

Universidad Nacional de La Plata
Facultad de Psicología
Carrera de Doctorado En Psicología

Tesis de Doctorado

Título

“Validez, confiabilidad y datos normativos de un test breve para la medición de la eficacia lectora en alumnos de escuela primaria”

Lic. Norma Nancy China

Director: Psic. Telma Piacente

Año 2018

***“VALIDEZ, CONFIABILIDAD Y DATOS
NORMATIVOS DE UN TEST BREVE PARA LA
MEDICIÓN DE LA EFICACIA LECTORA EN
ALUMNOS DE ESCUELA PRIMARIA”***

Dedicatoria

Dedico este trabajo a mi hija Kumi, a mi compañero Carlos y a mis padres, Marta y Teiichi.

Agradecimientos

Deseo expresar mi profundo agradecimiento a la profesora Telma Piacente, directora de esta tesis, por transmitirme sus conocimientos y experiencia, y especialmente, su desvelo por la niñez desfavorecida.

Al Dr. Aldo Ferreres, por impulsar, guiar y acompañar mi formación como investigadora y docente.

A la Dra. Valeria Abusamra, por su apoyo incondicional desde el inicio de esta investigación.

A las Lic. Micaela Difalcis, Florentina Morello García, Agustina Miranda, Andrea Casajús, a las Dras. Bárbara Sampedro y Romina Cartoceti, y a todos los que colaboraron en la obtención de datos para los estudios de confiabilidad y de elaboración de los datos normativos.

A las Dras. Silvia Jacobovich y Rufina Pearson, a las Lic. Bárbara Gottheil y Nancy Mantiñán por su contribución al estudio de validez de criterio.

A los Dres. Mario Squillace e Irene Injoque Ricle Giberti, y a la Lic. Samanta Leiva, por su asesoramiento acerca de procesamiento estadístico.

A las autoridades de los distritos escolares, a las directoras y maestros que abrieron las puertas de sus escuelas y sus aulas, a los niños y niñas participantes de esta investigación, y a sus padres, por haberme permitido realizar las evaluaciones para esta tesis.

Índice

Contenidos	Pág.
Resumen	1
Abstract	5
Introducción	8
Planteo del problema.....	12
Justificación del tema.....	13
Organización de la tesis.....	16
Capítulo 1. Marco teórico y antecedentes	19
1.1. Los sistemas de escritura.....	21
1.1.1. El modelo de doble ruta para el reconocimiento de las palabras escritas.....	23
1.2. El aprendizaje de la lectura en los sistemas alfabéticos.....	26
1.2.1. Los métodos de enseñanza de la lectura en los sistemas alfabéticos: el método global versus el método fónico.....	31
1.3. Las dificultades en el aprendizaje de la lectura.....	32
1.3.1. La dislexia en español.....	37
1.3.2. Fuentes de variabilidad que influyen en la manifestación de las dificultades en el aprendizaje de la lectura.....	39
1.4. La identificación de las dificultades específicas en el aprendizaje de la lectura.....	40
1.4.1. El enfoque basado en la discrepancia entre la capacidad intelectual general y el rendimiento académico.....	41
1.4.2. El enfoque basado en la respuesta a la intervención.....	44
1.4.2.1. Detección del riesgo. La efectividad en los tests de <i>screening</i> . Sensibilidad y especificidad.....	46
Capítulo 2. El Test Colectivo de Eficacia Lectora. Su adaptación	49
2.1. Elaboración y adaptación del TECLE al español rioplatense.....	50
2.1.1. Adaptación destinada a la población uruguaya.....	50
2.1.2. Adaptación destinada a población argentina.....	50
2.1.3. Modificaciones ulteriores realizadas para esta tesis.....	51

2.2. El constructo “eficacia lectora”: interpretación de los puntajes del test.....	52
2.2.1 Tests de potencia versus tests de velocidad.....	53
2.3. La presente investigación.....	53
2.3.1. Interrogantes que orientaron la investigación.....	54
2.3.2. Objetivos.....	54
2.3.3. Hipótesis.....	55
2.3.4. Metodología.....	56
2.3.4.1. Participantes, instrumentos y procedimientos de obtención, elaboración y análisis datos.....	56

Capítulo 3. Estudio de confiabilidad del TECLE Buenos Aires.....

3.1. Metodología.....	61
3.1.1. Participantes.....	62
3.1.2. Instrumentos.....	62
3.1.3. Procedimientos.....	62
3.2. Resultados.....	63
3.2.1. Estadísticos descriptivos.....	63
3.2.2. Pruebas de normalidad y homocedasticidad.....	64
3.2.3. Coeficiente de confiabilidad test-retest.....	65
3.3. Discusión y conclusiones.....	66

Capítulo 4. Estudios de validez del TECLE Buenos Aires.....

4.1. Estudio de validez de constructo.....	70
4.1.1. Metodología.....	70
4.1.1.1. Participantes.....	70
4.1.1.2. Instrumentos.....	71
4.1.1.3. Procedimientos.....	78
4.1.2. Resultados.....	79
4.1.2.1. Estadísticos descriptivos.....	79
4.1.2.2. Prueba de normalidad de las distribuciones.....	80
4.1.2.3. Análisis correlacionales entre los puntajes observados en el TECLE Buenos Aires y otras medidas relacionadas.....	81
4.1.3. Discusión y conclusiones del estudio de validez de constructo.....	84

4.2. Estudio de validez de criterio.....	87
4.2.1. Metodología.....	87
4.2.1.1. Participantes.....	87
4.2.1.2. Instrumento.....	88
4.2.1.3. Procedimientos.....	88
4.2.2. Resultados.....	89
4.2.2.1. Estadísticos descriptivos.....	89
4.2.2.2. Prueba de normalidad y homocedasticidad.....	89
4.2.2.3. Análisis de las diferencias entre las puntuaciones..	90
4.2.3. Discusión y conclusiones del estudio de validez de criterio	90
4.3. Discusión y conclusiones generales de los estudios de validez..	91

Capítulo 5. Los datos normativos del TECLE Buenos Aires para la Ciudad Autónoma de Buenos Aires y alrededores. Justificación estadística.....

5.1. Metodología.....	95
5.1.1. La muestra de tipificación.....	96
5.1.2. Instrumento.....	100
5.1.3. Procedimientos.....	100
5.2. Resultados.....	102
5.2.1. Desempeño en la prueba según grado escolar.....	103
5.2.1.1. Estadísticos descriptivos.....	103
5.2.1.2. Pruebas de normalidad y homocedasticidad.....	103
5.2.1.3. Análisis de las diferencias entre las puntuaciones..	104
5.2.2. Desempeño en la prueba según edad.....	105
5.2.2.1. Estadísticos descriptivos.....	105
5.2.2.2. Pruebas de normalidad y homocedasticidad.....	106
5.2.2.3. Análisis de las diferencias entre las puntuaciones..	106
5.2.3. Desempeño en la prueba según grado escolar versus edad.....	108
5.2.3.1. Estadísticos descriptivos según grado escolar y edad.....	108
5.2.3.2. Pruebas de normalidad.....	108
5.2.3.3. Análisis de las diferencias entre las puntuaciones según grado escolar y edad.....	109

5.2.3.3.1. Prueba de homocedasticidad y análisis de las diferencias entre las puntuaciones según la edad para cada grado escolar....	109
5.2.3.3.2. Prueba de homocedasticidad y análisis de las diferencias entre las puntuaciones según el grado escolar para los grupos de 8, 9, 10 y 11 años.....	110
5.2.4. Desempeño en la prueba según sexo.....	110
5.2.4.1. Estadísticos descriptivos.....	110
5.2.4.2. Pruebas de normalidad y homocedasticidad.....	111
5.2.4.3. Análisis de las diferencias entre las puntuaciones..	111
5.2.5. Desempeño en la prueba según nivel de oportunidades educativas.....	111
5.2.5.1. Estadísticos descriptivos.....	111
5.2.5.2. Pruebas de normalidad y homocedasticidad.....	112
5.2.5.3. Análisis de las diferencias entre las puntuaciones..	112
5.2.6. Desempeño en la prueba según tipo de gestión.....	113
5.2.6.1. Estadísticos descriptivos.....	113
5.2.6.2. Pruebas de normalidad y homocedasticidad.....	114
5.2.6.3. Análisis de las diferencias entre las puntuaciones..	114
5.2.7. Desempeño en la prueba según región geográfica.....	114
5.2.7.1. Estadísticos descriptivos.....	114
5.2.7.2. Pruebas de normalidad y homocedasticidad.....	115
5.2.7.3. Análisis de las diferencias entre las puntuaciones..	115
5.2.8. Resumen de los resultados del estudio del desempeño en la prueba según grado escolar, edad, sexo, nivel de oportunidades educativas, tipo de gestión y región geográfica.....	116
Capítulo 6. Discusión y conclusiones de la tesis.....	119
6.1. Síntesis de los resultados.....	119
6.2. Discusión general.....	120
6.2.1. Interpretación de los puntajes observados en el TECLE Buenos Aires. Una medida de eficacia lectora.....	122
6.2.2. Usos propuesto de los puntajes observados en el TECLE Buenos Aires. Detección del riesgo de desarrollar dificultades en el aprendizaje de la lectura. Evaluación de la calidad de la enseñanza de la lectura.....	124

6.3. Limitaciones de la tesis.....	127
6.4. Proyección para futuras investigaciones.....	128
6.5. Conclusiones.....	128
Referencias.....	130
Anexo A. Análisis no paramétricos.....	142
A.1. Estudio de confiabilidad test-retest.....	142
A.2. Estudio de del efecto de variables sobre el rendimiento en el test en la muestra para la obtención de los baremos.....	143
A.2.1. Efecto del grado escolar sobre el rendimiento en el test...	143
A.2.2. Efecto de la edad sobre el rendimiento en el test.....	144
A.2.3. Efecto del grado escolar versus la edad sobre el rendimiento en el test.....	144
A.2.4. Efecto del sexo sobre el rendimiento en el test.....	145
A.2.5. Efecto del nivel de oportunidades educativas (NOE) sobre el rendimiento en el test.....	145
A.2.6. Efecto del tipo de gestión sobre el rendimiento en el test..	146
A.2.7. Efecto de la región geográfica sobre el rendimiento en el test.....	146
Anexo B. Baremos del TECLE Buenos Aires para la Ciudad Autónoma de Buenos Aires y alrededores.....	148
B.1. Medidas de posición, rango intercuartílico y baremos en rangos percentiles para los grupos según grado escolar.....	149
B.2. Medidas de posición, rango intercuartílico y baremos en rangos percentiles para los grupos según grado escolar y nivel de oportunidades educativas.....	155
Anexo C. El TECLE Buenos Aires. Protocolo de respuestas y normas de aplicación y calificación.....	161
C.1. TECLE Buenos Aires: protocolo de respuestas.....	162
C.2. TECLE Buenos Aires: normas de aplicación y calificación.....	168
C.2.1. Normas de aplicación del test.....	168
C.2.2. Normas de calificación.....	171
C.3. Planilla de desempeño en el TECLE Buenos Aires de los alumnos del curso.....	174

Anexo D. Instrumentos administrados en el estudio de validez de constructo	175
D.1. Prueba de denominación automatizada rápida de dibujos del test DAR (Morello García, China, & Ferreres, 2015).....	175
D.2. Láminas de prueba de las partes 1 y 2 de la prueba de trazado de senderos diseñada para esta tesis (China, no publicado).....	176
D.2.1. Lámina de prueba de la parte 1.....	176
D.2.2. Lámina de prueba de la parte 2.....	177
D.3. Prueba de <i>screening</i> sobre el texto expositivo “Un almuerzo prehistórico” del Test Leer para Comprender (Abusamra, Ferreres, Raiter, De Beni, & Cornoldi, 2010).....	178

Índice de tablas

Tablas	Pág.
Capítulo 3. Estudio de confiabilidad del TECLE Buenos Aires	
Tabla 3.1. Media, desvío estándar y varianza de los puntajes observados en el test y el retest en la muestra total.....	63
Tabla 3.2. Media, desvío estándar y varianza de los puntajes observados en el test y el retest por grado escolar.....	64
Capítulo 4. Estudios de validez del TECLE Buenos Aires	
Tabla 4.1. Media, desvío estándar y varianza de los puntajes observados en el TECLE Buenos Aires y de medidas teóricamente relacionadas.....	80
Tabla 4.2. Correlaciones (coeficientes rho de Spearman) entre el rendimiento en el TECLE Buenos Aires y en 11 medidas de interés de la muestra completa.....	82
Tabla 4.3. Correlaciones (coeficientes rho de Spearman) entre el rendimiento en el TECLE Buenos Aires y en las medidas basadas en el número de aciertos en los grupos de 3°-4° y 5°-6° grados.....	83
Tabla 4.4. Correlaciones (coeficientes rho de Spearman) entre el rendimiento en el TECLE Buenos Aires y en las medidas del tiempo de ejecución en los grupos de 3°-4° y 5°-6° grados.....	84
Tabla 4.5. Correlaciones (coeficientes rho de Spearman) entre el rendimiento de la muestra completa en el TECLE Buenos Aires y el resto de las medidas, presentadas en orden decreciente, y agrupadas en términos de los procesos compartidos.....	86
Tabla 4.6. Media, desvío estándar y varianza de las puntuaciones observadas en el TECLE Buenos Aires en ambos grupos.....	89
Capítulo 5. Los datos normativos del TECLE Buenos Aires para la Ciudad Autónoma de Buenos Aires y alrededores. Justificación estadística	
Tabla 5.1. Distribución de las escuelas por tipo de gestión, región geográfica y nivel de oportunidades educativas.....	100
Tabla 5.2. Distribución de los participantes según sexo por cada grado escolar.....	100
Tabla 5.3. Distribución de los participantes según región geográfica para cada grado escolar.....	100

Tabla 5.4. Distribución de los participantes según tipo de gestión para cada grado escolar.....	101
Tabla 5.5. Distribución de los participantes según el nivel de oportunidades educativas.....	101
Tabla 5.6. Media, desvío estándar y varianza de los puntajes observados en los grupos según el grado escolar.....	104
Tabla 5.7. Diferencias entre las medias de los puntajes observados en los grupos según el grado escolar (prueba de Games-Howell).....	106
Tabla 5.8. Media, desvío estándar y varianza de los puntajes observados en los grupos según la edad en años.....	107
Tabla 5.9. Diferencias entre las medias de los puntajes observados en los grupos según la edad (prueba de Games-Howell).....	108
Tabla 5.10. Media, desvío estándar y varianza de los puntajes observados en los grupos según el grado escolar y la edad.....	109
Tabla 5.11. Media, desvío estándar y varianza de los puntajes observados en los grupos según el sexo.....	112
Tabla 5.12. Media, desvío estándar y varianza de los puntajes observados en los grupos según el nivel de oportunidades educativas.....	113
Tabla 5.13. Diferencias entre las medias de los puntajes observados en los grupos según el nivel de oportunidades educativas (prueba de Games-Howell).....	114
Tabla 5.14. Media, desvío estándar y varianza de los puntajes observados en el grupo que asistía a escuelas con NOE medio segmentado según el tipo de gestión.....	114
Tabla 5.15. Media, desvío estándar y varianza de los puntajes observados en el grupo que asistía a escuelas de gestión estatal segmentado según la región.....	116
Tabla 5.16. Diferencias entre las medias de los puntajes observados en el grupo que asistía a escuelas de gestión estatal segmentado según la región (prueba de Games-Howell)	117
Tabla 5.17. Tamaño del efecto de las variables independientes sobre el desempeño en el TECLE Buenos Aires.....	118

