

¿Qué es un diagnóstico de lectura?

Es una evaluación de aspectos específicos o componentes de la lectura, tales como la conciencia fonológica, las habilidades de decodificación, la fluidez, y la comprensión. Los diagnósticos de lectura pueden ser evaluaciones formales estandarizadas en base a normas, o evaluaciones menos formales referidas a criterios.

Las **evaluaciones estandarizadas** basadas en normas comparan el puntaje que un estudiante obtiene con el de una población usada para la estandarización de la prueba para determinar qué tan bien el estudiante se desempeñó en comparación a otros estudiantes de su mismo curso, edad, o grupo socioeconómico. Este tipo de evaluación no debe administrarse más que una vez en el año.

Las **evaluaciones con base a criterios** pueden ser formales o informales y generalmente se usan como evaluaciones diagnósticas. Estas evaluaciones comparan el grado de habilidad de un estudiante con relación a una meta (o criterio) de logro. Este tipo de evaluación puede administrarse al inicio y fin del año escolar o antes y después de una intervención o programa, para medir el crecimiento que los estudiantes logran en una determinada habilidad.

¿Qué es Dialect y qué mide?

Dialect es una plataforma digital desarrollada por **Colegium y la Universidad de los Andes** para el diagnóstico temprano de habilidades lectoras. La plataforma se basa en la noción de que un adecuado nivel de comprensión lectora depende de que ciertos subprocesos lectores se encuentren adecuadamente desarrollados. Estos componentes son: conciencia fonológica, conocimiento del alfabeto, decodificación, comprensión auditiva, vocabulario, concepto de lo impreso y fluidez. Así, un estudiante que no logra un adecuado nivel de desempeño en comprensión lectora puede tener alguna dificultad específica de comprensión, o bien puede tener alguna dificultad específica en alguno de estos componentes, que incide en su bajo desempeño en comprensión.

En lo relativo a la comprensión lectora, Dialect permite, por una parte, establecer el nivel de comprensión que un lector obtiene al leer diversos textos y responder preguntas relacionadas, y por otra, determinar las fortalezas y debilidades en los subprocesos lectores, cuando la comprensión lectora es inferior a lo esperable para su nivel de desarrollo. Como evaluación diagnóstica, Dialect establece criterios para comparar el grado de habilidad de un estudiante con relación a una meta. Estos criterios se han establecido a partir de reiteradas mediciones a estudiantes de Chile y Latinoamérica y se refieren a:

- Conciencia fonológica
- Comprensión auditiva
- Identificación de letras del alfabeto

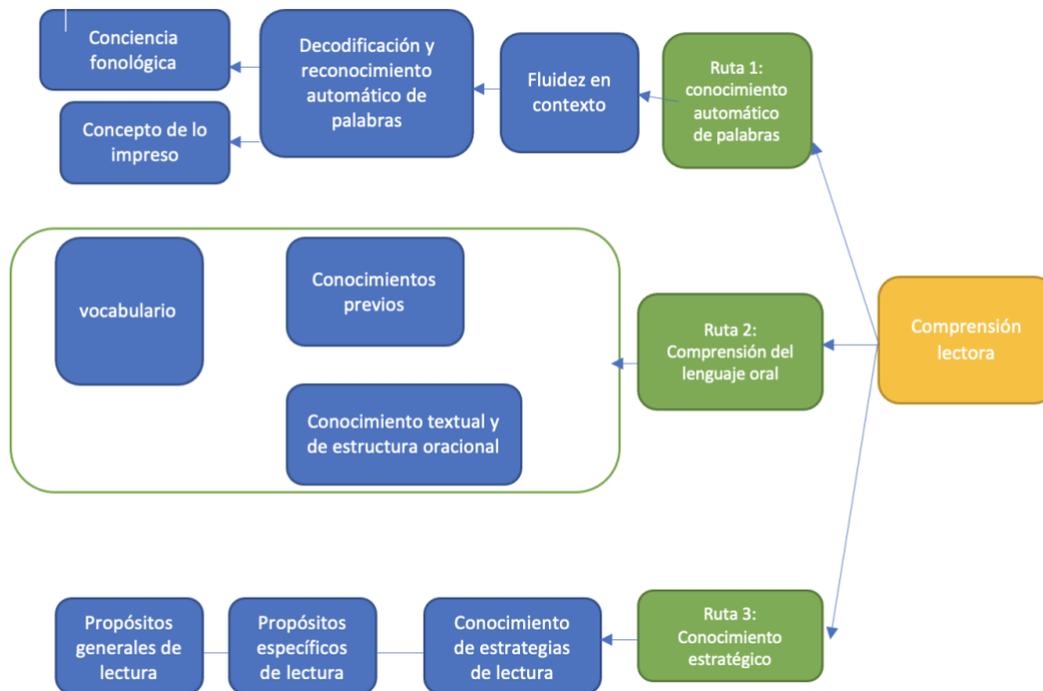
- Concepto de lo impreso
- Vocabulario
- Reconocimiento de palabras

