pool of items to all test takers, a CAT test repeatedly selects the optimal next item for the test taker, bracketing their ability estimate until some stopping criteria is met.

The algorithm is as follows:

1. Assign an initial ability estimate to the test taker

2. Ask the question that gives you the most information based on the current ability estimate

3. Re-estimate the ability level of the test taker

4. If stopping criteria is met, stop. Otherwise, go to step 2

This iterative approach is made possible by using Item Response Theory (IRT) models. IRT models generally estimate a single latent trait (ability) of the test taker and this trait is assumed to account for all response behavior. These models provide response probabilities based on test taker ability and item parameters. Using these item response probabilities, we can compute the amount of information each item will yield for a given ability level. In this way, we can always select the next item in a way that maximizes information gain based on student ability rather than percent correct or grade-level expectations.

Though the CAT algorithm is simple, it allows for endless variations on item selection criteria, stopping criteria and ability estimation methods. All of these elements play into the predictive accuracy of a given implementation and the best combination is dependent on the specific characteristics of the test and the test takers.

In developing Istation's CAT implementation, we explored many approaches. To assess the various approaches, we ran CAT simulations using each approach on a large set of real student responses to our items. To compute the "true" ability of each student, we used Bayes expected a posteriori (EAP) estimation on all 700 item responses for each student. We then compared the results of our CAT simulations against these "true" scores to determine which approach was most accurate, among other criteria.

## Ability Estimation

From the beginning, we decided to take a Bayesian approach to ability estimation, with the intent of incorporating prior knowledge about the student (from previous test sessions and grade-based averages). In particular, we initially chose Bayes EAP with good results. We briefly experimented with Maximum Likelihood (MLE) as well, but abandoned it because the computation required more items to converge to a reliable ability estimate.

To compute the prior integral required by EAP, we used Gauss-Hermite quadrature with 88 nodes from -7 to +7. This is certainly overkill, but because we were able to save runtime computation by pre-computing the quadrature points, we decided to err on the side of accuracy.

For the Bayesian prior, we used a standard normal distribution centered on the student's ability score from the previous testing period (or the grade-level average for the first testing period). We decided to use a standard normal prior rather than using σ from the previous testing period so as to avoid overemphasizing possibly out- of-date information.

## Item Selection

For our item selection criteria, we simulated 12 variations on maximum information gain. The difference in accuracy between the various methods was extremely slight, so we gave preference to methods that minimized the number of items required to reach a satisfactory standard error (keeping the attention span of children in mind). In the end, we settled on selecting the item with maximum Fisher information. This approach appeared to offer the best balance of high accuracy and least number of items presented.

## Stopping Criteria

ISIP Español has a stopping criterion based on minimizing the standard error of the ability estimate.

### Production Assessment

Item types were grouped according to key reading domains for the production assessment. Each grade level is given the same set of subtests.

These subtests are administered sequentially and treated as independent CAT tests. Items are selected from the full, non-truncated, item pool for each subtest, so students are allowed to demonstrate their ability

regardless of their gradelevel. Each subtest has its own ability estimate and standard error, with no crossing between the subtests. After all subtests are complete, an overall ability score is computed by running EAP on the entire response set from all subtests. Each subtest uses its own previous ability score to offset the standard normal prior used in EAP.

Scale scores used in the reporting of assessment results were constructed by a linear transformation of the raw ability scores (logits). The study resulted in a pool of 2,088 Grades K-3 items with reliable parameter estimates aligned on a common scale, with the majority of items ranging from 140 to 289 in difficulty.

After completing this study, which included determining an appropriate IRT model, calibrating the items, and constructing the CAT algorithm, the ISIP Español assessment went into full production starting in the 2012-2013 school year.

# Chapter 5: Reliability and Validity of ISIP Espanol for Grades K–3

## ISIP Español Validity Framework

The Istation ISIP Español assessment is designed to be a criterion-referenced assessment to measure specific skills in early Spanish literacy. It has been developed to be used for formative purposes and progress monitoring. These purposes will be supported to the degree to which the criterion-referencing is supported by the evidence. Aspects of the development of ISIP Español and proposed claims from test scores are described in previous chapters of this document.

To support further development of the assessment and begin the validation process, this chapter describes the following:

1. A summary of the validity as argument framework employed to provide validity evidence
2. A summary of the proposed claims from ISIP Español
3. A description of the ISIP Español pilot validity study, preliminary results, and ongoing analyses

## Validity and Validation

Current definitions of validity vary across fields. However, in educational testing, most agree with the framework described in the *Standards for Educational and Psychological Testing* (hereafter referred to as *Testing Standards*; AERA, APA, NCME, 1999). "Validity refers to the degree to which evidence and theory support the interpretations of test scores entailed by proposed uses of tests" (AERA, APA, NCME, 1999, p. 9). The *Testing Standards* describes validation as the process of gathering evidence to achieve these goals, including evidence related to

- The construct
- Test content
- Response processes
- Internal structure
- Relations to other variables
- Intended and unintended consequences

In all cases, validation is an ongoing process, and the most important sources of validity evidence are those that are most closely related to the immediate inferences and proposed claims we make regarding test results. What evidence do we need to support the intended meaning of ISIP Español results?

## Validity as Argument

The argument approach to validation (Kane, 1992, 2006a, 2006b) is a contemporary approach that does not rely on our ability to define a construct or specify relations between the construct of interest and other important constructs. The heart of this approach is to make explicit arguments regarding proposed interpretations and uses. This is accomplished through an interpretive argument that specifies the inferences and assumptions leading from the test scores to the interpretations and decisions generated (Kane, 2006a). The validation process must evaluate and articulate the interpretive argument, specifying the reasoning from the score to the intended conclusions and the plausibility of the associated inferences and assumptions. The validity argument provides not only an evaluation of the proposed interpretations and uses of scores, but also alternative interpretations.

The forms of validity evidence described by the *Testing Standards* can be used in the validity argument framework; these include claims, intended inferences, and assumptions. These forms of evidence can be gathered to support the validity argument. These conceptualizations of validation are complimentary, providing the strongest approach to securing evidence to support that a measure is appropriate, meaningful, and useful.

## The Interpretive Argument

The first component of this process is clarifying the interpretive argument. This component frames the validation efforts by identifying the issues that need to be addressed. As Kane (2006b) describes, the interpretive argument provides three critical elements: (a) a framework to allow for the test development process to accommodate important assumptions and requirements that can be met in the design process, (b) a framework to clarify the validity argument by identifying the inferences and assumptions requiring evaluation, and (c) a framework for evaluating the validity argument by specifying the questions that need to be addressed. There are four important elements of the interpretive argument:

1. The conclusions and decisions to be made from test scores

2. The inferences and assumptions leading from test scores to the conclusions and decisions

3. The potential competing interpretations

4. Evaluative evidence for and against the proposed interpretive argument

# Proposed Claims of ISIP Español

An important tool in the specification of the interpretive argument is the clarification of the conclusions and decisions to be made and the intended inferences and associated assumptions. This includes identifying the proposed claims we hope to make based on ISIP Español results. The following list of proposed claims is derived from the evidence provided by the functionality and development of the English version of ISIP and potential reconditioning for use with reading objectives in other languages. These considerations include the following: (a) the primary objective of this version was to determine how to accurately measure,

on the computer, early reading skills known to be predictive of later reading success (Mathes, & Torgesen, 1996-1999); (b) the functionality of prototype tasks created for each subtest using a technology-based platform; and (c) the possibility to compare technology-based assessments to paper and pencil evaluations of a similar construct (ISIP was compared against CTOPP, TOWRE, and DIBELS ORFA). A number of these claims have been addressed throughout the development of ISIP Español, and some claims were addressed during the 2010-2011 pilot validation study. We recognize that validation is an ongoing process, and therefore we will continue to evaluate results in 2014 and beyond, as described in Chapter 6 of this document.

## Claims Regarding Spanish Early Literacy Evaluation

1. To measure achievement of selected Spanish language arts standards (domains), focusing on reading.

2. To measure whether specific early literacy skills in Spanish (subtests) have been achieved.

3. To measure students' knowledge, skills, and performance level in the domains of Spanish reading that apply to each grade level, including:

   - Phonemic awareness and grapheme/sound correspondence

   - Oral language and listening comprehension

   - Vocabulary

   - Reading comprehension thinking skills

   - Written communication

   - Text fluency

4. To determine the progress students make in Spanish reading instructional programs (also described as progress monitoring of Spanish foundational reading skills).

The assumptions regarding the degree to which domains are appropriately defined and to which the item types are appropriate measures of the knowledge, skills, and abilities targeted in each domain are included in Chapter 1 of this document.

There is an assumption regarding the representativeness of the models, based on Spanish Language Arts and Reading (SLAR) standards, which suggests that they provide an appropriate context for the intended purposes of ISIP Español.

More specifically, there is an assumption that standards and domain definitions, based on SLAR standards from selected states and countries (i.e., California, Texas, WIDA consortium, Puerto Rico, Colombia, Mexico, and Spain), are appropriate for the proposed purposes and the intended population. This research-based evidence is thoroughly presented and reviewed in the description, domains and item development sections in Chapter 1 of this document.

## Claims Regarding Conclusions and Decisions

1. To report domain scores with sufficient precision to warrant independent scoring and reporting of distinct aspects of Spanish literacy.

2. To be functional for the purposes of formative assessment.

3. To determine whether students are progressing toward end-of-year expectations and achieving selected skills for each domain.