Anexo A. Análisis no paramétricos

A.1. Estudio de confiabilidad test-retest

Tabla A.1. Mediana y rango intercuartílico de los puntajes observados en el test y el retest por grado escolar.....	143
---------------------------------------------------------------------------------------------------------------------	-----

A.2. Estudio del efecto de variables sobre el rendimiento en el test en la muestra para la obtención de los baremos	
Tabla A.2. Mediana y rango intercuartílico de los puntajes observados en los grupos según el grado escolar.....	144
Tabla A.3. Mediana y rango intercuartílico de los puntajes observados en los grupos según la edad en años.....	145
Tabla A.4. Mediana y rango intercuartílico de los puntajes observados en los grupos según el grado escolar y la edad.....	146
Tabla A.5. Mediana y rango intercuartílico de los puntajes observados en el grupo que asistía a escuelas de gestión estatal segmentado según la región.....	147

Anexo B. Baremos del TECLE Buenos Aires para la Ciudad Autónoma de Buenos Aires y alrededores

Tabla B.1. Medidas de posición (P16, P25, P50, P75) y rango intercuartílico de los puntajes directos observados en cada grupo según grado escolar.....	150
Tabla B.2. Rangos percentiles de los puntajes directos observados en 2° grado.....	151
Tabla B.3. Rangos percentiles de los puntajes directos observados en 3° grado.....	152
Tabla B.4. Rangos percentiles de los puntajes directos observados en 4° grado.....	153
Tabla B.5. Rangos percentiles de los puntajes directos observados en 5° grado.....	154
Tabla B.6. Rangos percentiles de los puntajes directos observados en 6° grado.....	155
Tabla B.7. Medidas de posición (P16, P25, P50, P75) y rango intercuartílico de los puntajes directos observados en cada grupo según grado escolar y nivel de oportunidades educativas.	156
Tabla B.8. Rangos percentiles de los puntajes directos observados en los grupos 2° grado con NOE bajo, medio y alto.....	157
Tabla B.9. Rangos percentiles de los puntajes directos observados en los grupos 3° grado con NOE bajo, medio y alto.....	158
Tabla B.10. Rangos percentiles de los puntajes directos observados en los grupos 4° grado con NOE bajo, medio y alto.....	159
Tabla B.11. Rangos percentiles de los puntajes directos observados en los grupos 5° grado con NOE bajo, medio y alto.....	160
Tabla B.12. Rangos percentiles de los puntajes directos observados en los grupos 6° grado con NOE bajo, medio y alto.....	161

Resumen

El objetivo de la presente tesis fue contribuir al desarrollo de tests de *screening* para la evaluación de la lectura destinados al monitoreo del curso de los aprendizajes y a la detección oportuna del riesgo de desarrollar dificultades eventuales.

Los tests de *screening* son un componente fundamental del modelo de identificación de las dificultades del aprendizaje basado en la respuesta a la enseñanza/intervención, conocido como modelo RTI (por su denominación en inglés, *Response to Instruction/Intervention*). El modelo enfatiza la distinción entre el bajo desempeño en lectura debido a un trastorno específico del aprendizaje del bajo desempeño debido a otras causas, como enseñanza inadecuada o insuficiente o desventaja socio-económica-cultural; en el primer caso, se observa una persistencia de las dificultades pese a la provisión de una enseñanza adecuada y a los intentos de intervención intensiva y sistemática. Así, se considera que la identificación de los trastornos específicos es válida solo cuando se ha constatado una falta/insuficiencia de mejoría en las habilidades de lectura luego de la intervención.

En el modelo RTI, el procedimiento de identificación comienza con la administración universal de un instrumento de *screening*, en los primeros años de escolarización y en el ámbito escolar. Tiene por objeto la detección temprana de quienes presentan riesgo de desarrollar o presentar dificultades en el aprendizaje de la lectura y su incorporación oportuna en un programa de enseñanza/intervención intensiva y sistemática con estrategias probadamente eficaces. A lo largo y al final del programa, se monitorea el progreso del niño en lectura. Si su desempeño no muestra niveles aceptables de mejoría, se introducen cambios en el programa de enseñanza/intervención. El entrenamiento finaliza cuando el niño alcanza el nivel de sus pares, o bien cuando persiste el bajo rendimiento luego de sucesivos intentos de remediación. En este caso, se considera la derivación

del niño a una instancia de evaluación cognitiva en el ámbito clínico dirigida a obtener un diagnóstico de su condición.

En virtud de los beneficios de adoptar un modelo RTI que permite la detección temprana y oportuna de las dificultades en el aprendizaje, y de la necesidad de contar con un test de *screening* para la evaluación de la lectura, se adaptó un instrumento existente, el Test Colectivo de Eficacia Lectora (TECLE) a las particularidades del español de la Ciudad Autónoma de Buenos Aires y alrededores (TECLE Buenos Aires). El test original fue desarrollado para la población española (Marín & Carrillo, 1999). Los autores afirman que el rendimiento en el TECLE depende de habilidades específicas para la lectura y habilidades generales lingüísticas y cognitivas.

El test se presenta como una herramienta breve y económica en cuanto al material necesario para su aplicación y la capacitación de los usuarios en técnicas de evaluación. Se caracteriza como un test de velocidad y de ejecución máxima, por la facilidad de los ítems que lo componen, a los que el sujeto debe responder en un tiempo límite de 5 minutos. La tarea consiste en completar la mayor cantidad de oraciones a las que les falta la última palabra seleccionando, entre cuatro opciones de respuesta, aquella que la completa coherentemente. Las opciones son ortográfica y fonológicamente similares, y consisten en dos palabras (la opción correcta y otra palabra que no completa coherentemente la oración) y dos pseudopalabras.

En consonancia con las recomendaciones de la Comisión Internacional de Tests para la adaptación de un instrumento de un contexto lingüístico y cultural a otro (International Test Commission, 2017; Muñiz, Elosua, & Hambleton, 2013), se realizaron los estudios dirigidos a obtener evidencias de confiabilidad de los puntajes que el test arroja y de validez de sus interpretaciones y usos, y el estudio destinado a obtener los datos normativos necesarios para interpretar adecuadamente el desempeño de los sujetos evaluados.

Para el *estudio de confiabilidad* se adoptó el método test-retest, con un intervalo de 5 semanas entre ambas instancias de evaluación, en una muestra de 218 niños de 3° a 6° grados. Los resultados mostraron un alto nivel de estabilidad temporal de los puntajes observados en el TECLE Buenos Aires, con un coeficiente de correlación $r = .90$ entre ambas mediciones.

El *estudio de validez de constructo* consistió en el análisis de las asociaciones entre el desempeño en el TECLE Buenos Aires y medidas de las habilidades de decodificación, específicas para la lectura, así como medidas de habilidades generales lingüísticas y cognitivas, obtenidos de una muestra de 57 participantes de 3° a 6° grados. Los análisis revelaron que el desempeño en el test está fuertemente asociado con medidas de las habilidades de decodificación, y en menor grado, con el resto de las medidas estudiadas.

El *estudio de validez de criterio* estuvo dirigido a determinar las diferencias en el desempeño entre un grupo clínico de 28 sujetos de 2° a 6° grados con diagnóstico de dificultades en el aprendizaje y/o del lenguaje, y otro grupo control de sujetos con desarrollo típico, equiparados uno-a-uno con los sujetos del grupo clínico en términos del grado escolar, la edad, el sexo y el nivel de oportunidades educativas. La comparación de los puntajes observados mostró diferencias estadísticamente significativas entre ambos grupos, con un peor rendimiento del grupo clínico, y un tamaño del efecto mediano.

Para la obtención de datos normativos, se evaluó el desempeño de 3097 participantes de 2° a 6° grados que asistían a 24 escuelas de la Ciudad Autónoma de Buenos Aires y alrededores. Se estudió el efecto del grado escolar, la edad y el sexo de los participantes, y el nivel de oportunidades educativas, el tipo de gestión y la región geográfica de las escuelas, con el propósito de confeccionar los baremos en función de las variables que produjeran las mayores diferencias entre los grupos. El estudio reveló que tales variables fueron el grado escolar y el nivel de oportunidades

educativas. En consecuencia, se confeccionaron dos tipos de datos normativos: baremos segmentados en función del grado escolar, y en función del grado escolar y el nivel de oportunidades educativas.

Concluimos que el TECLE Buenos Aires constituye un instrumento de *screening* válido y confiable para la detección temprana del riesgo de desarrollar o presentar dificultades en el aprendizaje de la lectura.

Palabras clave: validez, confiabilidad, baremos, *screening*, lectura.

Abstract

The aim of this thesis was to contribute to the development of screening tests for reading assessment intended to monitor learning trajectories and to timely detect the risk for developing eventual difficulties.

Screening tests are a fundamental component of an identification model of learning disabilities based on response to instruction/intervention (RTI model). This identification model emphasize the distinction between low achievement in reading due to specific learning disability and low achievement due to other causes, like inappropriate or insufficient instruction or socio-economical-cultural disadvantage; in the first case, the difficulty persists despite the provision of appropriate instruction and attempts of intensive and systematic intervention. Thus, the identification of specific learning disability becomes valid only after a lack/scarcity of progress in reading ability following the intervention has been confirmed.

The identification procedure in the RTI model begins with the early universal screening, in the school setting, designed for early detecting those children at risk for developing or presenting reading learning disabilities as well as for timely including them in a program of intensive and systematic instruction/intervention using proven effective strategies. The child's reading progress is measured throughout and at the end of the program. If reading achievement doesn't show acceptable levels of improvement, changes are introduced in the instruction/intervention program. Training ends when the child's reading achievement either meets that of her/his classmates, or remains low after successive attempts to remediation. In this case, the clinical referral for cognitive assessment is considered in order to obtain a diagnosis of her/his condition.

By virtue of the benefits of adopting an RTI model that allows early and timely detection of learning difficulties, and the need to have a screening test available for the reading assessment, we adapted an existing instrument, the

Test Colectivo de Eficacia Lectora, TECLE test (Collective test of reading efficiency) developed for Spanish population (Marín & Carrillo, 1999), to the lexical-syntactic peculiarities of Spanish spoken in Buenos Aires City and surroundings (TECLE Buenos Aires test). The authors claim that the performance in the TECLE test depends on specific reading and general linguistic and cognitive abilities.

The test is proposed as a brief and economical tool in terms of the materials required for administration and of users training in assessment techniques. It is characterized as a maximum performance, speed test, given the ease of the items it includes, to which the examinee must respond within a time limit of 5 minutes. The task is to complete, as much as possible, sentences lacking the last word, by selecting the word that consistently completes the sentence, out from four response options. The options are orthographically and phonologically similar and include two words (the correct one and another that does not consistently complete the sentence) and two pseudowords.

In line with the recommendations given by the International Test Commission for adapting a test from one linguistic and cultural environment to another (International Test Commission, 2017; Muñiz, Elosua, & Hambleton, 2013), we carried out the studies aimed at obtaining evidences of reliability of the scores in the test, evidences of validity of the interpretations and uses of the scores, and a study to obtain normative data necessary for appropriately interpreting the examinees' performance.

We adopted the test-retest method for the *reliability study*, with 5 weeks between both administrations, in a sample of 218 third to sixth grade children. The results showed a high level of temporal stability of the scores observed in the TECLE Buenos Aires test, with a correlation coefficient $r = .90$ between both measures.

The *construct validity study* consisted of the analyses of the associations between the performance in the TECLE Buenos Aires test and measures of reading specific decoding abilities and linguistic and cognitive

general abilities, obtained from a sample of 57 third to sixth grade participants. The analyses revealed a strong association between the performance in the test and the measures of decoding abilities, and to a lower degree, the other measures included.

The *criterion validity study* aimed to determine the differences in the performances observed in the test from a clinical group conformed by 28 second to sixth grade subjects with diagnosis of learning and/or language disabilities, and from a control group of participants with typical development, one-to-one matched with the subjects from the clinical group in terms of school grade, age, sex, and level of educational opportunities. The comparison of the observed scores showed statistically significant differences between both groups, with poorer performance from the clinical one, and a medium effect size.

Normative data were obtained from a sample of 3097 second to sixth grade children from 24 school establishments from Buenos Aires City and surroundings. We studied the effect of children's school grade, age, and sex, and level of educational opportunities, type of management, and geographic region of the schools, in order to construct the tables of normative data segmented in terms of those variables producing the greatest differences between groups. The study revealed that school grade and level of educational opportunities emerged as such variables. Consequently, we present two type of tables of normative data: tables segmented according to school grade and according to school grade and level of educational opportunities.

We conclude that the TECLE Buenos Aires test is a valid and reliable screening tool for early detecting the risk of developing or presenting reading learning difficulties.

Key words: validity, reliability, norms, screening, reading.

Introducción

La tesis que se presenta está destinada al estudio de las propiedades psicométricas y a la elaboración de los datos normativos de un test breve de eficacia lectora, el TECLE Buenos Aires (Ferrerres, Abusamra, Casajús, & China, 2011), para la Ciudad Autónoma de Buenos Aires y alrededores.

Se inscribe en las líneas de investigación sobre la evaluación del desempeño en lectura destinadas a la prevención/intervención de las eventuales dificultades que pueden presentarse en el curso de su aprendizaje.

Si bien la evaluación es contemporánea de los comienzos de la educación obligatoria, la evolución que ha seguido hasta nuestros días ha estado influenciada por el desarrollo de modelos teóricos robustos de lectura que contemplan las variables involucradas en su dominio progresivo. Ello ha posibilitado el énfasis creciente de investigaciones que posibiliten el diseño de estrategias de intervención preventivas a partir de la detección temprana de posibles déficits en los conocimientos y habilidades de los que disponen los niños, que resulten oportunas, sistemáticas y adecuadas (Snow, Burns, & Griffin, 1998).

Esa detección requiere de instrumentos específicos, que contemplen competencias psicolingüísticas, mnésicas, perceptivas y cognitivas cuya integración permite la lectura. Dicho de otro modo, que vayan más allá de las consideraciones relativas al papel que juegan exclusivamente aspectos cognitivos, como los examinados en las pruebas de inteligencia, traducidos en un puntaje de cociente intelectual (en adelante, CI). Los mismos tradicionalmente han sido tomados en cuenta como un criterio excluyente para determinar el buen desempeño o el desempeño con dificultades.