Para determinar el nivel de habilidad en **comprensión lectora**, el puntaje de un estudiante en esta sección del diagnóstico se representa en Lexiles. Un **Lexile** es una medida científica de la dificultad de un texto y de la comprensión lectora de un individuo. Se representa con un número seguido de la letra L, que va desde los 0L (representado por las letras BR) para lectores principiantes y textos sencillos a los 1500L para lectores y textos avanzados. Existen muchas pruebas estandarizadas de lectura que utilizan la medida Lexile, la cual es una métrica estandarizada que tiene procesos estandarizados

Modelo Subyacente y Componentes de Dialect.

Dialect se basa en un modelo cognitivo de la lectura (e.g. McKenna & Stahl, 2015; Spear-Swerling, & Sternberg, 1996), en el que la comprensión lectora depende del nivel de desempeño en tres componentes: lenguaje oral, capacidad para decodificar el texto impreso y conocimiento estratégico (figura 1). De este modo Dialect considera la comprensión lectora silenciosa como primer *screening* o tamizaje (en el caso de estudiantes que ya leen), y un flujo de subtests en habilidades que inciden en el grado de eficiencia en la comprensión: habilidades de identificación de letras y palabras, vocabulario, decodificación y habilidades fonológicas.

Figura 1
Modelo de lectura de McKenna & Stahl, 2015.



Fuente: Traducción de McKenna & Stahl, 2015.

La prueba de comprensión lectora está compuesta por 45 ítems de complejidad progresiva creciente. Esta complejidad se determinó usando el marco Lexile[®]. Los Lexiles son una medida que permite determinar, por una parte, el grado de dificultad que un texto supone para el lector (en función de su complejidad semántica y sintáctica) y por otra, el grado de habilidad lectora de cada persona (Stenner, Stanford-Moore, & Williamson, 2012; Swartz, Burdick, Hanlon, Stenner, Kyngdon, Burdick & Smith, 2014). Se trata de una escala evolutiva mediante la cual es posible además asignar textos de lectura a los estudiantes en función de su capacidad de comprender más que en función del grado que cursan o la edad cronológica. Los 45 ítems son del tipo cloze. En cada ítem al lector se le presenta un texto breve y una oración que debe completar con una de cuatro palabras sugeridas como alternativas. Para elegir la alternativa correcta, es necesario realizar inferencias a partir de la información explícita del texto. Cuando el estudiante comete 5 errores consecutivos, la evaluación se interrumpe.

Los textos están ordenados según nivel de dificultad, y los primeros textos presentan al lector dibujos y palabras. Los ítems van incrementando su complejidad a medida que el lector avanza. Así, por ejemplo, después de los ítems con dibujo vienen textos simples de una o dos oraciones, aumentando progresivamente su extensión y la complejidad semántica y sintáctica. Se procuró abordar distintos temas y tipos textuales para que existiera mayor variedad en las lecturas. Los ítems se elaboraron en función de las temáticas más recurrentes en los textos de lectura

entregados por el Ministerio de Educación a los establecimientos de educación básica. Se procuró además que éstos no presentaran sesgo a nivel de género, etnia, o grupo etario.