4. To provide information on how groups of students are progressing toward achieving grade level expectations in each domain.

These claims depend on the plausibility of the first four claims and associated assumptions regarding content, described above. In addition, there are assumptions about the appropriateness of performance standards (end-of-year expectations and achieving expected levels of performance).


# The ISIP Español Pilot Validity Study

Istation is completing a validity and reliability study during the 2010-2011 school year. The scope of the validity and *reliability study is aimed at answering the following questions: (1) Do the scores produced by* ISIP Español show evidence of reliability, including internal consistency, and alternate form reliability? (2) Do the scores produced by ISIP Español show evidence of validity, including concurrent validity and predictive validity? (3) Do the scores produced by ISIP Español show evidence of accurate classification, as established by ROC analysis? The development of a technical report will be completed once the validity study is finished. The first part of the study, which includes content validity analysis, has been completed as described in this chapter.

In addition to the validity and reliability study, a concordance analysis also has been conducted, whereby the results of students assessed on ISIP Español were compared to results on other external measures obtained by the same group of students. The external measures selected for this analysis include Evaluación del desarrollo de la lectura (EDL2, Pearson), *Téjas* LEE (Brookes) for Grades Kindergarten through Grade 2, and Texas Assessment of Knowledge and Skills in Reading (TAKS Reading, TEA). This part of the research will not be completed until after results are obtained from 2011 administration of the TAKS and all data have been analyzed.

# Core Elements of the Validity Study, Evidence, and Analyses

It is important to recognize that validation is an ongoing process. Validity is not something that is present or absent; validity is the accumulation of evidence that supports ongoing and current (and new) interpretations and applications of test results. The analyses will be described in terms of the inferences that are important for a standardized measure of observable attributes, in this case ISIP Español. These four inferences (scoring, generalization, extrapolation, and implication) will provide the framework that encompasses all elements of the current validation design, analyses, and anticipated results (see Technical Note on Content-Related Validity Evidence, following this section).

**Scoring (supporting the inference from observations of performance to an observed score)**

Since the test is a standardized, objectively-scored instrument, the inference regarding scoring that supports the content-validity argument of ISIP Español is relatively easy to achieve. [Claim 5]

- Key confirmation

- Internal consistency (reliability and score accuracy), item discrimination

***Generalization* (moving from an observed score on a sample of tasks to an expected score on the universe of generalization)**

Since the test is based on relatively narrow subdomains with relatively homogenous items, this inference is fairly direct. [Claim 1]

- Alternate form correlations

- Documentation of construct representation

- Item discrimination

***Extrapolation* (moving from the universe score to the target score)**

Since the test is based on a broad range of language arts and reading standards, the degree to which the domain is covered by the represented skills is more challenging to support. [Claims 2, 3, 6]

- Correlations with criterion variables in same target domain (EDL; *Téjas* LEE, administered 3 times during year; TAKS, the spring state standards-based assessment)

- Documentation of content coverage, given review of existing standards

***Implication* (the translation of the estimated target score into a description of knowledge, skills, or ability level).**

The test is designed for multiple purposes, so the clarity of the attribute being measured and how these can be described is critical. [Claims 4, 6, 7, and 8]

- Performance level descriptors

- Instrument development process

- Agreement among users and relevant content experts

- Standard error of measurement, score precision, and test information functions

## Analysis Methods

To support the analyses of test scores from the ISIP Español pilot study, based on the relevant inferences described above, the following analyses of pilot data were conducted or are planned:

- Winsteps was used to provide classical test theory item statistics, including classical item difficulty (item p-values, proportion correct) and discrimination (point-biserial item-total correlations).

- Winsteps was used to provide a Rasch analysis of item performance, including item fit measures.

- With the scores on the 13 forms administered to a subsample within grade, inter-form correlations and form difficulty (means) and variability (variance) were examined.

- To the extent possible from the larger sample, Mplus will be used in subsequent studies to conduct confirmatory factor analysis to test the degree to which forms are unidimensional (on forms with large enough samples) versus multidimensional, based on the reading subdomains.

- Finally, once criterion measures are administered and scored, correlations between ISIP Español forms and criterion measures will be assessed (with attention to intended alignment of constructs across measures). This will be accomplished after TAKS results are analyzed.

## Study Sample and Form Design

The sample was obtained through careful selection of students from El Paso Independent School District to include a full range of ability levels among Spanish speaking students in grades K through 3, including 219 students with valid responses, as shown in Table 5-1.

Table 5-1: Validation Sample Size by Grade and Percentage Female

| Grade | n | % Female |
|-------|-----|----------|
| K | 52 | 56% |
| 1 | 56 | 43% |
| 2 | 52 | 46% |
| 3 | 59 | 54% |

For each grade level, 13 forms were constructed to cover parallel content. The 13 forms were administered online within three to four weeks, in random order across students.

Table 5-2: Skill Areas Assessed by Grade

| Grade | Skill Area | | | | |
|---|---|---|---|---|---|
| | Listening Comp | Phonological Awareness | Reading Comp | Vocabulary | Reading Fluency |
| K | ✓ | ✓ | ✓ | ✓ | |
| 1 | | ✓ | ✓ | ✓ | |
| 2 | | | ✓ | ✓ | ✓ |
| 3 | | | ✓ | ✓ | ✓ |

A total of 3,832 items were used across skill areas on the 13 forms.

## Scoring

Reading Fluency was assessed by a 90-second timed maze task and was scored with an algorithm that accounts for (a) number of tasks completed within the 90-second limit and (b) accuracy of responses. All other skill areas were scored in terms of percent correct.

# Score Reliability (Inferences Regarding Scoring, Implication)

Coefficient alpha is a typical form of reliability, which, under specific assumptions of the parallel measurement model, provides an estimate of item internal consistency. These coefficients are presented here for preliminary consideration only. In this context, results are promising. Test score reliability is a form of validity evidence, as it informs the precision of scores and supports related inferences.

## Rasch Model Reliabilities of the Item Pools

Another index of reliability comes from the Rasch analyses of each measure. In order to accomplish the Rasch analyses, all items across the 13 forms were combined to improve estimation of item functioning. This model considers the 13 forms to be samples of a larger pool of items, which is consistent with the future intent to create online adaptive forms. These estimates of reliability are based on (1) an estimate of true variance (model-based score variance), (2) a measurement error variance (based on the theoretical definition of reliability as the ratio between true-score variance), and (3) an observed-score variance (the proportion of observed variance that is true). Rasch analysis required larger samples, so the results presented in Table 5-3 should be interpreted strictly as preliminary.

Person Reliability, which is similar to traditional test-score reliability, indicates the capacity of the sample to generate a stable ordering of person abilities based on their test scores. Low person reliabilities among a pool of items suggest a high degree of randomness in responses (guessing).

Item Reliability, which has no traditional equivalent, indicates the capacity of the sample to generate a stable ordering of item difficulties.

Table 5-3: Person Reliability and Item Reliability

| Skill Area | Grade | *n* | # Items | Person Reliability | Item Reliability |
|---|---|---|---|---|---|
| Listening Comprehension | Kinder | 52 | 104 | .90 | .80 |
| Reading | Kinder | 52 | 104 | .48 | .88 |
| Comprehension | 1 | 56 | 130 | .83 | .74 |
| | 2 | 52 | 195 | .88 | .66 |
| | 3 | 59 | 260 | .89 | .77 |
| Reading | 2 | 52 | 585 | .92 | |
| Fluency | 3 | 59 | 503 | .94 | *1 |
| Phonological | Kinder | 52 | 351 | .96 | .83 |
| Awareness | 1 | 55 | 298 | .96 | .78 |
| Vocabulary | Kinder | 52 | 155 | .88 | .85 |
| | 1 | 55 | 194 | .92 | .91 |
| | 2 | 52 | 364 | .96 | .89 |
| | 3 | 59 | 310 | .92 | .90 |

These reliabilities are not associated with individual form-based scores. They provide an index of the measurement quality of the pool of items in each area, based on this specific sample. They might also be interpreted as a potential score reliability, based on the current pools of items.

## Within Skill-Level Analyses across Forms (Inferences Regarding Generalization)

Each skill area was assessed by 13 forms designed to cover parallel content. There are multiple indicators with respect to evaluating the degree to which forms are parallel or similar in means, variances, and total scores. As an initial set of analyses, the correlations between forms for each Skill area by Grade, are provided in Table 5-4. However, as the plans for ISIP Español include computer adaptive testing (CAT)

---

[1] Item reliability for Fluency needs to be measured differently because this is the only timed subtest in the battery and student responded to ⅓ - ½ of the total available items.

administration, the degree to which forms are parallel will become less important. All items will be placed on the same common scale, making items within domains exchangeable. The results reported in Table 5-4 include the analyses across all 13 forms.

Another indicator of the association among forms (or the degree to which forms are measuring a similar construct) is the correlation among form scores and the average across all forms (excluding the target form score). This analysis is more rigorous since it requires a student to respond to all 13 forms, and consequently the sample has been reduced.