Asimismo cabe señalar que la consideración de las dificultades de aprendizaje y de los denominados trastornos específicos, ha sufrido la

influencia tanto de los modelos teóricos como de la evidencia empírica acumulada en los últimos 30 años, que han redundado en definiciones más precisas para su caracterización. Entre los modelos más avanzados de identificación de las dificultades de aprendizaje de la lectura, presentes en un número sustantivo de niños, aparecen no solo aquellos que han logrado identificar las características específicas que presentan los sujetos que las detentan, como la explicitación proporcionada por el modelo de doble ruta (Coltheart, 2005; Coltheart, Rastle, Perry, Langdon, & Ziegler, 2001), sino también aquellos que consideran de manera conjunta los procesos de enseñanza y aprendizaje. Se trata en este caso del Modelo de Respuesta a la Intervención (en adelante, RTI, como se lo conoce por su denominación en inglés, *Response to Instruction/Intervention*), que ha resultado de particular interés para esta tesis.

El modelo RTI se basa en la falta de respuesta adecuada a la enseñanza/intervención. Permite diferenciar de manera válida y confiable dos grupos de niños con bajo rendimiento académico, los “respondedores” y los “no respondedores”. En estos últimos, que corresponden a sujetos disléxicos (ver capítulo 1 “Marco teórico y antecedentes” para mayor detalle), las dificultades suelen persistir en la adultez y, si bien es posible observar mejorías sustanciales en la precisión de la decodificación como resultado de la intervención, los beneficios en la fluidez son menos frecuentes y más limitados (Démonet, Taylor, & Chaix, 2004; Fletcher, Lyon, Fuchs, & Barnes, 2007; Shaywitz, Fletcher, Holahan, Shneider, Marchione, Stuebing, Francis, Pugh, & Shaywitz, 1999). Ambos grupos de niños también difieren en términos de su rendimiento académico, de sus correlatos cognitivos y de factores neurobiológicos tales como el patrón de activación cerebral que se observa en estudios de neuroimagen funcional (Fletcher, Simos, Papanicolaou, & Denton, 2004; Stage, Abbott, Jenkins, & Berninger, 2003; Vaughn, Linan-Thompson, & Hickman, 2003; Vellutino, Scanlon, & Jaccard, 2003).

Una ventaja adicional de los modelos RTI es que, a la vez que contribuyen a la identificación temprana de aquellos niños que presentarán déficits persistentes en la lectura, permiten aumentar las oportunidades educativas de todos los niños, reduciendo las posibilidades de aparición de dificultades eventuales. Para ello, proponen implementar el despistaje masivo o *screening* dirigido a todos los niños en cada uno de los dominios de rendimiento académico a fin de detectar a los que presentan características de riesgo de desarrollar dificultades. En estos niños deberían hacerse intentos inmediatos de enseñanza intensiva en el dominio académico específico, seguidos de evaluaciones periódicas breves basadas en el currículo con el objeto de monitorear el progreso en el rendimiento académico.

Más allá de esas consideraciones, desde la perspectiva de la evaluación, en la actualidad es posible distinguir distintos enfoques en relación con el desempeño en lectura, según se trate de diferentes ámbitos de aplicación. En un caso, particularmente el ámbito clínico y de investigación, se trata de la elaboración de instrumentos que proporcionen *indicadores diagnósticos*. En el otro, especialmente en el ámbito educativo, de *indicadores de riesgo*.

Para la utilización apropiada en los ámbitos clínicos y de investigación existen herramientas válidas y confiables en español que incluyen la evaluación de los múltiples componentes de decodificación y de comprensión de la lectura, como la batería PROLEC-R (Cuetos, Rodríguez, Ruano, & Arribas, 2009) y el test LEE (Defior Citoler, Fonseca, Gottheil, Aldrey, Jiménez Fernández, Pujals, Rosa, & Serrano Chica, 2006). Este último cuenta además con baremos para población argentina. Estas herramientas apuntan a la evaluación de la capacidad lectora, de las estrategias utilizadas por cada sujeto en la lectura de palabras, oraciones y textos, y de los procesos en los que se presentan dificultades.

Sin embargo, dada las características de su administración individual, el tiempo que demanda su completamiento y las calificaciones requeridas

para los usuarios de estas pruebas para una correcta interpretación de los resultados, su aplicación en el ámbito educativo resulta costosa para los fines de la detección oportuna del riesgo de desarrollar dificultades en el aprendizaje de la lectura.

En función de ello, una literatura creciente aboga en favor de tests de *screening*, caracterizados por la brevedad de su aplicación y la posibilidad de despistaje que proporcionan. Entre ellos, Marín y Carrillo (1999) diseñaron el Test Colectivo de Eficacia Lectora (en adelante, TECLE) destinado a población española, siguiendo la estructura del test D-OR-LEC de Lobrot (1980) para población francófona. Se trata de un test breve de administración colectiva en el que los sujetos evaluados deben completar la mayor cantidad de oraciones a las que les falta la última palabra en un tiempo máximo de 5 minutos. Cuadro, Costa, Trías y Ponce de León (2009) adaptaron el test y obtuvieron las normas para la población uruguaya. De acuerdo con estos autores, el test arroja una medida de “eficacia lectora”, que incluye componentes de decodificación, comprensión de las oraciones y administración adecuada de los recursos cognitivos del procesamiento lector.

En Argentina, Ferreres, Abusamra, Casajús y China (2011) adaptaron el test a las particularidades léxico-sintácticas del español de Buenos Aires y sus alrededores, versión sobre la que se realizaron modificaciones ulteriores para la presente tesis (en adelante, TECLE Buenos Aires).

Por sus características, el TECLE Buenos Aires, al igual que el TECLE, cumple con los requerimientos de un test de *screening*, componente fundamental de los modelos más actuales de identificación de las dificultades en el aprendizaje.

No obstante, en concordancia con las directrices de la Comisión Internacional de Tests (International Test Commission, 2017), es necesario distinguir entre la traducción y adaptación lingüística de un test a un idioma y región diferente de los de la versión original y la verdadera adaptación de la

prueba, que requiere la constatación de la equivalencia en ambas versiones de todos los aspectos involucrados.

Al respecto, se observa que aún no se han realizado estudios dirigidos a aportar evidencias empíricas de la confiabilidad de los puntajes que arroja y de la validez de la interpretación de los mismos, ni se cuenta con datos normativos específicos.

Planteo del problema

En Argentina, las habilidades de lectura se incluyen en el área de Lengua y la evaluación de los aprendizajes alcanzados por los alumnos a lo largo de la educación primaria se caracteriza por centrarse en el desempeño en comprensión de textos. Esta información aparece ilustrada en el Informe y Análisis de los resultados del XII Operativo Nacional de Evaluación llevado a cabo por la Dirección Nacional de Información y Evaluación de la Calidad Educativa del Ministerio de Educación de la Nación (DiNIECE) del año 2007, en el cual participaron todas las provincias. A partir de dicha información, la DiNIECE también publicó en octubre de 2010 las Recomendaciones Metodológicas para la enseñanza en el área de Lengua y en noviembre del mismo año las Pautas Pedagógicas para elaborar evaluaciones diagnósticas. En el mismo sentido se inscribe el análisis de los resultados del Operativo “Aprender” del año 2016, publicado por la Secretaría de Evaluación Educativa del Ministerio de Educación y Deportes en el año 2017.

Si bien en los informes mencionados aparece un número significativo de niños que presenta dificultades en la comprensión de textos, en ninguno de estos documentos se especifican las dificultades inherentes a los primeros aprendizajes relativos a la decodificación, esto es, a la precisión y la velocidad en el reconocimiento de palabras escritas. La falta de información acerca del desempeño en decodificación resulta por lo menos preocupante teniendo en cuenta que una comprensión lectora eficaz depende tanto del dominio en las habilidades de decodificación, específicas

para la lectura, así como de habilidades más generales de comprensión del lenguaje, elaboración de inferencias, memoria de trabajo y funcionamiento ejecutivo.

Al respecto, Fletcher, Lyon, Fuchs y Barnes (2007) y Stanovich (1993), entre otros, señalaron que cuando la decodificación no se realiza de manera precisa y rápida, demanda recursos cognitivos que no pueden ser asignados a procesos de alto nivel para la integración y la comprensión de textos.

Por otro lado, a partir del Operativo Nacional de Evaluación del año 2005, la Argentina adoptó un enfoque de evaluación basado en pruebas referidas a criterios. Tales pruebas permiten estimar los conocimientos y habilidades adquiridos por los alumnos en las distintas áreas curriculares. Aportan información que puede ser utilizada para la identificación de dificultades en el aprendizaje, la implementación de estrategias de intervención, y el monitoreo de los progresos. Sin embargo, resulta necesario que sean complementadas con evaluaciones referidas a normas que permitan comparar a los sujetos con su grupo de referencia.

En función de lo expuesto, resulta insuficiente la información disponible sobre las posibilidades de evaluación destinada a la detección temprana de dificultades en el desempeño en lectura, así como de su secuencia ulterior.

Justificación del tema

La justificación del tema propuesto reside en su *relevancia y viabilidad*. La relevancia surge, en primer lugar, de la *importancia* de contar con instrumentos válidos y confiables, con datos normativos adecuados que posibiliten la identificación del desempeño en la lectura, teniendo en cuenta la evidencia empírica sobre un porcentaje significativo de la población infantil que no alcanza el nivel esperado. En segundo lugar, de la *novedad* que supone contar con los requisitos psicométricos del test seleccionado, para un adecuado uso de sus resultados.

Por lo tanto, surge particularmente la necesidad de disponer de instrumentos que permitan detectar de manera oportuna a aquellos niños que presenten riesgo de desarrollar dificultades en el aprendizaje de la lectura, que puedan ser utilizados por los docentes, que resulten rápidos y económicos, y para cuya aplicación e interpretación no se requiera formación altamente especializada.

Teniendo en cuenta las ventajas de adoptar los modelos teóricos más actuales, en particular el modelo RTI para la identificación de las dificultades en el aprendizaje y la existencia de un test de *screening* de eficacia lectora adaptado lingüísticamente al medio local, el TECLE Buenos Aires, aparece con nitidez la viabilidad e interés de los estudios dirigidos a su tipificación. La misma está destinada a aportar evidencias empíricas de su *validez* y *confiabilidad*, que permitan una interpretación adecuada de las puntuaciones que arroja, para lo que es imprescindible contar con datos normativos. Al respecto, la originalidad de la tesis reside, además, en la elaboración de los *baremos locales generales según el grado escolar y baremos segmentados según el grado escolar y el nivel de oportunidades educativas*, así como la especificación de los criterios a sustentar para el uso e interpretación de los puntajes obtenidos en uno y otro caso.

En síntesis, contar con un instrumento que posibilite la estimación rápida de los niveles de aprendizaje logrados en la lectura en relación con la identificación de los estándares a alcanzar en el ámbito educativo local, de modo tal de contar con indicadores para arribar a la detección oportuna de eventuales dificultades, parece promisorio para coadyuvar al diseño de estrategias de intervención.

En función de ello, en esta tesis, con un diseño instrumental que corresponde a la elaboración/adaptación de un instrumento de medición en Psicología (Ato, López, & Benavente, 2013), se realizaron tres estudios: de confiabilidad, de validez y de elaboración de datos normativos.

El estudio de confiabilidad por el método test-retest, incluyó a 218 niños de escolaridad primaria de la región noroeste del Gran Buenos Aires, que arrojó como resultado un alto grado de estabilidad temporal en los puntajes observados en ambas instancias de evaluación realizadas con un intervalo de 5 semanas, con un coeficiente de estabilidad $r = .90$.

El segundo estudio estuvo dirigido a establecer la validez de la prueba, a través de evidencias de constructo y de criterio. En el caso de la *validez de constructo* se examinaron 57 sujetos de escolaridad primaria. Se analizaron las asociaciones convergentes y discriminantes entre el rendimiento en el TECLE Buenos Aires y otras medidas teóricamente relacionadas: precisión y tiempo de ejecución en una prueba de *screening* del Test Leer para Comprender; precisión y tiempo en la lectura en voz alta de palabras y de pseudopalabras y rendimiento en la prueba de comprensión de palabras y frases del test LEE; rendimiento en la prueba de comprensión de vocabulario de la escala Alfa-2; tiempo en la denominación automatizada rápida de dibujos, y tiempos de ejecución en las partes 1 y 2 de una prueba de trazado de senderos. Los resultados mostraron que todas las correlaciones alcanzaron significación estadística. Para la *validez de criterio* se estudiaron las diferencias entre los puntajes observados en el test en dos grupos contrastados de 28 participantes cada uno, apareados por edad, sexo, grado escolar y nivel de oportunidades educativas: un grupo clínico conformado por sujetos con diagnóstico de trastornos en el aprendizaje de la lectura, de la escritura, de las matemáticas y/o trastornos del lenguaje (variable criterio), y un grupo control que incluyó a sujetos sin estas dificultades. La prueba *t* para muestras independientes reveló una diferencia estadísticamente significativa entre las medias de los puntajes de ambos grupos, [$t(54) = -2.07, p = .044$], con un tamaño del efecto mediano ($d = .55$) (Cohen, 1988; Ledesma, Macbeth, & Cortada de Kohan, 2008). Ambos estudios han permitido establecer la validez de la prueba, que se especifican con mayor detalle en los capítulos respectivos.

Finalmente, el tercer estudio estuvo referido a la obtención de los datos normativos de la prueba. Se incluyeron, en una muestra por conveniencia, 3097 participantes de 2° a 6° grados de educación primaria, que asistían a 24 escuelas de la Ciudad Autónoma de Buenos Aires y alrededores. Además de la obtención de datos para la muestra total, se analizaron los efectos sobre el rendimiento en el test de las variables “grado escolar”, “edad en años” y “sexo” de los niños y “nivel de oportunidades educativas”, “tipo de gestión” y “región geográfica” de las escuelas a las que asistían. Ello permitió confeccionar los baremos generales para la población según grado escolar y baremos particularizados según grado escolar y nivel de oportunidades educativas, en función de que resultaron las variables en las que aparecen las mayores diferencias entre los grupos. En ambos casos se ofrecen indicaciones sobre la pertinencia y utilidad del uso de cada uno de ellos.

En síntesis, esta tesis aporta evidencias sobre la confiabilidad y validez del instrumento, adicionando información empírica acerca del alcance del constructo “eficacia lectora”, y proporciona baremos generales y diferenciados por nivel de oportunidades educativas. Con las moderaciones del caso, puede sostenerse que constituye una herramienta para la detección del riesgo de desarrollar o presentar dificultades en el aprendizaje de la lectura, así como evaluar la efectividad de las prácticas de enseñanza de la lectura.

Organización de la tesis

El capítulo 1 está dedicado al marco teórico y los antecedentes de la identificación de las dificultades en el aprendizaje de la lectura. Comienza con una breve descripción de los distintos tipos de sistemas de escritura, seguida de una presentación del clásico modelo de doble ruta para el reconocimiento de las palabras escritas. A continuación se desarrollan las etapas en el aprendizaje de la lectura en los sistemas alfabéticos, se definen las dificultades específicas en el aprendizaje de la lectura, que incluyen las particularidades de la dislexia en español, y se comparan dos enfoques para

su identificación. El capítulo finaliza con una consideración de la importancia de la detección temprana del riesgo de desarrollar estas dificultades en el marco de los modelos RTI.