A partir de más de 80.000 mediciones realizadas en Chile, Argentina y España, hemos podido establecer una tabla de niveles Lexiles esperables para cada curso o nivel del sistema escolar chileno. Este se ha relacionado con las bandas que Scholastic (2014) establece para sus pruebas, cuyos resultados también se reportan en Lexiles. En la tabla n. 1 detallamos estos niveles.

Tabla n.1

Niveles Lexile por curso o nivel.

Nivel (fin de año)	Deficiente	Básico	Satisfactorio	Avanzado	Lexile esperado por nivel
Primero	BR	20L-190L	215L-405L	>435L	210L-410L
Segundo	BR-190L	215L-310L	330L-490L	>520L	325L-510L
Tercero	BR-190L	330L-425L	460L-590L	>625L	445L-605L
Cuarto	450L-545L	570L-630L	660L-775L	>825L	550L-700L
Quinto	<619L	620L a 829L	839L a 1010L	>1010L	620L a 829L
Sexto	<729L	730L a 924L	925L a 1070L	>1070L	730L a 924L
Séptimo	<769L	770L a 969L	970L a 1120L	>1120L	770L a 969L
Octavo	<789L	790L a 1009L	1010L a 1185L	>1185L	790L a 1009L
I Medio	<849L	850L a 1049L	1050L a 1260L	>1260L	850L a 1049L
II Medio	<889L	890L a 1079L	1080L a 1335L	>1335L	890L a 1079L
III Medio	<984L	985L a 1184L	1185L a 1385L	>1385L	985L a 1184L
IV Medio	<984L	985L a 1184L	1185L a 1385L	>1385L	985L a 1184L

Si el evaluado no decodifica aún (e.g., Kinder o Prekinder), o bien obtiene un resultado inferior al rango esperable, se le aplica la batería de pruebas que miden el nivel de desempeño en los componentes asociados a la comprensión: conciencia fonológica, conocimiento del alfabeto, decodificación, comprensión auditiva, concepto de lo impreso, y vocabulario. En el caso de conciencia fonológica, se elaboraron ítems similares a los de la Prueba de Alfabetización Inicial (PAI) desarrollada por Rolla & Villalón (2000). Para la habilidad de vocabulario se eligió el Test de Vocabulario en Imágenes (TEVI) de Echeverría, Herrera, & Segure (1995) y se procedió de igual manera. Posteriormente se elaboró un segundo test de vocabulario (EVOC) que sigue el mismo formato del TEVI, pero que cuenta con imágenes más actualizadas (fotografías en lugar de dibujos) y que se elaboró a partir de un corpus lingüístico creado usando textos escolares como su repertorio. Tanto la prueba de Alfabetización Inicial como el TEVI y EVOC presentan bastante robustez desde una perspectiva de validación y son pruebas bastante más conocidas a nivel de hispanohablantes. La habilidad de reconocimiento de letras se evaluó siguiendo el modelo de Marie Clay (1993) en el que se presentan letras mayúsculas y minúsculas a los

alumnos y se les pide identificar una de ellas a partir de tres opciones. Para la evaluación del reconocimiento de palabras se adaptó el modelo de “*flash and analysis*” desarrollado por Cunningham & Erickson (1998) en el que el estudiante debe leer una palabra en forma automática (1/4 de segundo o menos) o bien decodificarla en un espacio de tiempo más largo (2-3 segundos). Inicialmente esta evaluación se realizaba en forma manual; es decir, un evaluador mostraba la palabra al estudiante, este la leía, y el evaluador ingresaba la respuesta como “correcta” o “incorrecta” en la plataforma. Sin embargo, esta forma de evaluar era lenta y producía algunos errores en el ingreso de las respuestas, por lo que se modificó el procedimiento. Actualmente se presentan tres palabras de uso frecuente con características morfológicas similares (ej., primera sílaba igual, como en “mamá”, “manta”, “mayor”), y el estudiante debe elegir la correcta según el estímulo que se le entrega en el audio.

Para evaluar comprensión auditiva, los estudiantes escuchan un texto breve y de nivel de dificultad similar a su nivel de habilidad y responden preguntas que se les formulan por audio, eligiendo de entre 3 imágenes la que mejor representa la alternativa correcta.