Table 5-4: Analyses across Forms, Including Students Responding to All 13 Forms
Reading Comprehension Corrected Correlations between Form and Average Score by Grade

| Reading Comp Form | Grade K N=19 | Grade 1 N=18 | Grade 2 N=27 | Grade 3 N=29 |
|---|---|---|---|---|
| 1 | .428 | .421 | .522 | .696 |
| 2 | .278 | .221 | .628 | .270 |
| 3 | .387 | .459 | .638 | .472 |
| 4 | - | .474 | .631 | .661 |
| 5 | .436 | .623 | .435 | .726 |
| 6 | .032 | .348 | .503 | .836 |
| 7 | .619 | .646 | .682 | .756 |
| 8 | .427 | .670 | .698 | .766 |
| 9 | .166 | .484 | .669 | .854 |
| 10 | .553 | .301 | .110 | .782 |
| 11 | .243 | .456 | .679 | .834 |
| 12 | - | .482 | .458 | .744 |
| 13 | .506 | .643 | .790 | .410 |

An example interpretation from this table is: The correlation between the Reading Comprehension Form 1 score and the average Reading Comprehension score across all other forms in Grade 3 (excluding Form 1) is .696.

Figure 5-A: Reading Comprehension Form Mean 95% Confidence Intervals by Grade
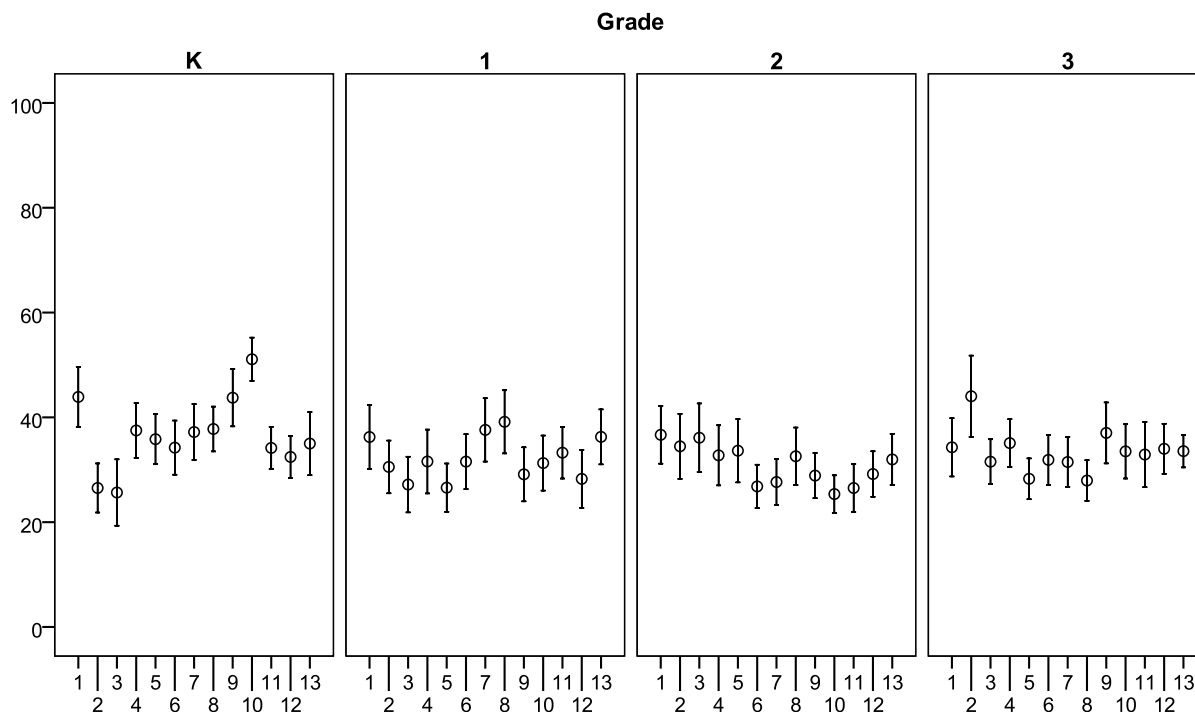


In this figure, the consistency in difficulty of forms can be seen graphically. Forms in Grade K show some degree of variability, compared to forms in other grade levels. In Grades 1 and 2, forms are within a narrow range of difficulty. In Grade 3, form 2 appears to be easier than some of the others.

The variability observed in the Reading Comprehension scores across forms in Kindergarten is likely due to the small number of items in each form (there are only 8 Reading Comprehension questions in each Kindergarten form). In measuring Reading Comprehension among Kindergarten students, results can be challenging to interpret due to students' ages and possible differences in instruction methods. Since the pilot was administered early in the fall, this may also indicate that Reading Comprehension could be more accurately measured beginning later in spring, not immediately at the beginning of the school year in the fall.

Reading Comprehension (among other grade levels) correlates quite strongly and fairly stably across forms, suggesting that the measures in each domain are strong across grades.

As is expected, Listening Comprehension scores are much more stable across forms for Kindergarten (See Table 5-5). These results coincide with instructional methodology and standards objectives for Kindergarten students in Texas.

Table 5-5: Listening Comprehension Corrected Correlations between Form and Average Score

| Form | Grade K N=18 |
|---|---|
| 1 | .532 |
| 2 | .711 |
| 3 | .108 |
| 4 | .696 |
| 5 | .417 |
| 6 | .753 |
| 7 | .619 |
| 8 | .747 |
| 9 | .824 |
| 10 | .793 |
| 11 | .783 |
| 12 | .786 |
| 13 | .768 |

Figure 5-B: Listening Comprehension Form Mean 95% Confidence Intervals by Grade
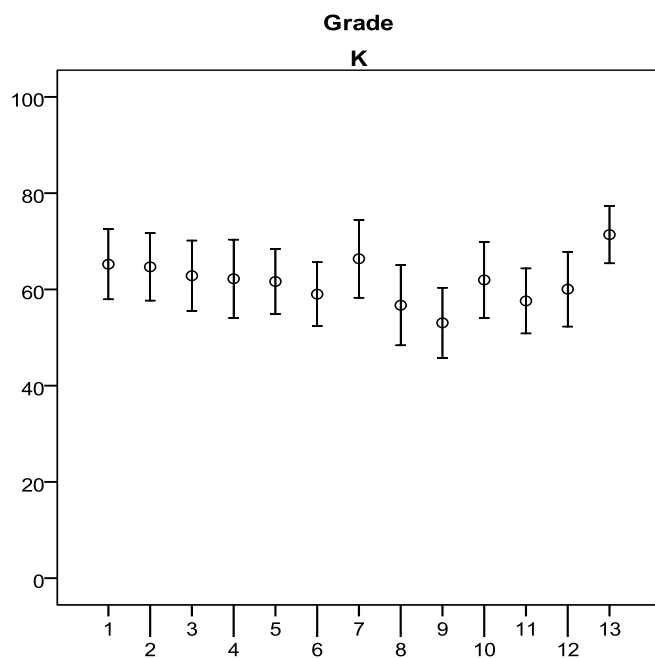
Table 5-6: Reading Fluency Corrected Correlations between Form and Average Score by Grade

| Form | Grade 2<br>N=27 | Grade 3<br>N=29 |
|------|------|------|
| 1 | .780 | .759 |
| 2 | .840 | .829 |
| 3 | .838 | .799 |
| 4 | .671 | .779 |
| 5 | .881 | .816 |
| 6 | .893 | .791 |
| 7 | .824 | .675 |
| 8 | .882 | .840 |
| 9 | .839 | .775 |
| 10 | .795 | .817 |
| 11 | .790 | .678 |
| 12 | .445 | .864 |
| 13 | .588 | .664 |

Figure 5-3: Reading Fluency Form Mean 95% Confidence Intervals by Grade
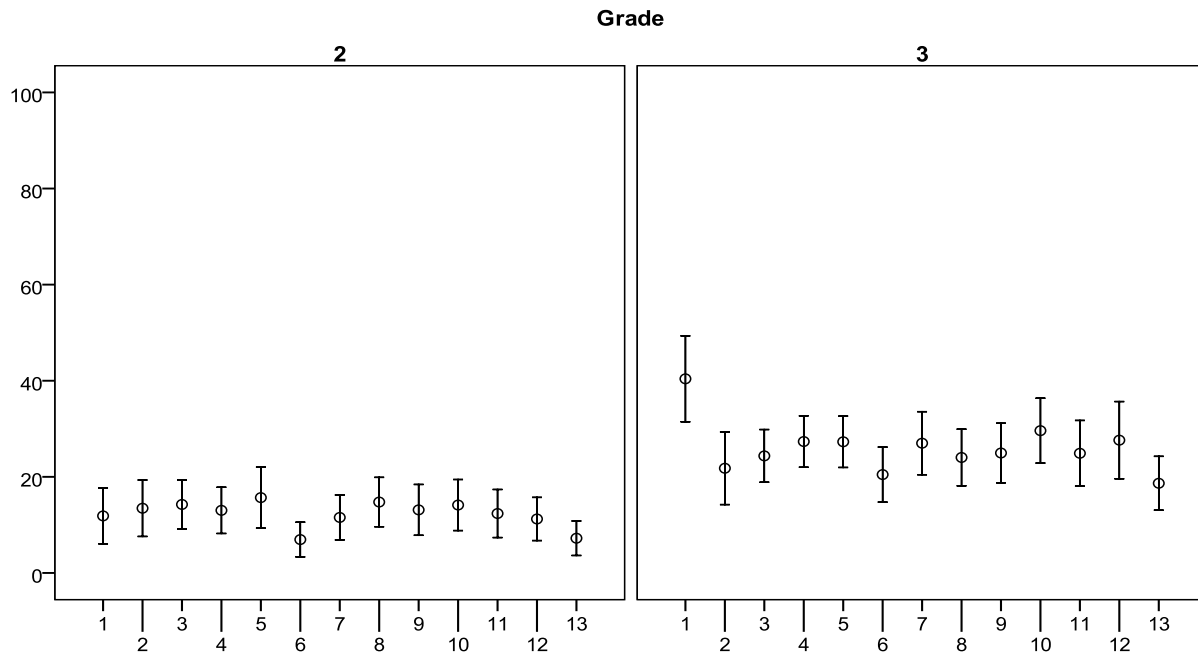
Table 5-7: Phonological Awareness Corrected Correlations between Form and Average Score by Grade

| Form | Grade K N=18 | Grade 1 N=19 |
|------|------|------|
| 1 | .913 | .860 |
| 2 | .964 | .898 |
| 3 | .765 | .932 |
| 4 | .739 | .929 |
| 5 | .727 | .855 |
| 6 | .684 | .825 |
| 7 | .916 | .786 |
| 8 | .890 | .891 |
| 9 | .876 | .821 |
| 10 | .882 | .801 |
| 11 | .921 | .885 |
| 12 | .824 | .867 |
| 13 | .928 | .876 |