En el capítulo 2, en la primera parte, se describe la estructura del test original (TECLE) así como sus características en cuanto a consignas, modalidades de administración y tiempo límite. Seguidamente, se propone una interpretación de la medida que arroja en términos del constructo “eficacia lectora”. Finalmente se describen los cambios realizados en la adaptación utilizada para la presente tesis, es decir el TECLE Buenos Aires. En la segunda, se presenta una descripción detallada de los interrogantes que orientaron esta investigación, de los objetivos e hipótesis formulados y de la metodología general utilizada. Se señala que se especificarán en cada uno de los capítulos sobre los 3 estudios realizados los participantes incluidos, las herramientas administradas y los procedimientos para la obtención y el análisis de los datos.

Los capítulos 3, 4 y 5 están dedicados a los estudios de confiabilidad, de validez de constructo y de criterio, y de obtención y justificación estadística de los datos normativos del test para la Ciudad Autónoma de Buenos Aires y alrededores, respectivamente. Los tres capítulos comparten la siguiente estructura: en primer lugar se introducen los conceptos pertinentes de la teoría psicométrica, a continuación, se amplía la metodología desarrollada en el capítulo 2, luego se reportan los resultados obtenidos, y finalmente, se presenta la discusión y las conclusiones.

En el capítulo 6 se resumen los resultados de los tres estudios realizados, se discute la interpretación de los puntajes en términos de “eficacia lectora” y los usos propuestos del test en el marco de un enfoque de la validez de las consecuencias de dichos usos y finalmente se exponen las conclusiones de la tesis.

Aparecen a continuación las referencias y, por último, se incluyen cuatro anexos. En el anexo A se presentan los resultados de análisis

estadísticos no paramétricos realizados como procedimiento alternativo a algunos de los análisis paramétricos en los cuales la distribución de los datos no cumplía con los requisitos de normalidad y/u homocedasticidad. El anexo B exhibe las tablas con dos tipos de baremos para la Ciudad Autónoma de Buenos Aires y alrededores: un baremo general según el grado escolar, y otro segmentado según el grado escolar y el nivel de oportunidades educativas. En el anexo C aparece la versión del test de *screening* utilizada en esta investigación, seguida de las normas de aplicación y puntuación y de un protocolo de puntuación sugerido para los usuarios. Finalmente, en el anexo D se presentan tres instrumentos utilizados en el estudio de validez de constructo.

Capítulo 1

Marco teórico y antecedentes

Cuando los lectores expertos se abocan a la tarea de comprender un texto, ponen en juego dos tipos de habilidades diferentes: las que resultan en la decodificación y el reconocimiento de las palabras escritas, y las que permiten extraer el significado del material leído. Más allá de la importancia que revisten ambas, sólo las primeras son específicas para la lectura, ya que dependen de los procesos por los cuales se accede a la pronunciación y al significado de las palabras a partir de su ortografía. Las segundas, en cambio, incluyen habilidades lingüísticas y cognitivas generales: riqueza del vocabulario, procesamiento morfológico, comprensión sintáctica, capacidad de almacenamiento en la memoria de trabajo, conocimiento general del mundo, habilidad para reponer información no explicitada en el texto (elaboración de inferencias), integración del significado de distintas partes del mismo, comprensión de los objetivos y demandas de la tarea, sensibilidad al tipo de texto, habilidades metacognitivas de monitoreo y regulación de la propia comprensión, y disponibilidad de estrategias para mejorar la comprensión (Defior Citoler, 1996; Fletcher, Lyon, Fuchs, & Barnes, 2007; Perfetti, 1985, citado en Perfetti, 2007; Scarborough, 2001). De acuerdo con Hoover y Gough (1990), cada uno de ambos componentes es necesario, pero no suficiente para una comprensión lectora eficaz. Así, por ejemplo, no será posible la comprensión del texto si el lector no logra evocar las palabras a partir de sus ortografías, o si no es capaz de comprender las palabras e ideas expuestas en el texto aun cuando éstas le son presentadas en forma oral.

Uno de los principales objetivos de la enseñanza en los primeros años de la educación primaria es que los niños adquieran el dominio de las habilidades de decodificación, en términos de precisión y fluidez en el reconocimiento de las palabras escritas así como de la recuperación de los aspectos prosódicos (ausentes en la escritura) en la lectura en voz alta. La automatización de la decodificación como consecuencia de la práctica

frecuente libera recursos atencionales para los procesos de alto nivel que demanda la comprensión de textos. En los años posteriores, el dominio de las habilidades de comprensión lectora constituye la principal herramienta para la continuidad exitosa de la escolarización y de la formación académica, así como la base del crecimiento del conocimiento del mundo.

No caben dudas, entonces, de la importancia de conocer las fases de la adquisición de la lectura, las manifestaciones de las dificultades en dicho aprendizaje, sus posibles causas, los métodos para su identificación temprana, así como las estrategias para su remediación.

En el primer apartado del presente capítulo, se presenta una breve caracterización de los distintos tipos de sistemas de escritura seguida de la descripción del modelo de doble ruta para el reconocimiento de las palabras escritas.

El segundo apartado se centra en las fases que atraviesan los niños en el aprendizaje de la lectura en los sistemas alfabéticos. Luego se comparan el método global y el método fónico de enseñanza de la lectura.

En el tercero, se enuncian las formas en que puede verse afectado el aprendizaje de la lectura, así como los factores que contribuyen a la aparición de las mismas. A continuación, se detallan las manifestaciones conductuales de las alteraciones en el reconocimiento de las palabras escritas, se define a la dislexia, se caracteriza su manifestación en español y se expone la hipótesis más ampliamente aceptada acerca del déficit cognitivo subyacente.

Por último, se abordan dos métodos de identificación de las dificultades específicas en el aprendizaje: el método tradicional basado en la discrepancia entre la capacidad intelectual general y/o la edad y el rendimiento académico, y el método adoptado por los modelos RTI.

1.1. Los sistemas de escritura

Los sistemas de escritura pueden ser clasificados, en primer lugar, en términos del tipo de unidad lingüística que los caracteres representan (Sampson, 1997). Los sistemas logográficos, como el chino mandarín, representan morfemas, las unidades mínimas portadoras de significado. En cambio, los sistemas fonográficos representan unidades fonológicas, sean sílabas (como el kana japonés), componentes silábicos (como el zhuyin, el sistema de notación fonética utilizado en Taiwán), o fonemas (como el español). Los sistemas que representan fonemas reciben el nombre de sistemas alfabéticos.

El tipo de unidad lingüística representada, junto con la estructura fonológica de la lengua, determina la extensión del conjunto de símbolos gráficos que componen un sistema de escritura. Así, por ejemplo, se espera que al finalizar la escuela primaria, los niños chinos logren reconocer unos 2.400 caracteres, que representan el 99% de los símbolos gráficos que aparecen en la mayoría de los textos (Mair, 1996, citado en Hanley, 2005). En cambio, el silabario kana japonés consta de las 116 sílabas permitidas en esa lengua, cuya estructura silábica puede ser V (vocal) o CV (consonante-vocal). Por último, el español se compone de 29 grafemas según Frost (2005), pero la Real Academia Española y la Asociación de Academias de la Lengua Española (2010) estipulan que se distinguen 27 letras y 5 dígrafos o combinaciones de dos letras (ch, ll, gu, qu y rr).

La comparación de los distintos sistemas de escritura muestra que cuanto menor es el tamaño de las unidades lingüísticas representadas, es más económico en términos de los requerimientos de memoria para el almacenamiento de los símbolos gráficos. Adicionalmente, los alfabéticos son los más productivos o generativos porque la combinación de un número limitado de caracteres permite representar ortográficamente todas las palabras de la lengua. Para ello, estos sistemas incluyen un conjunto de reglas que dan cuenta de las relaciones entre las unidades ortográficas y las fonológicas.

Una distinción importante entre los sistemas alfabéticos se refiere al grado de correspondencia entre los grafemas y fonemas, que se expresa en el conjunto de reglas de conversión entre ambos tipos de unidades (en adelante, RCGF). En los *sistemas transparentes* para la lectura, la ortografía de las palabras se caracteriza por ser regular o consistente, como en el caso del español, el italiano, el alemán y el serbocroata. Es decir que en estos sistemas la pronunciación de las palabras escritas se deriva directamente de la aplicación de las RCGF que en su mayoría son simples, de tipo uno-a-uno, por lo cual a un grafema le corresponde una única pronunciación. El resto de las reglas especifica el contexto ortográfico en el deben aplicarse las pocas pronunciaciones alternativas para un grafema; por ejemplo, en español rioplatense, el grafema “C” se pronuncia /k/ a menos que preceda a las vocales “E” o “I”, en cuyo caso se pronuncia /s/. Asimismo, en este sistema, el patrón acentual de las palabras escritas está totalmente especificado por reglas.

En cambio, en los *sistemas opacos* como el inglés, el alto grado de irregularidad o inconsistencia de la ortografía de las palabras da como resultado que su pronunciación no sea fácilmente predecible a partir de la aplicación de las RCGF, dada la complejidad de sus relaciones. En este sistema existen numerosos casos en los que una letra o secuencia de letras tiene más de una pronunciación; por ejemplo, la secuencia “OUGH” se pronuncia /ʌf/ en “TOUGH” (duro, fuerte, difícil), /əʊ/ en “THOUGH” (aunque) y /u:/ en “THROUGH” (a través de). El inglés incluye, además, una cantidad no menor de palabras de ortografía excepcional cuya pronunciación no surge de la aplicación de las RCGF; por ejemplo, las palabras “WEDNESDAY” (miércoles), “CHOIR” (coro) y “GAUGE” (calibre) se pronuncian, respectivamente /'wɛnzdeɪ/, /kwaɪə/ y /geɪdʒ/.

En la escritura, en cambio, el español tanto peninsular como rioplatense, no alcanza el grado de transparencia que caracteriza a la lectura, ya que existe un número de “sonidos” (fonemas) que poseen más de una realización ortográfica. Por ejemplo, según la palabra y la vocal con la

que forma sílaba, el sonido /s/ del español rioplatense se escribe con el grafema “S”, “Z”, o “C”, el sonido /x/ se escribe con “G” o “J”, el sonido /b/ se escribe con “B” o “V”; asimismo las vocales pueden estar precedidas por “H”, que carece de realización fonética. Además, existen numerosas palabras homófonas, es decir, palabras con ortografía diferente que comparten, sin embargo, una misma realización fonética, como “ASIA” y “HACIA”, “CIEN” y “SIEN”, “BOTAR” y “VOTAR”.

El inglés como el francés son sistemas más opacos para la escritura que el español. Por ejemplo, en inglés, la realización fonética /pi:s/ puede referir a las palabras “PEACE” (paz) y “PIECE” (trozo); y en francés, la expresión /mɛtʁ/, a las palabras “MAÎTRE” (maestro), “MÈTRE” (metro) y “METTRE” (poner).

En el caso del inglés, la opacidad de su sistema de escritura obedece a que la lengua oral sufrió grandes cambios a lo largo de mil años de historia, que no fueron acompañados por reformas sistemáticas en su sistema de escritura (Venezky, 1999). Tales transformaciones son solidarias de las influencias de otras lenguas, como el anglosajón antiguo proveniente del germano de base, del francés, del latín y del griego. En este sistema, un grafema –unidad de escritura que corresponde a un fonema– puede tener una longitud de una a cuatro letras. De ahí, que algunos autores consideren que se trata de “un sistema morfofonográfico” (Moats, 2010).

1.1.1. El modelo de doble ruta para el reconocimiento de las palabras escritas

Los lectores expertos disponen de dos procedimientos para el reconocimiento de las palabras escritas, involucrados tanto en las tareas de lectura en voz alta como en las de comprensión (Coltheart, 2005; Coltheart, Rastle, Perry, Langdon, & Ziegler, 2001). La *ruta léxica* o *directa* permite el reconocimiento de una palabra escrita conocida a través de la conexión entre tres tipos de información almacenada en el léxico: la representación ortográfica de la palabra y sus representaciones fonológica y semántica. En cambio, la decodificación de las palabras cuya ortografía no está

almacenada en el léxico requiere de la mediación fonológica. Esta *ruta subléxica* o *fonológica* involucra, en primer lugar, la segmentación de la secuencia de letras, seguida por la aplicación de RCGF, y finalmente, el ensamblaje de la cadena de fonemas en una secuencia de unidades silábicas. El procesamiento en esta ruta está influido por la longitud del estímulo ya que el ensamblaje de secuencias más largas, en relación con el de secuencias más cortas, impone mayores demandas al sostenimiento de información segmental o fonémica en la memoria de corto término fonológica.

En los lectores expertos, ambas rutas son funcionales y operan en paralelo. Sin embargo, la importancia relativa de cada una varía en función de las características del sistema de escritura, de las del estímulo, y de las demandas de la tarea.

La mayoría de los modelos de doble ruta fueron elaborados a partir de estudios realizados en inglés, cuyo sistema de escritura es considerado el más opaco entre los sistemas alfabéticos europeos, por lo que cabe preguntarse si resulta válida la generalización a otros. Sin embargo, pese a las diferencias descritas más arriba, todos los sistemas de escritura comparten características que habilitan cierto grado de generalización. En primer lugar, los símbolos gráficos o caracteres de todos ellos representan unidades del lenguaje oral, sean morfemas, sílabas, o fonemas. En segundo lugar, aún los sistemas de escritura logográficos, compuestos por caracteres que representan unidades léxico-semánticas, incluyen símbolos con información fonética (como el chino moderno), o bien se utilizan conjuntamente con elementos de otros sistemas de escritura que representan unidades fonológicas. Por ejemplo, en gran parte de China, el sistema de escritura logográfico coexiste con otro sistema, el pinyin, basado en el alfabeto latino; en Corea y Japón, se utilizan dos sistemas que toman los logogramas chinos, respectivamente, el hanja y el kanji, junto con el alfabeto hangul coreano y el silabario kana japonés (Frost, 2005; Hanley, 2005).

En cuanto al rol de las características del estímulo en la importancia relativa de las dos rutas de lectura en un sistema alfabético, las palabras excepcionales y las de ortografía irregular o inconsistente, por un lado, sólo pueden ser procesadas por la ruta léxica, que además sería la preferida para el procesamiento de las palabras de alta frecuencia, sean de ortografía irregular o regular. Por otro lado, las pseudopalabras, así como las palabras de ortografía regular de muy baja frecuencia, sólo pueden ser procesadas por la ruta fonológica ya que no cuentan o pueden no contar con una representación ortográfica almacenada en el léxico. En cambio, el reconocimiento de las palabras de ortografía irregular menos frecuentes depende de la contribución de ambas rutas. Esto último ha sido expresado a través de la metáfora de “la carrera de caballos” para hipotetizar que ambas vías operan en paralelo y que la respuesta en una tarea –la lectura en voz alta, por ejemplo– es el resultado de la actividad en la vía que logre completar antes el procesamiento. El rendimiento diferencial en función de las características de los estímulos ha sido interpretado como una manifestación de la actividad en una u otra vía. Así, con respecto a la ruta léxica, se considera que el mejor rendimiento con palabras que con pseudopalabras (efecto de lexicalidad) pone en evidencia que las primeras cuentan con representaciones almacenadas en el léxico, y que el mejor rendimiento con palabras más frecuentes que menos frecuentes (efecto de frecuencia) es una manifestación del rol que juega la frecuencia de uso en la cantidad de activación que requiere una representación léxica para que esté disponible para dar una respuesta. Por otro lado, en cuanto a la ruta fonológica, el mejor rendimiento con palabras de ortografía regular que irregular (efecto de regularidad ortográfica) refleja el uso de las RCGF para la ejecución de la tarea, así como el mejor rendimiento con estímulos cortos que largos (efecto de longitud) es una manifestación de las menores demandas de sostenimiento en la memoria fonológica para los primeros que para los segundos.