Para medir concepto de lo impreso se presentan imágenes de distintas partes de un libro y el estudiante debe identificar conceptos tales como autor, título, inicio de la lectura, direccionalidad, palabra, algunos signos de puntuación y línea de texto.

En las Figuras 2 a 5 se detallan los componentes de Dialect y los flujos o árboles de decisión correspondientes a cada curso.

Figura n. 2.
Flujograma y Componentes Dialect para Estudiantes de Kindergarten.



Figura n. 3.
Flujograma y Componentes Dialect para Estudiantes de Primero básico.

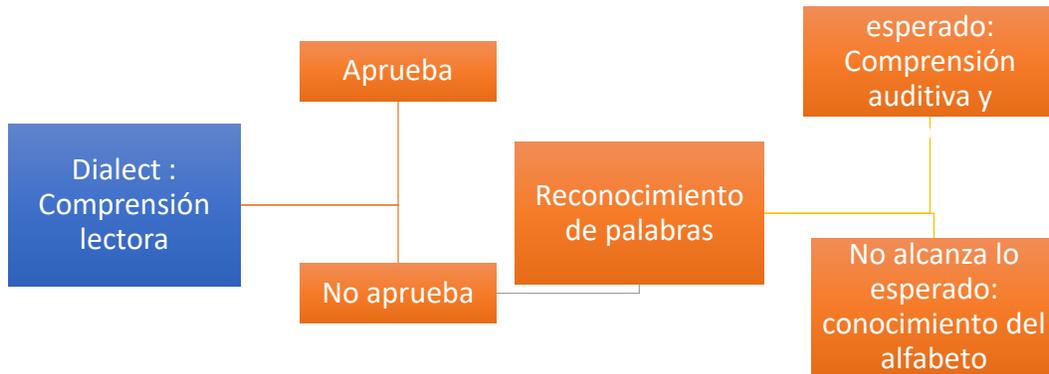


Figura 4.
Flujograma Dialect para segundo y tercero básico

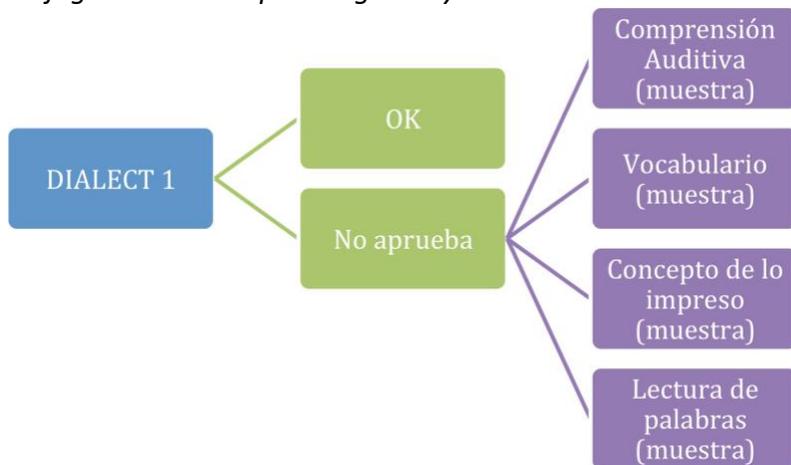


Figura 5.
*Flujograma Dialect para alumnos de cuarto básico**



Una vez obtenidos los resultados de cada una de las evaluaciones, es posible obtener informes individuales y grupales en forma inmediata. En éstos se interpretan los resultados en función de normas, estándares o criterios, según sea el caso. Se entregan sugerencias a los docentes respecto de los ámbitos lectores más descendidos y de las fortalezas grupales. Estas sugerencias están contenidas en el manual Enseñar Lenguaje, desarrollado por el equipo Dialect de la Universidad de los Andes y ofrece más de 60 actividades para cada componente de la lectura.

En los informes individuales se entrega además la medida Lexile para cada niño, con una sugerencia de títulos asociados a ese Lexile, y que el alumno podría leer y comprender en forma independiente con un 75% de logro. Los informes se pueden configurar de manera que informen a profesores, directivos o padres y apoderados.

Para diagnosticar habilidades de comprensión lectora en estudiantes de 10 años de edad en adelante contamos con dos formas paralelas de Dialect 5-8 y Dialect 9-12, además de las pruebas TEVI y EVOC que complementan el diagnóstico con información de la cantidad de vocabulario que manejan los estudiantes.