Figure 5-4: Phonological Awareness Form Mean 95% Confidence Intervals by Grade
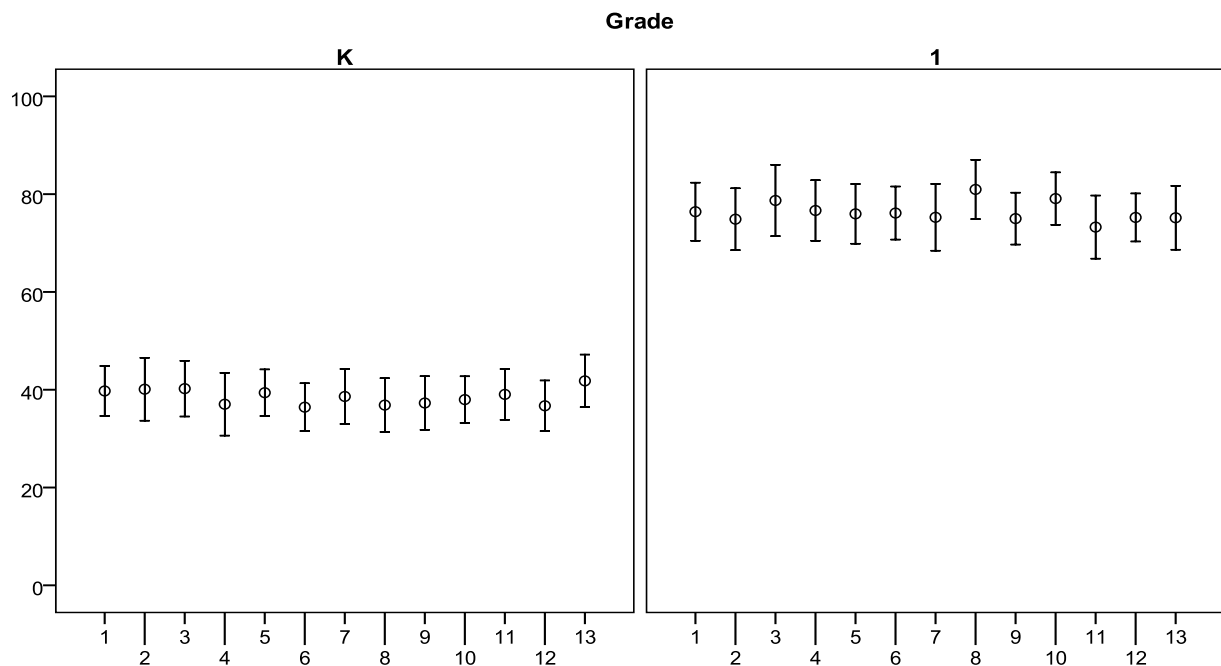
Table 5-8: Vocabulary Corrected Correlations between Form and Average Score by Grade

| Form | Grade K N=18 | Grade 1 N=18 | Grade 2 N=27 | Grade 3 N=30 |
|------|------|------|------|------|
| 1 | .748 | .892 | .693 | .716 |
| 2 | .718 | .584 | .749 | .564 |
| 3 | .665 | .787 | .739 | .661 |
| 4 | .454 | .820 | .806 | .768 |
| 5 | .131 | .637 | .815 | .799 |
| 6 | .728 | .654 | .718 | .779 |
| 7 | .486 | .662 | .785 | .812 |
| 8 | .588 | .430 | .782 | .870 |
| 9 | .438 | .723 | .788 | .676 |
| 10 | .715 | .566 | .757 | .770 |
| 11 | .564 | .905 | .731 | .671 |
| 12 | .864 | .521 | .663 | .691 |
| 13 | .634 | .791 | .889 | .824 |

Figure 5-5: Vocabulary Form Mean 95% Confidence Intervals by Grade
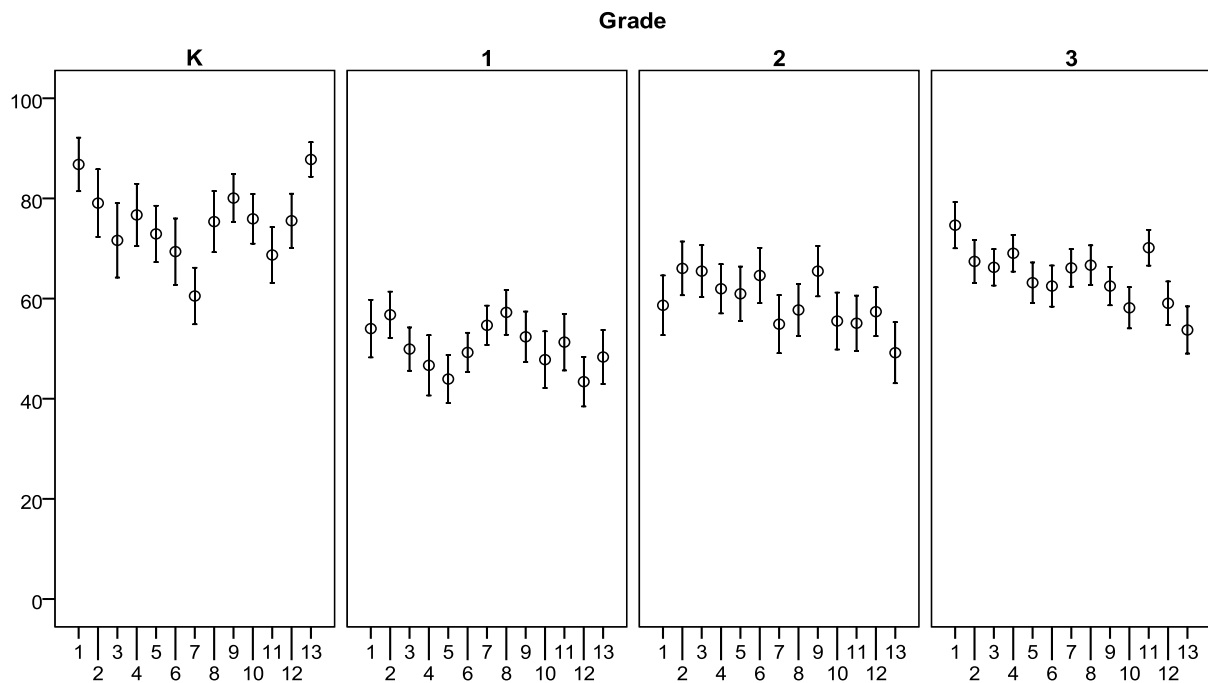
# Between Skill-Level Analyses across Forms (Inferences Regarding Generalizations)

Each of the 13 forms designed to cover parallel content were constructed to include grade-relevant skill areas. Each skill area is designed to assess an important component of reading readiness or reading itself, depending on grade level. As an initial set of analyses, the correlations between skill areas within forms by Grade are summarized in Table 5-9. In addition, an average skill area score was computed for each student to evaluate correlations between average performances among the skills assessed by grade, as shown in Table 5-10. Below, the average inter-skill correlations across forms are reported.

Table 5-9: Between Skill Correlations across Forms, by Grade

|  | Grade K | Grade 1 | Grade 2 | Grade 3 |
|---|---|---|---|---|
| Mean Correlation | .206 | .363 | .444 | .402 |
| Median Correlation | .197 | .348 | .489 | .386 |
| SD of Correlations | .145 | .151 | .166 | .151 |
| Minimum Correlation | .000 | .099 | .065 | .027 |
| Maximum Correlation | .491 | .662 | .692 | .631 |

The average inter-skill correlation across forms and grades is approximately .35. This indicates relatively unique skills, with approximately 12% shared variance. The skill areas appear to be measuring unique areas. The maximum correlation between any two skill areas across the forms is .69, indicating less than 50% shared variance.