Sin embargo, Coltheart (2005) ha señalado que los sujetos normales tardan más tiempo en leer las palabras irregulares que las regulares, aunque

no suelen producir errores de regularización. Esto es, si el rendimiento en la lectura de las palabras irregulares fuera el resultado del procesamiento en la ruta léxica más rápida, no deberían observarse tiempos más prolongados que en la lectura de palabras regulares. Por otro lado, si fuera el resultado del procesamiento en la ruta fonológica debería dar lugar a errores de regularización. En razón de ello, Coltheart ha afirmado que la metáfora de “las dos mangueras y el balde” de Baron (1977, citado en Coltheart, 2005) es una expresión más adecuada para el rendimiento de los sujetos normales en la lectura en voz alta de palabras irregulares. De acuerdo con esta metáfora, la contribución de ambas rutas al reconocimiento de las palabras se compara con la cantidad de agua que aportan dos mangueras para llenar un balde: el balde tarda menos tiempo en llenarse cuando recibe agua de ambas mangueras, independientemente de la cantidad aportada por cada una.

Por último, existen evidencias contradictorias para el rol de la mediación fonológica en tareas de comprensión en lectores expertos, que parecen estar influidas por la experiencia lectora de los sujetos, las características de los estímulos y las demandas particulares de las tareas (Van Orden & Kloos, 2005).

1.2. El aprendizaje de la lectura en los sistemas alfabéticos

Las distintas teorías del aprendizaje de la lectura, entendiendo por lectura a las habilidades específicas de decodificación y reconocimiento de las palabras escritas, describen la manera en que se desarrollan los procedimientos o estrategias por los cuales los lectores expertos leen las palabras.

En una de las teorías más influyentes acerca de la adquisición de la lectura, Uta Frith (1985) propuso la existencia de tres fases –logográfica, alfabética y ortográfica– que se caracterizan por el tipo de estrategia adoptada para la lectura de las palabras.

La *fase logográfica* comienza alrededor de los 4 o 5 años de edad, previamente al ingreso a la escuela primaria. Los niños de esta edad pueden reconocer un número reducido de palabras, que incluyen su propio nombre y nombres de marcas comerciales frecuentes, a partir de claves visuales salientes de la tipografía o contextuales. Si bien este procedimiento les permite reconocer algunas palabras, no constituye una estrategia de lectura en sentido estricto ya que, en lo concerniente a los sistemas de escritura alfabéticos, no involucra la conexión entre las letras o grafemas y lo que éstos representan, esto es, los sonidos de habla o fonemas.

La noción de que los grafemas guardan relaciones sistemáticas con los fonemas es conocida como el *principio alfabético* y su comprensión requiere el dominio de las habilidades de conciencia fonémica. La *conciencia fonémica* se define como la comprensión y la reflexión consciente acerca de que el habla está compuesta por sonidos discretos, los fonemas, y se manifiesta a través de la habilidad para manipularlos, como en las tareas de conteo, segmentación, adición, sustracción, sustitución y desplazamiento de fonemas. Esta habilidad para la identificación de los sonidos en el habla no se desarrolla de manera espontánea ya que la unidad de segmentación del habla es la sílaba, que coincide con un pulso de sonido (Mehler, Dommergues, Frauenfelder, & Seguí, 1981; Torgesen, & Mathes, 2000).

Además, en el habla se produce el fenómeno de la coarticulación por el cual los patrones articulatorios de fonemas adyacentes se influyen mutuamente asimilándose unos a otros, lo que resulta en una falta de invariancia acústica del fonema en los distintos contextos fonéticos en los que aparece. Es por ello que los niños difícilmente logran captar espontáneamente la esencia del fonema en la percepción del habla (Lieberman & Shankweiler, 1991, citado en Fletcher, Lyon, Fuchs, & Barnes, 2007). En consecuencia, la enseñanza y el entrenamiento de las habilidades de conciencia fonémica deben ser implementados de manera explícita y sistemática y vinculados a material escrito con el fin de habilitar la adquisición del principio alfabético. A su vez, y recíprocamente, la

adquisición de este principio redundando en un refinamiento de las habilidades de conciencia fonémica (Morais, Bertelson, Cary, & Alegría, 1986; Snow, & Juel, 2005).

El ingreso a la escuela y el comienzo de la enseñanza formal de la lectoescritura dan inicio a la *fase alfabética*, en la que los niños empiezan a adquirir las RCGF. En esta fase, la decodificación de las palabras se logra a partir de la aplicación de estas reglas y del ensamblaje de los fonemas resultantes en sílabas. A diferencia de la estrategia adoptada en la fase logográfica, la estrategia alfabética se caracteriza por ser analítica, fonológicamente mediada y altamente generativa. Le permite al niño decodificar todas las palabras de ortografía regular, aún la primera vez que se le presentan. A su vez, funciona como un mecanismo de auto-enseñanza ya que la decodificación exitosa de secuencias nuevas de letras resulta en el establecimiento de una conexión entre la pronunciación y el significado de las palabras y su ortografía, y en el almacenamiento de las restricciones ortográficas del sistema de escritura que se está adquiriendo. Consecuentemente, empieza a constituirse el léxico ortográfico (Share, 1995).

En la *fase ortográfica*, el niño es capaz de utilizar las representaciones almacenadas en su léxico ortográfico para recuperar directamente la pronunciación y el significado de las palabras, sin recurrir a la aplicación de las RCGF. Según Ehri (1999) los reiterados eventos de recodificación exitosa conducen a una suerte de amalgama que incluye la forma escrita, la forma fonológica y el contenido semántico de las palabras conocidas. No obstante, la estrategia fonológica de lectura se mantiene disponible en los lectores expertos y opera en paralelo con la estrategia ortográfica. De hecho, la estrategia fonológica es el único procedimiento que permite la lectura de secuencias ortográficas novedosas.

La estrategia ortográfica facilita la automatización del reconocimiento de la palabra. Conjuntamente con factores como la denominación automatizada rápida (que refleja la velocidad en la recuperación de una

etiqueta léxica), la regulación de la atención y la inhibición de respuestas no pertinentes, contribuye a la fluidez en la lectura de palabras, oraciones y texto, que se manifiesta a través de la precisión, la velocidad y la prosodia adecuadas en la lectura en voz alta de textos.

Por otra parte, de acuerdo con la idea de la “división del trabajo” planteada por Byrne (2005), todo aprendizaje es el producto de las contribuciones del aprendiz y del ambiente, y los distintos tipos de aprendizaje pueden ubicarse a lo largo de un *continuum* en función de la influencia relativa de uno y otro factor. Así, por ejemplo, en la adquisición del lenguaje oral, es el aprendiz quien realiza la contribución más significativa ya que la mayoría de los niños lo adquiere de manera espontánea, sin esfuerzo consciente y sin instrucción formal, mientras que en el otro extremo, el dominio de las nociones y las operaciones matemáticas requiere de varios años de enseñanza explícita y sistemática. De acuerdo con esta idea, las prácticas pedagógicas deben suministrar y entrenar oportunamente aquellas nociones y habilidades que el aprendiz no adquirirá sin instrucción formal.

En el caso particular del aprendizaje de la lectura, como se mencionó previamente en este apartado, los niños difícilmente logran dominar espontáneamente las habilidades de conciencia fonémica necesarias para comprender el principio alfabético. De modo que también es fundamental considerar los factores ambientales que incluyen el grado de transparencia/opacidad del sistema alfabético, la complejidad silábica de la lengua oral y los hábitos de lectura en el entorno de crianza del niño. Por otro lado, la contribución del niño está dada por su dominio de las habilidades necesarias para dicho aprendizaje.

Con respecto a las características del sistema alfabético, los niños que adquieren un sistema transparente con un predominio de sílabas que poseen una estructura silábica simple, como el español, el italiano, el griego y el finés, logran dominar más rápidamente la decodificación que los niños que adquieren un sistema opaco con un predominio de sílabas con estructura silábica compleja, como el inglés y el danés. Esto se debe, por un

lado, a que en los sistemas transparentes la correspondencia entre los grafemas y los fonemas es más confiable, mientras que en los sistemas opacos, en los que dichas correspondencias son menos confiables, el procedimiento de conversión también debe aplicarse de manera flexible a unidades de mayor tamaño: componentes silábicos, morfemas y palabras. Por otro lado, la mayor complejidad en la estructura silábica exige mayores habilidades de conciencia fonémica ya que es necesario segmentar los sonidos que componen los ataques y codas silábicos complejos a los fines de aplicar las RCGF. Seymour, Aro y Erskine (2003) compararon el rendimiento de niños de primer grado hablantes de inglés y de otras 12 lenguas europeas en tareas de conocimiento de letras, lectura en voz alta de palabras familiares y lectura en voz alta de no palabras simples; además incluyeron en la comparación, el rendimiento de niños de segundo grado hablantes de inglés, francés y danés, cuyos sistemas ortográficos son opacos. Los autores mostraron que en los sistemas transparentes, los niños lograron dominar la decodificación antes de la finalización del primer grado escolar, mientras que los niños que estaban adquiriendo el sistema del inglés necesitaron más de un año adicional de instrucción para alcanzar niveles comparables de precisión y fluidez. Adicionalmente, la complejidad de la estructura silábica también influyó en la eficacia de la lectura en voz alta de no palabras simples, lo cual fue interpretado como el resultado de la dificultad en adquirir las correspondencias de un grafema con un fonema en distintos contextos, simples y complejos (Seymour, Aro, & Erskine, 2003; Ziegler & Goswami, 2005).

En cuanto a las habilidades necesarias para el aprendizaje de la lectura, las contribuciones relativas de las habilidades de conciencia fonológica, de denominación automatizada rápida y de procesamiento ortográfico (definida en términos de la sensibilidad a la estructura ortográfica de las palabras), predictoras de la precisión y la fluidez en lectura, difieren según el grado de transparencia/opacidad del sistema alfabético. Por ejemplo, Georgiou, Parrila y Papadopoulos (2008) compararon la contribución relativa de dichas habilidades en dos grupos de niños que

estaban adquiriendo el sistema alfabético griego (transparente) y el inglés (opaco). Los autores observaron las siguientes relaciones concurrentes entre dichas habilidades y el rendimiento en precisión de la decodificación en primer grado: (a) las habilidades de conciencia fonológica predijeron mejor el rendimiento en inglés que en griego, (b) las habilidades de denominación automatizada rápida predijeron el rendimiento sólo en griego, y (c) las habilidades de procesamiento ortográfico predijeron el rendimiento sólo en inglés. En cuanto al rendimiento en la fluidez lectora, tanto en griego como en inglés, las variables predictoras concurrentes fueron la denominación automatizada rápida y el procesamiento ortográfico. Cuando estudiaron las relaciones longitudinalmente, encontraron que en ambos sistemas alfabéticos, el rendimiento en la precisión de la decodificación y en la fluidez lectora observados en primer grado fueron las variables predictoras más importantes del rendimiento de las mismas en segundo grado (ver también Defior & Serrano, 2011).

1.2.1. Los métodos de enseñanza de la lectura en los sistemas alfabéticos: el método global versus el método fónico

En el marco de la idea de la “división del trabajo” de Byrne (2005) y en virtud de lo expuesto en el apartado anterior, resulta pertinente describir brevemente los dos métodos de enseñanza de la lectura que son objeto de debate desde hace más de un siglo, a fin de considerar de qué manera contribuyen a dicho aprendizaje.

Los defensores del *método global*, por un lado, sostienen una visión holística del reconocimiento de las palabras escritas, evidenciada a través de lo que Cattell llamó “efecto de superioridad de la palabra” – el mejor rendimiento que exhiben los lectores expertos en el reconocimiento de palabras presentadas brevemente que en el reconocimiento de letras aisladas (Cattell, 1886, citado en Snow & Juel, 2005). Además, asumen que los niños poseen la capacidad para deducir de manera autónoma las relaciones entre las unidades del lenguaje escrito y las del lenguaje oral. En consecuencia, promueven en los niños la práctica de predecir o “adivinar” las

palabras basándose en el significado del texto. Sin embargo, estudios realizados en adultos demuestran que los sujetos predicen, en promedio, una de cuatro palabras funcionales de un texto y sólo una de diez palabras de contenido, y que predecir las palabras lleva más tiempo que leerlas (Gough, Alford, & Holley-Wilcox, 1981; Gough, 1983, citados en Snow & Juel, 2005).

Por el contrario, los defensores del *método fónico* señalan que lo que diferencia a los buenos y malos lectores no es la mayor habilidad para predecir las palabras a partir del contexto, sino el reconocimiento rápido y eficaz de las mismas. Asimismo, reconocen la dificultad que impone el fenómeno de la coarticulación a la habilidad para segmentar el habla en las unidades representadas por los símbolos gráficos. Por lo tanto, prescriben la enseñanza explícita y sistemática de las habilidades de conciencia fonémica y de las RCGF (Snow & Juel, 2005). En esta línea, cabe mencionar las recomendaciones expresadas en un reporte del Panel Nacional de Lectura (*NRP, National Reading Panel, 2000*, citado en Snow & Juel, 2005) para la enseñanza de la lectura en inglés, a partir de la revisión de la literatura acerca de aquellas prácticas pedagógicas más efectivas. Las recomendaciones incluyeron atender a las áreas de conciencia fonológica, relaciones sonidos-letras, fluidez, vocabulario, y estrategias de comprensión. Las recomendaciones para las habilidades de precisión y fluidez en la decodificación incluyeron la implementación del entrenamiento en conciencia fonológica antes de segundo grado, durante un mínimo de 5 a 20 horas en total, vinculado con la enseñanza sistemática de las relaciones sonido-letra, y la lectura en voz alta guiada.

1.3. Las dificultades en el aprendizaje de la lectura

El 6 de abril de 1963, en una conferencia organizada por el Fondo para los Niños con Discapacidad Perceptual (*Fund for Perceptually Handicapped Children, Inc.*), Samuel Kirk señaló para referirse a la variada nomenclatura destinada a caracterizar diferentes problemas escolares “que gran parte de la terminología utilizada hasta la fecha era inútil desde el punto de vista del

manejo educacional” (Kirk & Bateman, 1962). Fue quien primero habló de *trastornos de aprendizaje*, para describir niños con alteraciones en el lenguaje, el habla, la lectura o problemas agregados de comunicación, sin deficiencias sensoriales (ceguera o sordera) y sin retraso mental generalizado. En gran parte de la literatura específica posterior acerca de las dificultades en el aprendizaje de la lectura, se identificó al déficit específico con el término “dislexia”.