Evidencias de validez y confiabilidad

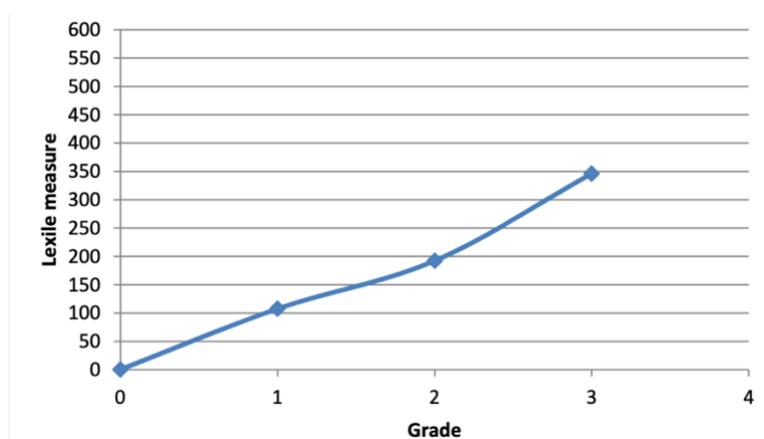
Validez. La validez se refiere al grado en que la evidencia y la teoría sustentan las interpretaciones de puntajes en el uso de las pruebas (AERA, APA, NCME, 2014). La validez de un instrumento entrega información acerca de qué tan bien dicho instrumento cumple con su función; es decir, mide lo que tiene que medir (Price, 2017). Existen distintas maneras de establecer la validez de un instrumento de medición, todas las cuales se van perfeccionando en la medida en que se obtienen más datos y realizan más análisis.

Validez de contenido. Este tipo de validez se asocia a un determinado dominio, contenido, tema, y las inferencias que se pueden realizar desde puntajes obtenidos en una prueba en relación a dicho contenido o dominio (Price, 2018). En el caso de Dialect, el objetivo es medir

comprensión lectora en textos literarios y no literarios. La prueba contiene pasajes de este tipo de textos, y se utilizó un procedimiento denominado *Spanish Lexile Analyzer* para asegurar que existiera gradualidad en la progresión de los textos que los estudiantes deben leer para responder las preguntas. De este modo, los textos, las preguntas y la estructuración de la prueba apuntan a medir comprensión lectora con textos adecuados, debidamente calibrados en función de dificultad textual, y rigurosamente revisados por jueces expertos.

Validez de constructo. La validez de constructo se refiere a la manera en que el contenido de una prueba mide adecuadamente un constructo, en este caso, la comprensión lectora en español. La validez de constructo analiza la estructura interna de una prueba y el grado en el que los distintos ítems de ésta se relacionan y se configuran respecto del constructo. La validez de constructo de Dialect se basa en el uso del Sistema Lexile para Leer que ha demostrado ser un procedimiento efectivo para crear mediciones de comprensión lectora. Una evidencia de esta validez de constructo se observa en el hecho de que Dialect logra capturar la naturaleza evolutiva de la comprensión lectora (Anastasi, 1982). Esto significa que el instrumento debe ser capaz de reflejar incrementos en los puntajes de quienes toman la prueba conforme avanza su edad. La lectura es una habilidad que mejora con la edad. Por otra parte, la evidencia indica que existen períodos de mayor crecimiento en la habilidad lectora, como son los primeros 2-3 años de escolaridad formal. Los análisis realizados con datos obtenidos de aplicaciones de Dialect muestran que, a medida que aumenta el curso (y por ende la edad) de los estudiantes, aumenta también su desempeño en Dialect, tal como lo muestra la figura 6.

Figura 6.
Puntajes promedio Dialect por curso (N=142), 2012.



Este incremento monótonico es evidencia de la naturaleza evolutiva de la lectura, capturada por Dialect, y especialmente por el uso de la medida Lexile.