Based on these findings, it is appropriate to use scores from ISIP Español to report skills independently, as intended. Research regarding Spanish-English bilingual development and assessment will drive further validity studies of ISIP Español domains and their inter-skill correlation to English reading (See Chapter 5). As reviewed by the Center for Early Education and Development at the University of Minnesota, English literacy development has similar relations between skill development in domains of oral language development, phonological awareness, and Spanish reading development (Farver et al., 2007; Gorman & Gillam, 2003; Signorini, 1997). Regarding national goals to improve the English proficiency of all children, a growing body of research provides evidence that Spanish-English Bilingual children's performances on Spanish early literacy measures predict later reading success (Cárdenas-Hagan, Carlson, & Pollard-Durodola, 2007; Cisero & Royer, 1995). Researchers have provided evidence of cross-linguistic transfer of early literacy skills, with higher achievement in Spanish phonological awareness, letter and word knowledge, print concepts, and sentence memory. Cross-linguistic transfer predicts improved reading achievement in English in Kindergarten and Grades 1, 3, and 4 (Lindsey, Manis, & Bailey, 2003; Manis, Lindsey, & Bailey, 2004). Early development of native oral vocabulary may also be related to improving English reading comprehension in elementary grades (Proctor, August, Carlo, & Snow, 2006). These

findings are consistent with the small to moderate correlations found with ISIP Español domains, such that unique information from each domain is relevant in understanding child development of readings skills in Spanish and English and their interrelations.

Average skill scores were computed by taking the average across forms for each student. The average skill scores across forms are much more stable, as they each comprise all of the items administered within a skill area. These average skill scores are correlated and summarized here.

The independent scoring and reporting of ISIP Español is an appropriate way to measure these skills.

Table 5-10: Correlations between Average Skill Scores across Forms, by Grade

|  | Grade K | Grade 1 | Grade 2 | Grade 3 |
|---|---|---|---|---|
| Mean Correlation | .456 | .678 | .688 | .638 |
| Median Correlation | .448 | .644 | .682 | .663 |
| SD of Correlations | .164 | .094 | .050 | .051 |
| Minimum Correlation | .275 | .606 | .641 | .580 |
| Maximum Correlation | .744 | .784 | .741 | .672 |

Generally, all of the skill areas resulted in higher inter-skill correlations by using average scores across forms. This is largely expected, due to the higher stability of average scores.

Among Kindergarten students, Phonological Awareness resulted in the lowest correlations among the skill areas. This reinforces the idea that Phonological Awareness is a unique skill, compared to the others.

Association between Skills

- The skill areas appear to be relatively independent and are appropriately reported separately.

- The average inter-skill correlations across forms and grades are highest in Grade 2 (about .44) and Grade 3 (.40), with lower correlations in Kindergarten (about .21). Correlations of average skill scores across forms are higher (.46 to .69).

- The correlations among average skill scores (scores combined across forms) were consistently higher, ranging from .46 (Kindergarten) to .69 (Grade 2). This suggests a moderate association among the skills being measured.

# Item-Level Analysis (Inferences Regarding Scoring, Generalization, Implication)

The data obtained at the item level, although it was collected from a small sample, provides initial information about item functioning. The item responses were analyzed using the Rasch model software Winsteps, which provides (1) classical test statistics of item difficulty (*p*-value), (2) discrimination (point-biserial correlation), and (3) Rasch item statistics (including item fit). Ideally, Infit and Outfit (z-scores) should be within -2 and +2; point-biserials should be positive; p-values should range between .2 and 1.0.

To maximize the information available, all items were analyzed concurrently within skill areas (items across all forms were analyzed simultaneously) by grade. Because of the timed nature of the fluency items, they were not included here.

As can be seen in the following tables (5-11–5-18), few items resulted in poor quality statistics (poor fit or discrimination) across skill area—generally less than 6%. No areas stood out in terms of Rasch item fit. With respect to item discrimination, Reading Comprehension resulted in a high number of items in the poorly functioning range: 43% of Kindergarten items, 21% of Grade 1 items, 21% of Grade 2 items, and 16% of Grade 3 items. In part, the Reading Comprehension result is due to the variable performance of items and forms in Kindergarten. This is also likely a result of having a fewer items in this skill area (only 8 in Kindergarten and 10 in Grade 1).

Regarding the Reading Fluency items, approximately 15–20 of the 50+ items were answered by more than 20 students in Grade 2 and approximately 11–17 of the 50+ items were answered by more than 20 students in Grade 3 across forms. Of these items, 16% in Grade 2 and 8% in Grade 3 had negative discrimination values, indicating potentially poorly fitting items.

Table 5-11: Percent of Items with Infit or Outfit z-Values Larger than 2.0, by Skill Area — Kindergarten

| Form | Skill Area | | | |
|---|---|---|---|---|
| | Listening Comp N=8 | Reading Comp N=8 | Phonological Awareness N=27 | Vocabulary N=12 |
| 1 | 0% | 0% | 0% | 0% |
| 2 | 13% | 0% | 4% | 8% |
| 3 | 13% | 0% | 4% | 8% |
| 4 | 0% | 0% | 0% | 8% |
| 5 | 25% | 0% | 0% | 0% |
| 6 | 0% | 0% | 19% | 8% |
| 7 | 0% | 0% | 4% | 0% |
| 8 | 13% | 0% | 4% | 0% |
| 9 | 0% | 0% | 0% | 17% |
| 10 | 0% | 0% | 11% | 0% |
| 11 | 13% | 0% | 0% | 0% |
| 12 | 0% | 0% | 0% | 0% |
| 13 | 0% | 0% | 0% | 0% |
| **Total** | 6% | 0% | 3% | 4% |

Table 5-12: Percent of Items with Infit or Outfit z-Values Larger than 2.0, by Skill Area — Grade 1

| Form | Skill Area | | |
|---|---|---|---|
| | Reading Comp N=10 | Phonological Awareness N=23 | Vocabulary N=15 |
| 1 | 0% | 9% | 7% |
| 2 | 10% | 9% | 0% |
| 3 | 0% | 0% | 0% |
| 4 | 0% | 4% | 0% |
| 5 | 0% | 4% | 0% |
| 6 | 0% | 9% | 7% |
| 7 | 0% | 4% | 13% |
| 8 | 0% | 0% | 13% |
| 9 | 0% | 17% | 0% |
| 10 | 0% | 0% | 13% |
| 11 | 0% | 4% | 0% |
| 12 | 0% | 13% | 0% |
| 13 | 0% | 0% | 7% |
| **Total** | 1% | 6% | 5% |

Table 5-13: Percent of Items with Infit or Outfit z-Values Larger than 2.0, by Skill Area — Grade 2

| Form | Skill Area | |
|---|---|---|
| | Reading Comp N=15 | Vocabulary N=28 |
| 1 | 0% | 11% |
| 2 | 7% | 7% |
| 3 | 7% | 4% |
| 4 | 0% | 7% |
| 5 | 0% | 4% |
| 6 | 0% | 4% |
| 7 | 0% | 4% |
| 8 | 0% | 11% |
| 9 | 0% | 4% |
| 10 | 0% | 7% |
| 11 | 0% | 7% |
| 12 | 0% | 4% |
| 13 | 0% | 7% |
| **Total** | 1% | 6% |

Table 5-14: Percent of Items with Infit or Outfit z-Values Larger than 2.0, by Skill Area — Grade 3

| Form | Skill Area | |
| --- | --- | --- |
| | Reading Comp N=20 | Vocabulary N=24 |
| 1 | 5% | 4% |
| 2 | 0% | 0% |
| 3 | 0% | 0% |
| 4 | 5% | 0% |
| 5 | 5% | 4% |
| 6 | 0% | 0% |
| 7 | 5% | 0% |
| 8 | 0% | 4% |
| 9 | 0% | 0% |
| 10 | 0% | 4% |
| 11 | 0% | 0% |
| 12 | 0% | 0% |
| 13 | 0% | 0% |
| Total | 2% | 1% |

Table 5-15: Proportion of Point-Biserial Correlations (Item Discrimination) Less than 0, by Skill Area — Kindergarten

| Form | Skill Area | | | |
|---|---|---|---|---|
| | Listening Comp N=8 | Reading Comp N=8 | Phonological Awareness N=27 | Vocabulary N=12 |
| 1 | 0% | 25% | 0% | 25% |
| 2 | 0% | 50% | 4% | 0% |
| 3 | 13% | 38% | 0% | 8% |
| 4 | 0% | 75% | 0% | 8% |
| 5 | 0% | 50% | 19% | 33% |
| 6 | 13% | 50% | 15% | 0% |
| 7 | 0% | 25% | 7% | 17% |
| 8 | 0% | 25% | 4% | 0% |
| 9 | 0% | 63% | 4% | 8% |
| 10 | 0% | 25% | 4% | 8% |
| 11 | 13% | 63% | 4% | 8% |
| 12 | 0% | 50% | 7% | 0% |
| 13 | 0% | 25% | 0% | 0% |
| Total | 3% | 43% | 5% | 9% |

Table 5-16: Proportion of Point-Biserial Correlations (Item Discrimination) Less than 0, by Skill Area — Grade 1

| | Skill Area | | |
|---|---|---|---|
| Form | Reading Comp N=10 | Phonological Awareness N=23 | Vocabulary N=15 |
| 1 | 30% | 4% | 13% |
| 2 | 40% | 13% | 40% |
| 3 | 0% | 4% | 20% |
| 4 | 30% | 0% | 0% |
| 5 | 0% | 0% | 7% |
| 6 | 20% | 13% | 33% |
| 7 | 20% | 0% | 33% |
| 8 | 10% | 4% | 13% |
| 9 | 30% | 13% | 0% |
| 10 | 10% | 0% | 0% |
| 11 | 30% | 0% | 7% |
| 12 | 20% | 4% | 7% |
| 13 | 30% | 4% | 13% |
| Total | 21% | 5% | 14% |