Sin embargo, este término había sido acuñado por Berlin en el año 1887 (Berlin, 1887, citado en Elliott & Grigorenko, 2014) para describir una dificultad para decodificar los símbolos escritos a causa de una lesión cerebral, en sujetos adultos. Por otro lado, el autor utilizó el término “alexia” para referir a una total incapacidad para leer. Posteriormente se distinguió entre la condición adquirida y la congénita del trastorno y se utilizaron otros términos tales como *ceguera verbal* (Kussmaul, 1877, citado en Elliott & Grigorenko, 2014) para referir a la primera y *ceguera verbal congénita* (Hinshelwood, 1900, 1917, citado en Vellutino & Fletcher, 2005) y *estrefosimbolia* (Orton, 1925, citado en Vellutino & Fletcher, 2005) para la segunda.

Progresivamente fue ganando atención el primitivo término dislexia, aunque su delimitación no fue precisa, en consonancia con los conocimientos de los que se disponía hasta hace algunas décadas.

Desde entonces un extenso recorrido ha marcado diferentes perspectivas sobre tales trastornos, que han incidido en su definición y consecuentemente en su caracterización, con modificaciones importantes en la medida en la que se fueron desarrollando modelos robustos de lectura que guiaron la investigación.

Aproximadamente hasta la última década, la identificación del trastorno aparecía de acuerdo a los criterios enunciados en el DSM-IV, cuya versión en español fue publicada en el año 1995 (Asociación Americana de Psiquiatría, 1995, p. 52):

- A) El rendimiento en lectura, medido mediante pruebas de precisión o comprensión normalizadas y administradas individualmente, se sitúa sustancialmente por debajo de lo esperado para la edad cronológica del sujeto, su coeficiente de inteligencia y la escolaridad propia de su edad.
- B) La alteración del criterio A interfiere significativamente el rendimiento académico y las actividades de la vida cotidiana que exigen habilidades de lectura.
- C) Si hay un déficit sensorial, las dificultades para la lectura exceden de las habitualmente asociadas a él.

Básicamente hasta esa fecha se sostuvo el *criterio de discrepancia* entre el nivel intelectual y el rendimiento en lectura para la identificación de niños con dificultades específicas de aprendizaje (Fuchs & Fuchs, 2006; Jiménez, 2012).

Sin embargo, en el año 2003, la Asociación Internacional de Dislexia y el Instituto Nacional de Salud Infantil y Desarrollo Humano (*International Dyslexia Association / National Institute of Child Health and Human Development*), adoptaron la siguiente definición (Lyon, Shaywitz, & Shaywitz, 2003, p.2):

La dislexia es un trastorno (*disability*) específico del aprendizaje, de origen neurobiológico. Se caracteriza por la presencia de dificultades en la precisión y/o la fluidez en el reconocimiento de palabras y por habilidades pobres de ortografía (*spelling*) y decodificación. Estas dificultades típicamente son el resultado de un déficit en el componente fonológico del lenguaje, a menudo inesperado en relación con otras habilidades cognitivas y con la provisión de una enseñanza escolar efectiva. Las consecuencias secundarias pueden incluir problemas en la comprensión lectora y reducción de la experiencia lectora que pueden impedir el crecimiento del vocabulario y el conocimiento general del mundo.

Tales criterios de delimitación, basados en la evidencia empírica de múltiples investigaciones, se han considerado en el DSM-5 recién en el año 2013. Las dificultades en la lectura aparecen ahora bajo la denominación de

trastorno específico del aprendizaje, sobre el que se hacen las siguientes consideraciones (Asociación Americana de Psiquiatría, 2013, pp. 38-39) :

A. Dificultad en el aprendizaje y en la utilización de las aptitudes académicas, evidenciado por la presencia de al menos uno de los siguientes síntomas que han persistido por lo menos durante 6 meses, a pesar de intervenciones dirigidas a estas dificultades:

1. Lectura de palabras imprecisa o lenta y con esfuerzo (p. ej., lee palabras sueltas en voz alta incorrectamente o con lentitud y vacilación, con frecuencia adivina palabras, dificultad para expresar bien las palabras).
2. Dificultad para comprender el significado de lo que lee (p. ej., puede leer un texto con precisión pero no comprende la oración, las relaciones, las inferencias o el sentido profundo de lo que lee).
3. Dificultades ortográficas (p. ej., puede añadir, omitir o sustituir vocales o consonantes).
4. Dificultades con la expresión escrita (p. ej., hace múltiples errores gramaticales o de puntuación en un oración; organiza mal el párrafo; la expresión escrita de ideas no es clara).

B. Las aptitudes académicas afectadas están sustancialmente y en grado cuantificable por debajo de lo esperado para la edad cronológica del individuo, e interfieren significativamente con el rendimiento académico o laboral, o con actividades de la vida cotidiana, que se confirman con medidas (pruebas) estandarizadas administradas individualmente y una evaluación clínica integral. En individuos de 17 y más años, la historia documentada de las dificultades del aprendizaje se puede sustituir por la evaluación estandarizada.

C. Las dificultades de aprendizaje comienzan en la edad escolar pero pueden no manifestarse totalmente hasta que las demandas de las aptitudes académicas afectadas superan las capacidades limitadas del individuo (p. ej., en exámenes programados, la lectura o escritura de informes complejos y largos para una fecha límite inaplazable, tareas académicas excesivamente pesadas).

D. Las dificultades de aprendizaje no se explican mejor por discapacidades intelectuales, trastornos visuales o auditivos no corregidos, otros trastornos mentales o neurológicos, adversidad psicosocial, falta de dominio en el lenguaje de instrucción académica o directrices educativas inadecuadas.

Asimismo se consigna que se debe especificar si se trata de *dificultades en la lectura*, diferenciándola de las que refieren a la *expresión escrita y a las matemáticas*. Respecto de las primeras se aclara que corresponden a la “precisión en la lectura de palabras, velocidad o fluidez en la lectura, comprensión en la lectura” (p. 39). A continuación se observa que la *dislexia* “es un término alternativo utilizado para referirse a un patrón de dificultades del aprendizaje que se caracteriza por problemas con el reconocimiento de palabras en forma precisa o fluida, deletrear mal y poca capacidad ortográfica” (p. 39). Desde la perspectiva diagnóstica se indica, además, “*especificar*” la gravedad actual: leve, moderada o grave.

La modificación de los criterios de delimitación, reflejan el cambio operado. No obstante, fue a partir de la definición de la Asociación Internacional de Dislexia y el Instituto Nacional de Salud Infantil y Desarrollo Humano, basada en los resultados de múltiples investigaciones, que comenzó a aparecer un amplio acuerdo acerca de las manifestaciones del trastorno y sobre su causa cognitiva: *un déficit en el componente fonológico del lenguaje*. Es necesario notar, sin embargo, que el mismo no afecta a las habilidades básicas de percepción y producción del lenguaje necesarias para la comunicación oral, sino a un conjunto de procesos fonológicos implicados en tareas más demandantes: las habilidades de conciencia fonológica, de denominación automatizada rápida, la capacidad de memoria de corto plazo verbal y la memoria de trabajo (Defior & Serrano, 2011). En una tarea típica de conciencia fonológica, el sujeto debe sostener la representación fonológica de una palabra y manipular –segmentar, sustraer, agregar, intercambiar, comparar– de manera explícita las unidades que la componen, lo cual, además, involucra necesariamente a la memoria de

trabajo. En las tareas de denominación automatizada rápida, el sujeto debe nombrar lo más rápidamente posible una serie de elementos familiares.

Este trastorno, así como el déficit cognitivo subyacente, suele persistir en la adolescencia y la adultez, aun cuando las personas afectadas han sido oportunamente expuestas a material de lectura adecuado para su edad y han recibido una enseñanza de la lectura centrada en el entrenamiento de las habilidades de conciencia fonémica y en la comprensión del principio alfabético. Estas prácticas, implementadas en el aula y/o en clases de remediación, pueden producir una notable mejoría en la precisión de la decodificación en estos niños, que eventualmente puede aproximarse al nivel de sus pares. Sin embargo, el déficit fonológico subyacente, sumado al retraso en la familiarización con las formas escritas de las palabras, impide la automatización de la decodificación manifestándose en la esfera de la fluidez en la lectura en voz alta y en menor medida en la comprensión de textos, a la vez que dificulta el establecimiento de representaciones ortográficas (deletreo) (Fletcher, Lyon, Fuchs, & Barnes, 2007).

En la actualidad existe consenso en considerar que las dificultades en el aprendizaje de la lectura, están entre las causas más frecuentes de fracaso escolar. Particularmente en el caso de las que afectan a la decodificación, constituye la forma mejor estudiada así como la más habitual de las dificultades del aprendizaje en general (Fletcher, Lyon, Fuchs, & Barnes, 2007; Shaywitz, Morris, & Shaywitz, 2008). La importancia de comprender cuáles son los factores que influyen en la manifestación de estas últimas reside en el rol que juegan en la adquisición de nuevos aprendizajes académicos mediada por las habilidades de comprensión de textos.

1.3.1. La dislexia en español

Del mismo modo en que el grado de transparencia/opacidad del sistema alfabético influye en el ritmo con que los niños logran dominar las estrategias para el reconocimiento de las palabras escritas, también influye

en la manifestación de las dificultades. En español, el alto grado de consistencia de las RCFG y la estructura silábica relativamente simple de la lengua oral confluyen para que, pese al déficit fonológico, los niños disléxicos que adquieren el español logren un mayor desarrollo de la estrategia fonológica que los niños disléxicos que adquieren sistemas opacos. Sin embargo, esta estrategia no alcanza el nivel de automatización ni constituye un mecanismo de auto-enseñanza para el desarrollo de la estrategia ortográfica (Share, 1995). En virtud de estas características del sistema del español, los niños disléxicos presentan un rendimiento relativamente más pobre en tareas de fluidez, conocimiento ortográfico y comprensión que en tareas de decodificación (Carrillo & Alegría, 2009).

Por las mismas causas, los niños disléxicos que aprenden a leer inglés presentan más frecuentemente el patrón de alteraciones del subtipo de la “dislexia fonológica” que el de la “dislexia de superficie”. Los niños con dislexia fonológica presentan un déficit más marcado en la estrategia fonológica que en la estrategia ortográfica y exhiben un peor rendimiento en la lectura en voz alta de pseudopalabras que en la de palabras. Por otro lado, en los niños con dislexia de superficie, la estrategia ortográfica es más deficitaria que la estrategia fonológica, lo cual se manifiesta en un peor rendimiento en la lectura en voz alta de palabras de ortografía irregular que en la de palabras regulares y pseudopalabras. En cambio, en español, los niños disléxicos presentan más frecuentemente el patrón de la dislexia de superficie que el de la dislexia fonológica. Cabe aclarar que los estudios en español no son totalmente comparables a los estudios en inglés ya que en español prácticamente no existen palabras de ortografía irregular, por lo que la estrategia ortográfica se evalúa a través de pruebas de dictado de palabras con mapeos fonema-grafema inconsistentes, decisión léxica con pseudohomófonos y comprensión de homófonos (Carrillo & Alegría, 2009; Carrillo, Alegría, & Calvo, 2008; Jiménez, Rodríguez, & Ramírez, 2009).

1.3.2. Fuentes de variabilidad que influyen en la manifestación de las dificultades en el aprendizaje de la lectura

Al igual que la mayoría de los atributos psicológicos, las aptitudes/dificultades en el aprendizaje en general, y en el aprendizaje de la lectura en particular, constituyen variables continuas que resultan de la interacción de un conjunto heterogéneo de factores vinculados a las características del niño y del ambiente. La ubicación de los distintos factores en un esquema de niveles de análisis ha contribuido a asignarles un peso relativo tanto en la definición de las alteraciones, su identificación y en las estrategias de intervención.

En un primer nivel de análisis, el conductual, las dificultades en el aprendizaje de la lectura se definen en términos del desempeño de los niños en uno o más de un dominio académico específico: la precisión, la fluidez, y la comprensión en la lectura.

En un segundo nivel de análisis, se ubican los factores cognitivos y psicosociales que subyacen directamente al bajo rendimiento académico. Por ejemplo, en cuanto a los factores cognitivos, está ampliamente aceptado que las dificultades en la precisión en la decodificación se deben a un déficit fonológico, que los déficits en la velocidad de denominación contribuyen a las dificultades en la fluidez, y que los niños con dificultades en la comprensión lectora se caracterizan por no disponer de estrategias cognitivas y metacognitivas para elaborar una representación integrada del significado del texto. A su vez, respecto de los factores psicosociales, el bajo desempeño escolar suele provocar baja autoestima y escasa motivación en la tarea. Recíprocamente, los déficits atencionales y motivacionales impactan en el funcionamiento cognitivo contribuyendo también al bajo rendimiento académico.

Finalmente, en un tercer nivel de análisis se ubican los factores neurobiológicos y ambientales. Por un lado, entre los factores neurobiológicos se encuentran las influencias genéticas y las anomalías en la estructura y la función de las regiones cerebrales hipotéticamente

involucradas en el procesamiento de la información relevante. Por otro lado, los factores ambientales también contribuyen a la variabilidad en el rendimiento académico. Así, por ejemplo, las características socioeconómicas del entorno de crianza influyen en las nociones que tienen los niños acerca de los usos del lenguaje escrito, mientras que las particularidades del sistema de escritura y el método de alfabetización adoptado en la escuela repercuten en el dominio de las habilidades necesarias que requieren de enseñanza explícita ya que típicamente no se adquieren de manera espontánea (Fletcher, Lyon, Fuchs, & Barnes, 2007; Frith, 1995).

En una propuesta similar, Vellutino y Fletcher (2005) hacen una distinción entre las causas manifiestas y las causas subyacentes del trastorno de lectura que caracteriza a la dislexia. Las causas manifiestas involucran la dificultad para adquirir los conocimientos y las habilidades necesarios para lograr una lectura competente. Entre estos conocimientos y habilidades se incluyen las habilidades de conciencia fonémica, el conocimiento del sonido de las letras y la sensibilidad a las regularidades ortográficas, que redundan en la identificación de las palabras escritas. Por otro lado, las causas subyacentes refieren a los déficits cognitivos de origen neurobiológico y a las características ambientales que interfieren en dicha adquisición. Como se mencionó en el párrafo anterior, la existencia de un déficit fonológico es la hipótesis más ampliamente aceptada como la causa que subyace a las pobres habilidades de conciencia fonémica en la dislexia. En cuanto a los factores ambientales, por ejemplo, se ha demostrado que las habilidades de conciencia fonémica dependen de la provisión de una enseñanza explícita y que la sensibilidad a las regularidades ortográficas dependen del grado de exposición a material escrito.