De la misma forma, la comprensión lectora ha sido definida como un proceso de construcción de significado a partir de un texto (RAND, 2002; Landi, 2010; Shankweiler, 1989). Sabemos que la construcción de significado depende de que el lector sea capaz de desarrollar eficientemente distintos subprocesos como decodificar, leer fluidamente, conocer el significado de las palabras en el texto y saber utilizar algunas estrategias de comprensión, como predecir, inferir, o secuenciar. Se esperaría, por tanto, que hubiera una correlación moderada entre el resultado en una evaluación de comprensión lectora y el resultado de evaluaciones que midan algunos de estos subprocesos. En 2012 se observaron correlaciones positivas, moderadas y altas entre Dialect (comprensión lectora) y algunos de los subtests que miden desempeño en subprocesos lectores, tal como lo muestra la tabla 2.

Tabla n. 2.
Correlaciones entre Dialect 1 y Subtests Dialect 2 (N=119).

Subtest	N	Correlación con Comp. Lectora
TEVI (Vocabulario)	119	.655**
Conciencia de lo impreso	119	.569**
Segmentación fonémica	119	.560**
Aislar fonema inicial	119	.533**
Rima	119	.503**
Segmentación fonémica	119	.358**
Alfabeto	119	.524**
Identificación de palabras	119	.745**

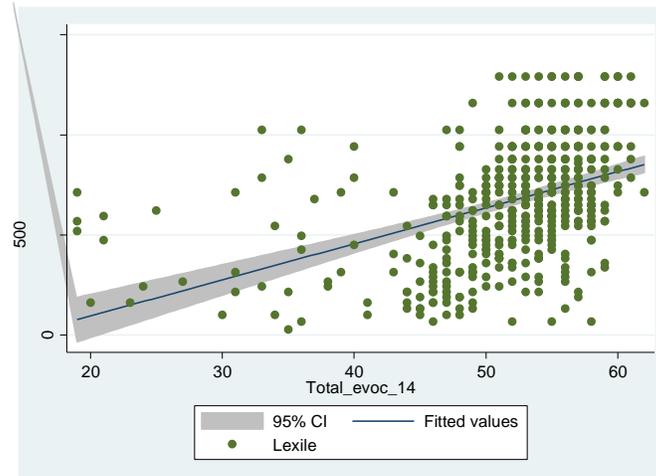
La tabla muestra una alta correlación entre comprensión lectora e identificación de palabras, lo que es consistente con otros estudios sobre comprensión en español (e.g, Infante, Coloma & Himmel, 2012) e inglés (e.g., Perfetti & Hart, 2001). En segundo lugar de magnitud aparece la correlación entre comprensión y vocabulario, también ampliamente validada en la literatura especializada (Kamil & Hiebert, 2005; Tannebaum, Torgesen & Wagner, 2006).

Validez concurrente

En un estudio (Orellana, Valenzuela, Kung, Elmore, & Stenner, 2020) examinamos la validez concurrente entre las pruebas EVOC y TEVI, que forman parte de la plataforma. Ambas miden vocabulario de manera similar. Las pruebas, aplicadas a 400 estudiantes mostraron una correlación de Pearson pequeña pero positiva y significativa (.34**). A su vez, la validez concurrente de la prueba de comprensión lectora de Dialect se midió con EVOC, observándose que un aumento en el puntaje de EVOC explica, en promedio, un incremento de 18 Lexiles en comprensión lectora ($p < .000$). El gráfico n. 1 ilustra esta relación.

Gráfico n. 1

Validez concurrente entre EVOC y comprensión lectora.



Validez predictiva. La validez predictiva se refiere a la capacidad de un puntaje de una determinada prueba de predecir el desempeño futuro. Por ejemplo, una prueba de vocabulario en Kinder puede predecir los resultados en comprensión lectora de esos estudiantes 3 o 4 años después. Un análisis de datos obtenidos de mediciones sucesivas usando Dialect (Orellana, en prensa) permitió establecer que las pruebas de Dialect que evaluaron conocimiento de lo impreso y decodificación en Kinder predicen la comprensión lectora de los estudiantes en segundo año básico. Otros análisis realizados previamente en contextos de vulnerabilidad identificaron el conocimiento del alfabeto y la decodificación en Kinder (ambos medidos con las pruebas Dialect) como principales predictores de la comprensión lectora en primer y segundo año básico.