Table 5-17: Proportion of Point-Biserial Correlations (Item Discrimination) Less than 0, by Skill Area — Grade 2

| Form | Skill Area | |
| --- | --- | --- |
| | Reading Comp N=15 | Vocabulary N=28 |
| 1 | 20% | 11% |
| 2 | 13% | 7% |
| 3 | 13% | 4% |
| 4 | 13% | 11% |
| 5 | 20% | 4% |
| 6 | 27% | 0% |
| 7 | 20% | 4% |
| 8 | 20% | 7% |
| 9 | 27% | 7% |
| 10 | 33% | 4% |
| 11 | 20% | 7% |
| 12 | 20% | 7% |
| 13 | 27% | 7% |
| Total | 21% | 16% |

Table 5-18: Proportion of Point-Biserial Correlations (Item Discrimination) Less than 0, by Skill Area — Grade 3

| Form | Skill Area | |
| --- | --- | --- |
| | Reading Comp N=20 | Vocabulary N=24 |
| 1 | 10% | 17% |
| 2 | 30% | 25% |
| 3 | 25% | 17% |
| 4 | 20% | 21% |
| 5 | 10% | 17% |
| 6 | 10% | 4% |
| 7 | 15% | 13% |
| 8 | 10% | 8% |
| 9 | 5% | 13% |
| 10 | 15% | 21% |
| 11 | 5% | 8% |
| 12 | 15% | 13% |
| 13 | 40% | 8% |
| **Total** | 16% | 14% |

# Correlations with External Measures (Inferences Regarding Extrapolation)

In addition to completing ISIP Español assessments, students completed beginning of school year evaluation using two external measures: *Evaluación del Desarrollo de la Lectura* (EDL2, Pearson) and *Tejas LEE* (Brookes). Students in Kindergarten through Grade 2 will also complete mid-year and year-end evaluations using the same external measures. Students in Grade 3 will also complete the State of Texas Assessment of Academic Readiness (STAAR Reading, TEA) in the spring of 2015 (these data will also be analyzed at a later date). Correlations with these external measures provide evidence of association with similar measures (traditionally referred to as concurrent criterion-related validity evidence). Correlations with the STAAR provide predictive criterion-related validity evidence, as those scores are obtained several months later. At this time, scores from beginning of school year (BOY) EDL2 and Tejas LEE have been obtained and correlated with each skill area by form and grade.

Decisions and suggestions to be made based on available data include the following:

- These data provide promising results for future administration in a computer adaptive testing system. The larger data set being gathered across multiple districts will be used to estimate item parameters to support computer adaptive testing development.

- These data indicate that a number of items in each domain will be reviewed, potentially revised, or removed from further administration and the computer adaptive testing model.

- Data suggest that students in Kindergarten might be better prepared for administration of Reading Comprehension items during spring semester, rather than beginning Kindergarten.

These data indicate that the small to moderate correlations between domain areas support separate reporting of domain scores. These domains are relatively independent, providing unique information about separate skills.

# Chapter 6: Determining Norms

Norm-referenced tests are designed so that test administrators have a way of comparing the results of a given test taker to the hypothetical "average" test taker to determine whether they meet expectations. In the case of the Computerized Adaptive Testing (CAT)-based ISIP Español test, we are interested in comparing students to a national sample of students. We are also interested in knowing what the expected growth of a given student is over time, and in administering our test regularly to students to determine how they are performing relative to this expected growth. By determining and publishing these norms, we enable teachers, parents, and students to know how their scores compare with a representative sample of children in their particular grade for the particular period (month) in which the test is administered.

The norming samples were obtained as part of Istation's ongoing research in assessing reading ability. The samples were drawn from all enrolled ISIP Español users during the 2011-2012 school year. In the case of ISIP Español, we felt that demographic considerations were moot, in that most of the students taking the test would be of a similar demographic, and it is difficult to say what constitutes a representative sample of Spanish speaking students in the United States. As such, all users of the program were considered in determining the norms for the test.

Table 21. *Demographics* for ISIP Español Reading *Norming Sample*.

| | Grade | | | | | |
| | PK-3 | PK | K | 1 | 2 | 3 |
|---|---|---|---|---|---|---|
| Gender | | | | | | |
|   Male | 51.2 | 50.4 | 51.5 | 51.6 | 51.0 | 50.8 |
|   Female | 48.8 | 49.6 | 48.5 | 48.4 | 49.0 | 49.2 |
| Race | | | | | | |
|   African American | 2.6 | 2.6 | 4.3 | 2.8 | 1.7 | 1.1 |
|   American Indian/Native | 11.6 | 8.6 | 12.8 | 12.4 | 9.3 | 13.0 |
|   Asian | 0.3 | 0.2 | 0.5 | 0.3 | 0.2 | 0.1 |
|   Other | 6.0 | 6.1 | 6.3 | 5.8 | 6.0 | 5.5 |
|   Pacific Islander | 0.6 | 0.5 | 0.7 | 0.7 | 0.6 | 0.6 |
|   No Answer | 13.7 | 14.3 | 13.7 | 14.1 | 13.2 | 13.7 |
|   White | 65.2 | 67.8 | 61.7 | 63.9 | 68.9 | 66.1 |
| Economically Disadvantaged | | | | | | |
|   Yes | 57.6 | 70.8 | 59.5 | 54.7 | 53.9 | 57.8 |
|   No/NA | 42.4 | 29.2 | 40.5 | 45.3 | 46.1 | 42.2 |
| Ethnicity | | | | | | |
|   Hispanic | 76.0 | 79.5 | 74.7 | 75.4 | 75.8 | 77.2 |
|   Non-Hispanic | 5.2 | 4.3 | 8.5 | 5.5 | 3.8 | 2.5 |
|   No answer | 19.1 | 16.8 | 17.1 | 19.4 | 20.6 | 20.6 |

*Note*: Each category is percent of total responding.

# Instructional Tier Goals

Consistent with other reading assessments, Istation has defined a three-tier normative grouping, based on scores associated with the 20th and 40th percentiles. Students with a score above the 40th percentile for their grade are placed into Tier 1. Students with a score below the 20th percentile are placed into Tier 3. These tiers are used to guide educators in determining the level of instruction for each student. That is, students classified as:

- Tier 1 are performing at grade level.

- Tier 2 are performing moderately below grade level and in need of intervention.

- Tier 3 are performing seriously below grade level and in need of intensive intervention.


# Computing Norms

Istation's norms are time-referenced to account for expected growth of students over the course of a semester. The ISIP Español test consists of several subtests and an overall score. Each of these is normed separately so that interested parties can determine performance in various areas independently. To compute these norms, the ISIP Español test was given to the students in the sample described above, once per month throughout a school year. Because of the test design, including computer-adaptive subtests, retakes of the test result in different test items for a given student, so it is expected that improved scores on the test reflect actual growth over time. Norms were computed for each time period, so that over time a student's score on ISIP Español is expected to go up. Norming tables for each of the ISIP Español subtests, as well as Overall, can be found at Istation's website, and these represent the results of norming all subtests and the overall score across all the periods of testtaking. For each time period, these scores were averaged and a standard deviation was computed. Then, to determine expected Tier 2 and Tier 3 scores, the 20th and 40th percentiles on a true normal bell curve were computed, and these numbers are given as norms for those Tier groups.

# References

Abadzi, H., Crouch, L., Echegaray, M., Pasco, C., & Sampe, J. (June 1, 2005). Monitoring basic skills acquisition through rapid learning assessments: A case study from Peru. Prospects: *Quarterly Review of Comparative Education*, 35, 2, 137-156.

AERA, APA, NCME. (1999). *Standards for educational psychological testing*. Washington DC: AERA.

Agrás G, Molina H, & Bareche Monclús, R., (2007).  Serie Didáctica de la Lengua y de la Literatura. Editorial Graó.

Armbruster, B. B., Lehr, F., & Osborn, J., National Institute for Literacy (U.S.), & RMC Research Corporation. (2006). A child becomes a reader: Proven ideas from research for parents: kindergarten through grade 3. Washington, D.C.: National Institute for Literacy.

August, D., & Shanahan, T. (Eds.). (2006). Developing literacy in second-language learners: Report of the National Literacy Panel in Language-Minority Children and Youth. Studies in second language acquisition, 30, 1, 116-117.

Ballantyne, K. G., Sanderman,  A. R., McLaughlin, N., & National Clearinghouse for English Language Acquisition & Language Instruction Educational Programs. (2008). *Dual language learners in the early years: Getting ready to succeed in school*. Washington, D.C: National Clearinghouse for English Language Acquisition.

Bazán, A. R., Acuña L., & Vega, F. Y., (2002).  Efectos de un método para la enseñanza de la lectura y la escritura en el primer grado de primaria. Instituto Tecnológico de Sonora, Departamento de Psicología y Educación, México.

Beck, I. L., McKeown, M. G., & Kucan, L. (2002). *Bringing words to life: Robust vocabulary instruction*. New York: The Guilford Press.

Bravo, M. V., Silva, M., Razmilic, T., & Swartz, S. (January 1, 2005). Educando juntos: a long-term intervention to promote literacy learning in low-performing primary schools in Chile. *Early Years, 25,* 2, 97-111.

Broer, M., Castillo, M., Malespín, M., & Gómez, P. (2009). Report on the Results of the EGRA 2008 Pilot Assessment, Oficina de Desarrollo Económico, Agricultura y Comercio (EGAT/ED) MIN ED.

Cao, J. & Stokes, S. L. (2006). Bayesian IRT guessing models for partial guessing behaviors, manuscript submitted for publication.