1.4. La identificación de las dificultades específicas en el aprendizaje de la lectura

La identificación de las dificultades específicas en el aprendizaje en general y del aprendizaje de la lectura en particular tiene objetivos diferentes

según el ámbito en el que se lleve a cabo. En el ámbito de la investigación básica, la identificación sirve para la selección de los sujetos –en estudios de caso único o de grupos– a los cuales se le administrarán las pruebas experimentales dirigidas a probar las hipótesis de investigación. En cambio, en los ámbitos de aplicación educativa y clínica, el objetivo es guiar la toma de decisiones acerca de las estrategias de intervención probadamente eficaces para la mejoría en el rendimiento académico.

En todo caso, la identificación es un proceso que depende de la definición de las “dificultades específicas en el aprendizaje de la lectura”. Los distintos enfoques para la identificación han focalizado en uno u otro componente de tal constructo. En los apartados que siguen se caracterizan el enfoque tradicional basado en la discrepancia entre la capacidad intelectual general y el rendimiento académico en lectura y el enfoque basado en la respuesta a la intervención (RTI).

1.4.1. El enfoque basado en la discrepancia entre la capacidad intelectual general y el rendimiento académico en lectura

El enfoque tradicional para la identificación de las dificultades específicas del aprendizaje de la lectura ha focalizado en el carácter inesperado del bajo rendimiento académico, dado el nivel de inteligencia general (o bien la edad, tal como se señala en el DSM-5). Este componente del constructo ha sido operacionalizado por el hallazgo de una discrepancia entre una medida de la capacidad intelectual general y una medida de la habilidad lectora. Por lo tanto, una vez que se ha determinado el bajo rendimiento en lectura, la identificación apunta diferenciar a los niños cuya capacidad intelectual general es superior a su habilidad lectora, de los niños con capacidades intelectuales generales descendidas y acordes con su bajo rendimiento en lectura. Adicionalmente, se evalúa una serie de criterios de exclusión, es decir, la influencia de otros factores que podrían ser la causa del bajo desempeño, que entonces resultaría esperado: déficits sensoriomotores no corregidos, alteración atencional, trastornos

emocionales, desventaja socio-económica, diversidad lingüística y enseñanza inadecuada o insuficiente.

Restori, Katz y Lee (2009) afirman que el carácter inesperado dado el nivel de inteligencia adquirió relevancia como componente del constructo de las dificultades específicas en el aprendizaje de la lectura a partir de un trabajo de Rutter y Yule (1975, citado en Restori, Katz, & Lee, 2009) en el que los autores diferenciaron dos tipos de dificultades. Por un lado, definieron el retraso general en lectura (*general reading backwardness*) como un nivel de lectura inferior al esperado para la edad cronológica del niño, y por otro lado, el retraso específico en lectura (*specific reading retardation*) como un nivel de lectura inferior al esperado para la inteligencia del niño. La búsqueda de una discrepancia continuó predominando en el proceso de identificación hasta hace poco tiempo, lo cual refleja la importancia atribuida al carácter inesperado de las dificultades específicas del aprendizaje dado el nivel de inteligencia.

Sin embargo, Restori, Katz y Lee (2009) notan que el enfoque ha generado problemas que fundamentan las críticas surgidas desde los comienzos del planteo y para las cuales existe evidencia acumulada. En primer lugar, el enfoque no permite realizar una detección y una intervención tempranas debido a que se requiere esperar a que el niño exhiba un retraso significativo en su habilidad lectora –usualmente entre 1 y 2 desvíos típicos por debajo del rendimiento promedio de acuerdo a la edad, según las distintas definiciones. Al respecto, los autores hacen referencia al hecho alarmante de que la probabilidad de identificación son máximas en tercero y cuarto grados de primaria. Es decir que en el momento de ser identificados, la mayoría de los niños con dificultades ha perdido entre uno y dos años de exposición a la lectura, lo cual impacta negativamente en la automatización de los procesos de decodificación y reconocimiento léxico necesaria para liberar los recursos atencionales destinados a la comprensión de textos.

En segundo lugar, la aplicación de los criterios de discrepancia carece de uniformidad entre distintos evaluadores. Esta falta de uniformidad es

observada no solo en la elección de los instrumentos para medir la inteligencia y la habilidad lectora y la determinación del grado de discrepancia requerido entre ambas medidas, sino también en el cumplimiento mismo del criterio de discrepancia.

Otro problema es que los niños que exhiben dificultades en el aprendizaje de la lectura y una capacidad intelectual disminuida que no resultan en un grado significativo de discrepancia quedan excluidos del proceso de identificación y, en consecuencia, de la intervención dirigida a mejorar su rendimiento en lectura. Si el descenso en el nivel de inteligencia general no cumple el criterio estadístico para retraso mental, estos niños quedan también al margen de todo intento explícito de mejorar su rendimiento académico (Fuchs & Fuchs, 2006).

Además, se cuestiona la premisa que fundamenta la búsqueda de una discrepancia. Dicha premisa asume una correlación alta entre las medidas de inteligencia y las de rendimiento académico. Sin embargo, el coeficiente de correlación raramente alcanza a ser de 0.60. Más allá de las interpretaciones acerca del tamaño del efecto, este valor equivale a afirmar que ambas medidas comparten sólo un 36 % de la varianza.

Finalmente, los autores ponen en duda la confiabilidad de las mediciones utilizadas para determinar la existencia de una discrepancia así como la validez de su interpretación. Por un lado, como sucede con toda medición, la confiabilidad de la medición del nivel de inteligencia y del rendimiento académico resulta potencialmente inestable cuando el instrumento se administra una única vez ya que la falta de medidas repetidas no permite que se anulen los efectos aleatorios del error de medición. Por otro lado, distintos estudios que comparan los correlatos cognitivos, la respuesta a la intervención, el pronóstico, y factores neurobiológicos como la heredabilidad y los patrones de activación cerebral, en dos grupos de niños cuyo bajo rendimiento en lectura es discrepante respecto del nivel de inteligencia, o bien, es acorde con el nivel de inteligencia, no resultaron en diferencias entre ambos grupos. Es decir, la evidencia empírica no permite

interpretar de manera válida el hallazgo de una discrepancia entre el nivel de inteligencia y el rendimiento en lectura como un buen indicador de la existencia de una dificultad específica en el aprendizaje de la lectura.

Sin embargo, cabe aclarar que el uso del enfoque de discrepancia está justificado en los estudios que comparan el desempeño de niños disléxicos con el de niños con desarrollo típico de la lectura. En estos estudios, las tareas utilizadas pueden imponer demandas cognitivas no relacionadas con los procesos de lectura y la comparación del rendimiento de grupos de niños que difieren en el nivel de inteligencia podrían arrojar diferencias atribuidas a la habilidad de lectura que resultarían espurias (Snowling, 2000).

1.4.2. El enfoque basado en la respuesta a la intervención

El enfoque para la identificación de las dificultades específicas del aprendizaje basado en la respuesta a la intervención (RTI) focaliza en la persistencia del bajo rendimiento en lectura pese a una enseñanza/intervención adecuada, es decir, probadamente eficaz para la adquisición de las habilidades de decodificación.

Al igual que el enfoque tradicional, el objetivo es detectar a los niños con dificultades “verdaderas”. Sin embargo, aquí el rasgo que distingue a estos niños de otros que presentan bajo rendimiento es que no responden a la enseñanza/intervención probadamente eficaz. Por lo tanto, la evaluación dirigida a la identificación se divide en varias etapas sucesivas, cada una de las cuales se compone de una instancia de evaluación seguida de un período de intervención.

La evaluación en la primera etapa se propone la detección temprana y oportuna de aquellos niños que presentan *riesgo* de desarrollar dificultades en el aprendizaje de la lectura. En esta primera instancia se evalúa a todos los niños y se recomienda utilizar instrumentos de *screening*, es decir, breves y que provean un panorama general del desempeño en lectura de cada niño, más que una descripción detallada de los procesos cognitivos afectados y conservados. La detección del riesgo también puede ser

implementada antes del ingreso a la escuela primaria, en cuyo caso se examinan las habilidades consideradas predictoras del rendimiento en decodificación, a saber, las habilidades de conciencia fonológica, el conocimiento de las letras y la denominación automatizada rápida. Los niños que exhiben bajo rendimiento en esta primera instancia de evaluación son incluidos en un programa de enseñanza más intensiva y sistemática con estrategias probadamente eficaces durante una cantidad limitada de tiempo.

Una vez finalizada la primera etapa de enseñanza/intervención, las instancias posteriores de evaluación están dirigidas a medir el progreso en las habilidades del niño. Los resultados de la eficacia del programa sirven para la toma de decisiones en torno al ajuste de los diversos aspectos del mismo: intervención individual o grupal, tamaño de los grupos, duración de las sesiones en minutos, frecuencia de las mismas, duración del programa en semanas, materiales a utilizar, tareas a ejercitar. Una vez que el niño alcanza el nivel de sus pares del aula se da por finalizada la intervención. Si, en cambio, el niño no cumple con un criterio previamente establecido de mejoría en el rendimiento luego de la implementación de un número previamente estipulado de estrategias de intervención, se realiza la derivación a una instancia clínica de evaluación. La evaluación clínica está dirigida a estudiar el perfil cognitivo del niño y puede resultar en la identificación de una dificultad específica en el aprendizaje de la lectura, o bien, de otra condición causal del bajo rendimiento.

A diferencia de la identificación y la intervención en el enfoque basado en la discrepancia entre la capacidad intelectual general y el rendimiento académico, que típicamente se ha implementado en el ámbito clínico, en el enfoque RTI, se propone que la identificación y la intervención se realicen en el marco de la escuela, hasta el momento de la derivación a una evaluación clínica. La identificación y la intervención tempranas evitan así las pérdidas insalvables en la cantidad de exposición al material escrito que indefectiblemente interfieren en el logro de la fluidez y la comprensión lectora, a la vez que se eluden las consecuencias de la “estigmatización”.

1.4.2.1. Detección del riesgo. La efectividad de los tests de *screening*. Sensibilidad y especificidad

De lo expuesto en el apartado anterior se deduce que el *screening* universal constituye un rasgo central del enfoque RTI ya que minimiza los costos materiales, de tiempo y de formación de los docentes en técnicas de evaluación, a la vez que permite la asignación de los recursos limitados (cuando no escasos) de los sistemas de educación a la implementación temprana y oportuna de intervenciones probadamente eficaces. Esto redundará en un beneficio para todos los niños en general, y para aquellos que presentan dificultades en particular.

Tras un período que abarca generalmente uno o más años escolares, es posible estudiar la efectividad de la herramienta de *screening* como medida predictora de las dificultades en el aprendizaje de la lectura a través de la implementación de una segunda instancia de evaluación dirigida a la identificación de las mismas. Las herramientas utilizadas en esta segunda instancia incluyen exámenes basados en el currículum con material acorde a la edad, que arrojan medidas de comprensión lectora y/o lectura en voz alta de un texto de manera fluida y con prosodia adecuada.

La efectividad de una medida del riesgo de desarrollar dificultades depende de un balance entre su sensibilidad y su especificidad. Por un lado, la sensibilidad refiere al poder para predecir la identificación de los sujetos que desarrollarán dificultades persistentes. Se expresa mediante el porcentaje de casos “positivos verdaderos” (PV), es decir, sujetos que obtuvieron puntajes de riesgo en la instancia de *screening* y que pese a recibir una instrucción adecuada durante el período estipulado mostrarán bajo desempeño en la instancia de identificación, del total de casos con bajo rendimiento en la instancia de identificación incluidos aquéllos que no exhibieron riesgo (“falsos negativos”; FN). De modo que la ecuación sería:

$$\text{Sensibilidad} = [\text{PV} / (\text{PV} + \text{FN})] \times 100$$

Por otro lado, la especificidad refiere al poder para excluir de la identificación a aquellos sujetos que no presentarán dificultades persistentes. Se expresa mediante el porcentaje de casos “negativos verdaderos” (NV), vale decir, sujetos con buen desempeño en ambas instancias de evaluación, del total de casos con buen rendimiento en la instancia de identificación incluidos aquéllos que obtuvieron puntajes de riesgo (“falsos positivos”; FP). En este caso, la ecuación sería:

$$\text{Especificidad} = [\text{NV} / (\text{NV} + \text{FP})] \times 100$$

Finalmente, la efectividad de una medida de *screening* es el porcentaje de sujetos clasificados correctamente en la primera instancia de evaluación, del total de sujetos evaluados, esto es:

$$\text{Efectividad} = [(\text{PV} + \text{NV}) / (\text{PV} + \text{NV} + \text{FP} + \text{FN})] \times 100$$

Habida cuenta de que los errores de clasificación constituyen un rasgo inherente a las herramientas de *screening* al producir un número variable de falsos positivos (sobre-identificación) así como de falsos negativos (sub-identificación), un número de trabajos se ha centrado en el estudio de la efectividad de distintos modelos de identificación utilizando distintas medidas para la detección del riesgo y para la identificación de las dificultades, distintos puntajes de corte, en distintos momentos del desarrollo (Andrade, Andrade, & Capellini, 2015; Clemens, Shapiro, & Thoemmes, 2011; Petscher & Kim, 2011; Vellutino, Scanlon, Zhang, & Schatschneider, 2008; Wilson & Lonigan, 2010). No obstante la diversidad de modelos investigados, existe un consenso entre los diferentes autores acerca del nivel mínimo aceptable de sensibilidad, estipulado en el 90%, mientras que el nivel de especificidad deseable ronda el 85%.

Adicionalmente, el estudio de la correlación entre la medida de *screening* y la de identificación de las dificultades constituye una fuente de evidencia de validez de criterio para la primera.

En consecuencia, desde una perspectiva preventiva y de intervención, se hace necesario contar con instrumentos de *screening* destinados a la detección temprana y oportuna del riesgo de presentar dificultades en el aprendizaje de la lectura.

En cuanto al TECLE Buenos Aires, consideramos que por sus características permite su aplicación rápida y oportuna a todos los alumnos para la identificación de los que pueden considerarse “en riesgo” y el monitoreo de su progreso en lectura en distintos momentos de la escolaridad a partir de una intervención adecuada.

Los estudios realizados para la presente tesis, dirigidos a la obtención de evidencias de confiabilidad, de validez de constructo y de criterio, así como a la confección de los baremos, constituyen una instancia necesaria para las investigaciones acerca de la efectividad de esta herramienta. De este modo sería posible contar con indicadores que posibiliten distinguir aquellos niños que presentan dificultades específicas, que responden de manera diferencial a la intervención de aquellos que exhiben bajo rendimiento pero que responden con mayor rapidez y adecuación. Se trataría en estos casos de niños cuyo desempeño refleja oportunidades educativas insuficientes, y no una discapacidad (Fuchs & Fuchs, 2006).

En función de lo expuesto, en los próximos capítulos se presentan los resultados de los estudios realizados así como las conclusiones a las que han dado lugar.

Capítulo 2

El Test Colectivo de Eficacia Lectora. Su adaptación

El Test colectivo de eficacia lectora (en adelante, TECLE) fue desarrollado en español por Marín y Carrillo (1999) siguiendo la estructura de la prueba en francés L3 del test D-OR-LEC de Lobrot (1973). Está destinado a incluir una medida única de lectura, referida al constructo *eficacia lectora*, sin intentar un análisis diferenciado de los procesos que la componen.