Confiabilidad

La confiabilidad se refiere a la consistencia tanto interna como externa de un instrumento de medición. La consistencia interna es el hecho de que todas las partes o ítems de un instrumento contribuyan de igual forma a lo que está siendo medido. Análisis como el Split-half reliability permiten determinar la consistencia interna. En el caso de las pruebas Dialect, diversos estudios ratifican una buena consistencia interna, tal como lo muestra la tabla n. 3. En el caso de la prueba EVOC de vocabulario se obtuvo un alfa de .86, demostrando alta consistencia interna. Está pendiente el análisis de confiabilidad para las pruebas de conocimiento del alfabeto, comprensión auditiva, decodificación, concepto de lo impreso y conciencia fonológica.

Tabla n. 3.
Consistencia Interna Pruebas Dialect.

Prueba	Alfa de Cronbach
Dialect comprensión lectora	.91
Conocimiento Alfabeto	pend.
Comprensión Auditiva	pend.
Decodificación	pend.
Concepto de lo Impreso	pend.
Conciencia Fonológica	pend.
Vocabulario	.86

Referencias

- American Educational Research Association, American Psychological Association, and National Council on Measurement in Education. (2014). *Standards for educational and psychological testing*. Washington, DC. American Educational Research Association.
- Cunningham, J.W. (1993). Whole-to-part reading diagnosis. *Reading and Writing Quarterly: Overcoming Learning Difficulties*, 9: 31-49.
- Echeverría, M., Herrera, M.O., & Segure, M.T. (1995) *TEVI-R Test de Vocabulario en Imágenes*. Concepción, Chile: Editorial Universidad de Concepción.
- Infante, M., Coloma, C.J. & Himmel, E. (2012). Comprensión lectora, comprensión oral y decodificación en escolares de segundo y cuarto básico de escuelas municipales. *Estudios Pedagógicos XXXVIII*, 1, 149-160.
- Kamil, M. y Hiebert, E. (2005). Teaching and learning vocabulary: Perspectives and persistent issues. En E. H. Hiebert y M. L. Kamil (Eds.), *Teaching and learning vocabulary: Bringing research to practice* (pp. 1-23). Mahwah: Lawrence Erlbaum.
- Landi N. (2010). An examination of the relationship between reading comprehension, higher-level and lower-level reading sub-skills in adults. *Reading and writing*, 23(6), 701–717. <https://doi.org/10.1007/s11145-009-9180-z>
- McKenna, M. & Stahl, K.D. (2015). *Assessment for reading instruction*, 3rd Ed. New York, NY: Guilford Press.
- Orellana, P. (en prensa). Trayectorias de crecimiento lector en estudiantes chilenos de segundo grado con y sin dificultades lectoras iniciales. En En K. Hess & L. Alarcón (Eds.) *Desarrollo Lingüístico Tardío en Poblaciones Hispanohablantes*. Mexico.
- Orellana-García, P., Valenzuela-Hasenohr, M. F., Kung, M., Elmore, J., & Stenner, A. J. (2020). EVOC: un Instrumento para Evaluar Vocabulario en Español en Chile. *Magis, Revista Internacional De Investigación En Educación*, 13(Primera parte dossier Etnografía), 1-26. <https://doi.org/10.11144/Javeriana.m13.eiev>

- Perfetti CA, Hart L. The lexical basis of comprehension skill. In: Gorfien DS, editor. *On the consequences of meaning selection: Perspectives on resolving lexical ambiguity*. American Psychological Association; Washington, DC: 2001. pp. 67–86.
- Price, L. R. (2017). *Psychometric Methods. Theory into Practice*. New York: Guilford Press.
- Shankweiler D. How problems of comprehension are related to difficulties in decoding. In: Shankweiler D, Liberman IY, editors. *Phonology and reading disability: Solving the reading puzzle*. The University of Michigan Press; Ann Arbor, MI: 1989. pp. 35–68.
- Tannenbaum, K. R., Torgesen, J. K. y Wagner, R. K. (2006). Relationships between word knowledge and reading comprehension in third-grade children. *Scientific Studies of Reading*, 10, 381-98.