Cárdenas-Hagan, E., Carlson, C. D., & Pollard-Durodola, S. D. (2007). The cross-linguistic transfer of early literacy skills: The role of initial L1 and L2 skills and language of instruction. L*anguage, Speech, and Hearing in the Schools*, 249-259.

Carrillo, M. (1994): Development of phonological awareness and reading acquisition. Reading and Writing: *An Interdisciplinary Journal* 6, 279-298.

Conte, K.L, & Hintz, J. M. (2000). The effect of performance feedback and goal setting on oral reading fluency with CBM. *Diagnostique, 25, 85-98.*

Cisero, C. A. & Royer, J.M. (1995). The development and cross-language transfer of phonological awareness. *Contemporary Educational Psychology*, 20, 275-303.

Crawford, J., & Krashen, S. D. (2007). *English learners in American classrooms: 101 questions, 101 answers*. New York: Scholastic.

Cronbach, L.J. (1988). Five perspectives of the validity argument. In H. Wainer & H.I. Braun (Eds.), *Test validity* (pp. 3-18). Hillsdale, NJ: Lawrence Erlbaum.

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, 16, 297-334.

Degraff A. & Torgesen, J. ( 2005). Monitoring growth in early reading skills: Validation of a computer adaptive test. Florida State University College of Arts and Sciences. Dissertation, spring 2005.

Deno, S. (January 1, 2003). Developments in curriculum-based measurement. *The Journal of Special Education*, 37, 3, 184-192.

Deno, S. L., (1985). Curriculum-based measurement: The emerging alternative. *Exceptional Children.*

DeThorne, L., Petrill, S., Deater-Deckard, K., Thompson, L., & Schatschneider, C. (January 1, 2006). Reading skills in early readers. *Journal of Learning Disabilities*, 39, 1, 48-55.

Díaz-Rico, L. T., & Díaz-Rico, L. T. (2008). *Strategies for teaching English learners*. Boston: Pearson/Allyn and Bacon.

Dickinson, D. & Tabors, P. (2001). *Beginning literary with language.* Baltimore: Paul E. Brookes.

Ehri, L. (2000). Learning to read and learning to spell: Two sides of a coin. *Topics in Language Disorders, 20*(3), 19-49.

Escamilla, K. (November 1, 2006). Semilingualism Applied to the literacy behaviors of Spanish-speaking emerging bilinguals: Bi-illiteracy or emerging bi-literacy. Teachers College Record, 108, 11, 2329-2353.

Escamilla, Kathy, (2000). Bilingual means two: Assessment issues, early literacy and Spanish-speaking children in a research symposium on high standards in reading for students from diverse language groups: Research, practice & policy (pp. 100-128). Washington, DC: U.S. Office of Bilingual Education and Minority Affairs.

Espin, C., Deno, S., Maruyama, G. & Cohen, C. (1989). *The basic academic skills samples (BASS): An instrument for the screening and identification of children at-risk for failure in regular education classrooms.* Paper presented at the annual American Educational Research Association Conference, San Francisco, CA.

Espinosa, Linda M., & Lopez, Michael. (2007). Assessment considerations for young English language learners across different levels of accountability Philadelphia: National Early Childhood Accountability Task Force.

Farver, J. M., Nakamoto, J., & Lonigan, C. J. (2007). Assessing preschoolers' emergent literacy skills in English and Spanish with the Get Ready to Read! screening tool. *Annals of Dyslexia*, 57, 161–178. doi: 10.1007/s11881-007-0007-9.

Ferreiro, E. (1997). Alfabetización: Teoría y práctica. México: Siglo Veintiuno Editores, Chile.

Ferreiro, E. (January 1, 2009). Ensayos e investigaciones - La desestabilización de las escrituras silábicas: alternancias y desorden con pertinencia. *Lectura y vida : Revista Latinoamericana de lectura*, 30, 2, 6.

Fletcher, J. M., Foorman, B. R., Francis, D. J., & Schatschneider, C. (1997). Prevention of reading failure. *Insight*, 22-23.

Fletcher, J. M., Foorman, B. R., Boudousquie, A., Barnes, M., Schatschneider, C., & Francis, D. J. (2002). Assessment of reading and learning disabilities: A research-based, treatment-oriented approach. *Journal of School Psychology*, *40*, 27-63.

Foorman, B. R., Anthony, J., Seals, L., & Mouzaki A.(2002). Language development and emergent literacy in preschool*, Seminars in Pediatric Neurology*, 9, 172-183.

Foorman, B. R., Santi, K., & Berger, L. (in press). Scaling assessment-driven instruction using the Internet and handheld computers. In B. Schneider & S. McDonald (Eds.), *Scale-up in education, vol. 1: Practice*. Lanham, MD: Rowan & Littlefield Publishers, Inc.

Foorman, B. R., & Torgesen, J. (2001). Critical elements of classroom and small-group instruction promote reading success in all children. *Learning Disabilities Research & Practice, 16,* 203-212.

Fuchs, L. S. & Deno, S. L. ( 1991). Paradigmatic distinction between instructionally relevant measurement models. *Exceptional Child, 57.*

Fuchs, L. S., Deno, S. L.& Marston D. (1983). Improving the reliability of curriculum- based measures of academic skills for psycho education decision-making, *Diagnotique, 8.*

Fuchs, L. S., Deno, S. L., & Mirkin, P. K. (1984). The effects of frequent curriculum-based measurement and evaluation on pedagogy, student achievement, and student awareness of learning. *American Educational Research Journal*, 21, 449-46.

Fuchs, D., and Fuchs, L. (1990). Making educational research more important. *Exceptional Children*, 57, 102-108

Fuchs, L. S., Fuchs, D., Hamlett, C. L., & Ferguson, C. (1992). Effects of expert system consultation within curriculum-based measurement using a reading maze task. Exceptional Children, 58, 436-450.

Fuchs, L. S., Hamlett, C., & Fuchs, D. (1995). *Monitoring basic skills progress: Basic reading – version 2* [Computer program]. Austin, Tx:PRO-ED.

Fuchs, L. S., Fuchs, D., Hamlett, C. L., & Ferguson, C. (1992). Effects of expert system consultation within curriculum-based measurement using a reading maze task. *Exceptional Children*, 58, 436-4.

Fuchs, L. S., Fuchs, D., Hamlett, C. L., & Stecker, P. M. (1991). Effects of curriculum-based measurement and consultation on teacher planning and student achievement in mathematics operations. *American Educational Research Journal*, 28, 617-641.

Garcia, E. E., & Miller, L. S. (August 01, 2008). Findings and recommendations of the national task force on early childhood education for Hispanics. *Child Development Perspectives, 2,* 2, 53-58.

Gersten, R. M., & National Center for Education Evaluation and Regional Assistance (U.S.). (2007). Effective literacy and English language instruction for English learners in the elementary grades. Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education.

Gilks, W. R. & Wild, P. (1992). Adaptive rejection sampling for Gibbs sampling. *Applied Statistics* 41, pp 337-348.

Good, R. H., Shinn, M. R., & Bricker, D. (1991). *Early Intervention to prevent special education: Direct assessment of student progress on pre-reading skills*. Paper submitted to the Field-Initiated Research Projects (84.023C) Competition United States Department of Education Office of Special Education and Rehabilitative Services. Oregon.

Good, R. H. & Kaminski, R. A. (2002). DIBELS. Oral reading fluency passages for first through third grade (Technical Report No. 10). Eugene, OR: University of Oregon.

Good, R. H. & Kaminski, R. A. (1996). Assessment for instructional decisions: Toward a proactive/prevention model of decision-making for early literacy skills. *Scholastic Psychology, 11, 325-336.*

Good, R. H., & Kaminski, R. A. (Eds.). (2002b). Dynamic indicators of basic early literacy skills (6th ed.). Eugene, OR: Institute for the Development of Education Achievement.

Good, R. H., Simmons, D., & Kame'enui, E. J. (2001). The importance and decision-making utility of a continuum of fluency-based indicators of foundational reading skills for third-grade high-stakes outcomes. *Scientific Studies of Reading*, 5, 257-288.

Gorman, B., & Gillam, R. (2003). Phonological awareness in Spanish: A tutorial for speech-language pathologists. Communication *Disorders Quarterly*, 25(1), 13-22.

Hirsch, E. D. J. (June 06, 2003). Reading Comprehension Requires Knowledge--of Words and the World. *American Educator, 27,* 1, 10.

Howard, E. R., Lindholm-Leary, K. J., & Center for Applied Linguistics. (2007). *Guiding principles for dual language education*. Washington, D.C.: Center for Applied Linguistics.

Jenkins, J. R., & Jewell, M. (1993). Examining the validity two measures for formative teaching: Reading aloud and maze. *Exceptional Children*, 59, 421-432.

Jenkins, J. R., Pious, C.G., & Jewell, M. (1990). Special education and the regular education initiative: Basic assumptions. *Exceptional Children, 56*, 479-91.

Jiménez González, J. E., & Ortiz, R. (2000). Metalinguisitc awareness and reading acquisition in the Spanish language. *The Spanish Journal of Psychology*, 3, 37-46.

Jiménez, G., & O'shanahan, I. (2008). *Evolución de la escritura de palabras de ortografía arbitraria en lengua española*. 20.

Junqué, P.C., Bruna, R.O., & Mataró, S.M. (2004). *Neuropsicología del lenguaje: Funcionamiento normal y patológico rehabilitación*. Barcelona: Masson.

Kaminski, R. A., & Good, R. H. (1996). Toward a technology for assessing basic early literacy skills. *School Psychology Review, 25,* 215-227.

Kane, M.T. (1992). An argument based approach to validity. *Psychological Bulletin, 112*, 527-535.

Kane, M.T. (2006a). Content-related validity evidence. In S.M. Downing & T.M. Haladyna (Eds.), *Handbook of test development* (pp. 131-154). Mahwah, NJ: Lawrence Erlbaum.

Kane, M.T. (2006b). Validation. In R.L. Brennan (Ed.), *Educational measurement* (4th ed.) (pp. 17-64). Westport, CT: American Council on Education/Praeger.

Kudo, I., Bazan, J., & Banco Mundial (2009). Measuring beginner reading skills: An empirical evaluation of alternative instruments and their potential use for policymaking and accountability in Peru. Washington, D.C: World Bank. Main Report. Tempe, AZ: Author. Retrieved from September 23, 2008.

Lindsey, K. A., Manis, F. R., & Bailey, C. E. (2003). Prediction of first-grade reading in Spanish-speaking English-language learners. *Journal of Educational Psychology, 95,* 482-494. doi: 10.1037/0022-0663.95.3.482.

Lonigan, C. J., Burgess, S. R., & Anthony, J. L. (2000). Development of emergent literacy and early reading skills in preschool children: Evidence from a latent variable longitudinal study. *Developmental Psychology, 36, 596-613.*

Manis, F. R., Lindsey, K. A., & Bailey, C. E. (2004). Development of reading in grades K-2 in Spanish-speaking English-language learners. *Learning Disabilities Research and Practice, 19,* 214-224.

Marco común europeo de referencia para las lenguas: enseñanza, aprendizaje y evaluación: Propuestas para la enseñanza de ELE. (2005). Madrid: Sociedad General Española de Librería.

Marston, D. B. (1989). *A curriculum-based measurement approach to assessing academic performance. What is it and why do it?* New York. Guillford Press.

Mathes, P. G., Fuchs, D., & Roberts, P. H., (1998). The impact of Curriculum-Based Measurement on Transenvironmental Programming. *The Journal of Learning Disabilities*, 31(6), 615-624.

Messick, S. (1989). Validity. In R.L. Linn (Ed.), *Educational measurement* (3rd ed.) (pp. 13-104). New York: American Council on Education and Macmillan.

National Task Force on Early Childhood Education for Hispanics/La Comisión Nacional para la Educación de la Niñez Hispana. (2007). Para nuestros niños. Expanding and Improving Early Education for Hispanics:

National Reading Panel. (2000). *Teaching children to read: An evidence based assessment of the scientific research literature on reading and its implications for reading instruction: Reports of the subgroups*. Bethesda, MD: National Institute of Child Health and Human Development.

O'Connor, R. E., & Jenkins, J. R. (1999). The prediction of reading disabilities in kindergarten and first grade. *Scientific Studies of Reading, 3,* 159-197.

Pearson, P. D. (2002). Handbook of reading research. Mahwah, N. J: L. Erlbaum Associates.

Proctor, C. P., August, D., Carlo, M.S., & Snow, C. (2006). The intriguing role of Spanish language vocabulary knowledge in predicting English reading comprehension. *Journal of Educational Psychology, 98,* 159-169. doi:10.1037/0022-0663.98.1.159

Rayner, K., Foorman, B. R., Perfetti , C. A., Pesetsky, D. & Seidenberg, M.S. (2001) How psychological science informs the teaching of reading. *Psychological Science in the Public Interest,* 2 (2), 31-74.

Roseberry-McKibbin, C., Brice, A., & O'Hanlon, L. (January 01, 2005). Serving English language learners in public school settings: A national survey. *Language, Speech, and Hearing Services in Schools*, 36, 1, 48-61.

Scarborough, H. S. (1998). Early identification of children at risk for reading disabilities: Phonological awareness and some other promising predictors. In B. K. Shapiro, P. J. Accardo, & A. J. Capute (Eds.), *Specific reading disability: A view of the spectrum* (pp. 75-107). Timonium, MD: York Press.

Scarborough, H. S., & Dobrich, W. (January 01, 1990). Development of children with early language delay. *Journal of Speech and Hearing Research, 33,* 1, 70-83.

Serrano, F., Defior, S. & Jiménez, G., (2009). *Evolución de la relación entre conciencia fonológica y lenguaje escrito en niños españoles de primer curso de Educación Primaria*. Universidad de Granada y Ministerio Español de Ciencia y Tecnología.

Seymour, P. H. K., Aro, M., & Erskine, J. M. (2003). Foundation literacy acquisition in European orthographies. *British Journal of Psychology*, 94, 143-174.

Share, D. L., & Stanovich, K.E. (1995). Cognitive processes in early reading development: Accommodating individual differences into a model of acquisition. *Issues in Education*, *1*, 1-57.

Shaywitz, S. E. (1996, November). Dyslexia*. Scientific American*, 98-104.

Shaywitz, S.E. (1996) Dyslexia, *Scientific American*.

Shin, J., Deno, S. L., & Espin, C. (2000). Technical adequacy of probes for Curriculum-Based Measurement of reading growth. *The Journal of Special Education,* 34(3), 140-153.

Shinn, M. R., Good, R. H., Knutson, N., Tilly, W. D., & Collins, V. L. (1992). Curriculum-based measurement reading fluency: A confirmatory analysis of its relation to reading. *School Psychology Review, 21*, 459-479.

Signorini, A. (1997). Word reading in Spanish: A comparison between skilled and less skilled beginning readers. *Applied Psycholinguistics*, 18, 319-344.

Snow, C. E., Griffin, P., and Burns, M. S. (Eds.) (2005). *Knowledge to support the teaching of reading: Preparing teachers for a changing world.* Indianapolis, NJ: Jossey-Bass.

Snow, C. E., Burns, S., & Griffin, P. (Eds.). (1998). *Preventing reading difficulties in young children.* Washington, DC: National Academy Press.

Snow, C. E., Carlo, M. S., August, D., Mclaughlin, B., Dressler, C., Lippman, D. N., Lively, T. J., & White, C. E. (January 01, 2004). Closing the gap: Addressing the vocabulary needs of English-language learners in bilingual and mainstream classrooms (Cerrando la brecha: Acerca de las nesecidades de vocabulario de aprendices de inglés en aulas bilingües y comunes). *Reading Research Quarterly, 39,* 2, 188-215.

Sprenger-Charolles, L., Carré, R., & Demonet, J. F. Serniclaes, W. (January 01, 2001). Perceptual discrimination of speech sounds in developmental dyslexia. *Journal of Speech, Language, and Hearing Research*: Jslhr, 44, 2, 384-99.

Stanovich, K. E. (1991). Word recognition: Changing perspectives. In R. Barr, M. L. Kamil, P. Mosenthal, & P. D. Pearson (Eds.), *Handbook of reading research* (Vol. 2, pp. 418-452). New York: Longman.

Stecker, P. M., & Fuchs, L. S. (2000). Effecting superior achievement using curriculum-based measurement: The importance of individual progress monitoring. *Learning Disabilities Research and Practice, 15*, 128-134.

Stiggins, R. J., & Assessment Training Institute. (2006). Grading & reporting in standards-based schools. Portland, OR: Assessment Training Institute.

Swaminathan, H., and Rogers, H. J. (1990). Detecting differential item functioning using logistic regression procedures. *Journal of Educational Measurement*, 27, 361-370.

Torgesen, J.K. (2002). The prevention of reading difficulties. *Journal of School Psychology, 40 ,7-26.*

Torgesen, J.K., Rashotte, C.A., Alexander, A.W. (2002). Principles of fluency instruction in reading: Relationships with established empirical outcomes. In M. Wolf (Ed.) *Time, Fluency, and Dyslexia*. Parkton, MD: York Press.

Vaughn, S., & Linan-Thompson, S. (2004). *Research-based methods of reading instruction, grades K-3*. Alexandria, Va: Association for Supervision and Curriculum Development.

Vellutino, F. R., Scanlon, D. M., & Lyon, G. R. (2000). Differentiating between difficult-to-remediate and readily remediated poor readers: More evidence against the IQ-achievement discrepancy definition of reading disability. *Journal of Learning Disabilities, 33,* 223-238.

Vellutino, F. R. (1991). Introduction to three studies on reading acquisition: Convergent findings on theoretical foundations of code-oriented versus whole-language approaches to reading instruction. *Journal of Educational Psychology, 83*, 437-443.

Wagner, R. K., Torgesen, J. K., Laughon, P., Simmons, K., & Rashotte, C. A. (1993). Development of young readers' phonological processing abilities. *Journal of Educational Psychology*, 85, 83-103.

Wimmer, H., Mayringer, H., & Landerl, K. (December 01, 2000). The double-deficit hypothesis and difficulties in learning to read a regular orthography. *Journal of Educational Psychology, 92,* 4, 668-80.

Wood, F., Hill, D., & Meyer, M. (2001). *Predictive Assessment of Reading*. Winston-Salem, NC: Wake Forest University School of Medicine.

Ziegler, J. C., Perry, C., Ma-Wyatt, A., Ladner, D., & Schulte-Körne, G. (January 01, 2003). Developmental dyslexia in different languages: language-specific or universal. *Journal of Experimental Child Psychology, 86,* 3, 169-93.