Se trata de un test de lápiz y papel integrado por 64 ítems, compuestos por oraciones incompletas. La tarea consiste en completar, en un tiempo límite de cinco minutos, la mayor cantidad posible de oraciones a las que les falta la última palabra, seleccionándola entre cuatro opciones de respuesta. Las cuatro opciones son similares en cuanto a la forma ortográfica y fonológica e incluyen dos palabras (la opción correcta y otra palabra que no completa coherentemente la oración) y dos pseudopalabras. El puntaje se obtiene de la sumatoria de ítems respondidos correctamente en el tiempo permitido.

Previamente al inicio de la prueba, y luego de impartida la consigna, se les brinda a los participantes tres ejemplos de complejidad léxica, sintáctica y semántica similar a la de los ítems de prueba con el objeto de verificar la comprensión de la consigna.

Podemos afirmar que el TECLE constituye una herramienta de *screening*, que podría resultar adecuada para la detección de aquellos alumnos que presentan riesgo de desarrollar o presentar dificultades en el aprendizaje de la lectura, teniendo en cuenta la estructura y características del test, entre las que destacamos las siguientes:

- Su administración es sencilla y puede ser colectiva.
- Su realización demanda un tiempo breve.

- Los puntajes que proporciona ofrecen una estimación global del constructo “eficacia lectora”, sin la posibilidad de inferir el grado de dominio en cada una de las habilidades que hipotéticamente intervienen en su ejecución.

2.1. Elaboración y adaptación del TECLE al español rioplatense

2.1.1. Adaptación destinada a la población uruguaya

El test fue adaptado por Cuadro, Costa, Trías y Ponce de León (2009) para su aplicación en población escolar de Uruguay, para la que están disponibles los datos normativos obtenidos de una muestra de 2336 participantes de 2° a 6° grados de 19 escuelas primarias de gestión privada de la ciudad de Montevideo.

Los autores realizaron modificaciones en el contenido de un total de 6 ítems, que consistieron en la sustitución de alguna/s palabra/s de la frase a completar, o bien en la de alguna/s de la/s opción/es. En sólo uno de esos 6 ítems se modificaron totalmente la oración y las opciones. Además, modificaron el orden de presentación de 3 ítems, ubicándolos al final, ya que en un estudio realizado con el objeto de calcular su índice de dificultad observaron que el porcentaje de alumnos que los respondieron correctamente rondaba el 50%. La razón por la cual estos ítems “más difíciles” fueron ubicados al final es que se consideró que el TECLE era un test de velocidad, en el que todos los ítems deberían ser fáciles.

2.1.2. Adaptación destinada a población argentina

Ferreres, Abusamra, Casajús y China (2011) realizaron una primera adaptación del test a las particularidades léxico-sintácticas del español de la Ciudad Autónoma de Buenos Aires y alrededores. La misma estuvo dirigida a sustituir expresiones, formas gramaticales y términos inusuales en el español local por versiones más familiares. El procedimiento resultó en la modificación de 52 de los 64 estímulos originales: en 17 ítems se modificó la oración a completar, en 4 ítems se modificaron las opciones de respuesta,

en 19 ítems se modificaron la oración a completar y las opciones de respuesta, y en 12 ítems se reemplazó todo el ítem. Además, se redujo el número de ítems de práctica, de tres a dos. Cabe señalar que se trata en este caso de una adaptación que refiere a las características lingüísticas del español hablado en la Ciudad Autónoma de Buenos Aires y alrededores, que si bien resulta similar al hablado en Uruguay, presenta peculiaridades que deben contemplarse. Asimismo, para esta versión no se cuenta con datos referidos a la confiabilidad y a la validez de la prueba ni con datos normativos.

2.1.3. Modificaciones ulteriores realizadas para esta tesis

Se incluyeron modificaciones adicionales a la adaptación referida en el apartado anterior que consistieron principalmente en el agregado de información en la oración a completar, con el objeto de restringir el campo semántico de las respuestas correctas posibles, y así minimizar la dificultad de los ítems, (por ejemplo, las oraciones “Se rompió el > cartón”, “Se oye al > trompetista” y “Han atrapado un > delfín” fueron sustituidas por “Para abrir más rápido su regalo, Lucía rompió el > cartón”, “Jorge no logra dormirse porque desde su habitación oye al > trompetista” y “En las redes del barco pesquero, quedó atrapado un > delfín”).

Además, se cambió el orden de las opciones con el objeto de que la opción correcta ocupara las distintas posiciones aproximadamente el mismo número de veces a lo largo de los 64 ítems y se cambió el orden de los ítems modificados a fin de que no quedaran más de dos oraciones consecutivas con la opción correcta en la misma posición.

Por último, se cambiaron las opciones de respuesta del primer ítem de práctica (el ítem “Tu pelota es de color > rogo/roco/robo/rojo” fue sustituido por “Tu pelota es de color > vedre/veque/verte/verde”) ya que resultaban similares a las opciones del segundo, y se sustituyeron dos opciones del segundo ítem, ya que una había sido incluida erróneamente como una pseudopalabra y la otra era un pseudohomófono (el ítem “El caballo tenía la

pata > ropa/rola/rota/roka” fue sustituido por “El caballo tenía la pata > ropa/rota/rofa/rona”).

En adelante, nos referiremos a esta versión como TECLE Buenos Aires (ver Anexo C “El TECLE Buenos Aires. Protocolo de respuestas y normas de aplicación y calificación”).

2.2. El constructo “eficacia lectora”: interpretación de los puntajes del test

El test está destinado a ofrecer una medida global del constructo “eficacia lectora”, que comprende, como se señalara precedentemente, los aspectos específicos y no específicos incluidos en la capacidad de lectura. Los aspectos específicos refieren a la precisión y la velocidad en el reconocimiento ortográfico y la decodificación, necesarios tanto para la lectura de la oración como para el análisis rápido de las opciones (por ejemplo, para descartar rápidamente las pseudopalabras). Los aspectos no específicos incluyen habilidades lingüísticas generales (sintácticas y semánticas) requeridas para comprender la oración y descartar la palabra que no la completa de manera coherente, así como habilidades cognitivas generales que permiten administrar los recursos de procesamiento para llevar a cabo la tarea en el tiempo permitido (Carrillo, Alegría, & Marín, 2013; Cuadro, Costa, Trías, & Ponce de León, 2009; Ferreres, Abusamra, Casajús, & China, 2011; Marín, Cuadro, & Pagán, 2007).

En la adaptación del test a la población escolar de Uruguay (ver apartado 2.1.1), Cuadro, Costa, Trías y Ponce de León (2009) presentaron datos normativos de dos tipos de puntaje: el puntaje directo y el puntaje corregido. El puntaje directo consiste en la sumatoria de ítems respondidos correctamente, mientras que el puntaje corregido se obtiene luego de penalizar los errores y omisiones. En su estudio, los autores consideraron que los puntajes corregidos eran un reflejo más acertado de la capacidad lectora verdadera de los niños, si bien afirmaron que las diferencias entre ambas medidas no fueron significativas. La razón para incluir los puntajes

corregidos fue que se consideró que el TECLE era, dentro del tipo de tests de ejecución máxima, un test de velocidad. En la presente tesis, en cambio, se consideró únicamente el puntaje directo.

2.2.1. Tests de potencia versus tests de velocidad

Los tests de potencia y los tests de velocidad se distinguen por el factor que produce la mayor variabilidad en el rendimiento. Así, en los tests de potencia, la variabilidad en el rendimiento se explica por el nivel alcanzado en una aptitud o conocimiento específico, mientras que en los tests de velocidad depende de la velocidad de respuesta.

Por lo tanto, los tests de potencia típicamente incluyen ítems con distintos grados de dificultad con el objeto de revelar las diferencias interindividuales en la aptitud o conocimiento específico de interés. Los tests de velocidad, en cambio, incluyen un mayor número de ítems con niveles comparables de dificultad y una alta probabilidad de ser respondidos correctamente por la mayoría de los sujetos. Además, habitualmente se administran con tiempo límite.

Cabe aclarar que consideramos que el TECLE no puede ser caracterizado como un test puro de velocidad ya que, tal como afirmamos, los puntajes observados se interpretan en términos de “eficacia lectora”. Esto es, nuestra hipótesis refiere a que la variabilidad en el rendimiento depende de una combinación de factores vinculados con los aspectos específicos y no específicos de la lectura, entre los cuales se incluye la velocidad de procesamiento (ver capítulo 4 “Estudios de validez del TECLE Buenos Aires”).

2.3. La presente investigación

Entre las directrices proporcionadas por la Comisión Internacional de Tests para la adaptación de un instrumento de un contexto lingüístico y cultural a otro, se encuentra la necesidad de establecer sus propiedades psicométricas de validez y confiabilidad, contar con datos normativos

propios, así como la de proveer a los usuarios del test de la documentación necesaria para su uso correcto en la población a la que está dirigido (International Test Commission, 2017; Muñiz, Elosua, & Hambleton, 2013).

2.3.1. Interrogantes que orientaron la investigación

A partir del análisis de los antecedentes fue posible formular los siguientes interrogantes que orientaron la elaboración de este trabajo:

- ¿El rendimiento en el TECLE Buenos Aires mostrará estabilidad a lo largo de un lapso razonable de tiempo?

- ¿Se observarán diferencias en las correlaciones entre el rendimiento en el test y el desempeño en pruebas que miden habilidades específicas de la lectura, por un lado, y en pruebas que miden habilidades generales, por el otro?

- ¿Existirán diferencias en el rendimiento en el test entre los niños con y sin diagnóstico de dificultades en el aprendizaje y/o dificultades del lenguaje?

- ¿Existirán diferencias en el desempeño en el test entre los niños según el grado escolar, la edad y el sexo, y según el nivel de oportunidades educativas, el tipo de gestión y la región geográfica de la escuela?

2.3.2. Objetivos

General

- Contribuir al desarrollo y aplicación de tests de *screening* para la evaluación de la lectura destinados a facilitar los procedimientos de monitoreo del curso de los aprendizajes y la detección temprana de dificultades eventuales.

Específicos

- 1) Estimar la estabilidad temporal de los puntajes que arroja el TECLE Buenos Aires para determinar su confiabilidad.

- 2) Obtener evidencias de validez de constructo y de criterio que posibiliten interpretaciones adecuadas de los puntajes observados en el test como una medida de “eficacia lectora”.
- 3) Examinar las variables “grado escolar”, “edad en años” y “sexo” de los niños, y “nivel de oportunidades educativas”, “tipo de gestión”, y “región geográfica” de las escuelas para identificar su influencia sobre el rendimiento en el test.
- 4) Confeccionar los baremos del test para la población de niños de 2° a 6° grados de educación primaria de la Ciudad Autónoma de Buenos Aires y alrededores.
- 5) Elaborar baremos segmentados en función de los resultados en el desempeño del test de acuerdo con las variables incluidas, para proporcionar elementos más adecuados para la interpretación de los puntajes.

2.3.3. Hipótesis

Las hipótesis planteadas para la presente tesis fueron:

- Se observará una correlación positiva estadísticamente significativa entre los puntajes observados en el TECLE Buenos Aires en las instancias de test y retest.
- Los puntajes observados en el test mostrarán asociaciones más fuertes con medidas que dependen de las habilidades específicas para la lectura, es decir, de decodificación, que con medidas que dependen de habilidades generales, lingüísticas y cognitivas.
- Los puntajes observados en el test en un grupo clínico con diagnóstico de dificultades del aprendizaje y/o del lenguaje serán significativamente inferiores a los observados en el grupo control con desarrollo típico.

- Los puntajes observados en el test mostrarán diferencias de distinto alcance según se trate del grado escolar y la edad en años de los niños, y del nivel de oportunidades educativas, el tipo de gestión y la región geográfica de las escuelas a las que asisten.

2.3.4. Metodología

La presente tesis se enmarca en un enfoque cuantitativo, dentro de los *Diseños Instrumentales* (Ato, López, & Benavente, 2013), destinados a la elaboración/adaptación de instrumentos en Psicología.

Se compone de tres estudios: 1) estudio de confiabilidad, 2) estudio de validez, que comprende la validez de constructo y la validez de criterio, y 3) estudio para la elaboración y justificación estadística de los datos normativos.

2.3.4.1. Participantes, instrumentos y procedimientos de obtención, elaboración y análisis datos

El desarrollo por separado de los tres estudios, que involucran participantes, instrumentos y procedimientos acordes a la naturaleza de cada uno, se ofrecen en los capítulos respectivos (capítulo 3 “Estudio de confiabilidad del TECLE Buenos Aires”, 4 “Estudios de validez del TECLE Buenos Aires” y 5 “Los datos normativos del TECLE Buenos Aires para la Ciudad Autónoma de Buenos Aires y alrededores. Justificación estadística”), con el propósito, además, de facilitar la lectura de la tesis.

Capítulo 3

Estudio de confiabilidad del TECLE Buenos Aires

Entre los requisitos fundamentales que aparecen entre las directrices de la Comisión Internacional de Tests para el desarrollo y uso de instrumentos (International Test Commission, 2017; Muñiz, Elosua, & Hambleton, 2013), aparece, como se señalara en el capítulo 2 “El Test Colectivo de Eficacia Lectora. Su adaptación”, la necesidad de proporcionar evidencia que respalde la confiabilidad de la versión adaptada del test en las poblaciones previstas.

De acuerdo con la Teoría Clásica de los Tests, el puntaje observado (empírico) de un sujeto en un test está determinado por su puntaje verdadero (teórico) – su nivel verdadero en el atributo que el test pretende medir – y por el error presente en toda medición. El error de medición resulta de la influencia de variables aleatorias no controladas (por ej., estado de ánimo, factores ambientales, diferencias en el grado en que los ítems de un test miden el atributo de interés, falta de uniformidad entre evaluadores en los criterios de calificación) y da lugar a diferencias entre los puntajes observados y los puntajes verdaderos. En razón de que estos últimos son inobservables, surge la necesidad de estimar la confiabilidad de los primeros.

Así, la confiabilidad refiere al grado en que los *puntajes observados en un test* están libres de error de medición, es decir, a la exactitud o precisión con que reflejan los puntajes verdaderos. Puesto que los errores de medición pueden ser de naturaleza muy diversa, se considera que existen varias dimensiones de la confiabilidad: la estabilidad temporal, la consistencia interna, y el acuerdo inter-evaluadores, que son estimadas a través de diferentes métodos (Olaz, 2008). Todos los métodos para estimar la confiabilidad se basan en el estudio de la correlación entre diferentes puntajes observados.

En el *método de test-retest* se estima la confiabilidad a partir del cálculo de un coeficiente de estabilidad, que expresa el grado de correlación entre los puntajes observados de un mismo grupo de sujetos en dos instancias de evaluación separadas por un intervalo de tiempo.

En el *método de división por mitades* se estudia la correlación entre los puntajes observados de un grupo de sujetos en las dos mitades de un test. Usualmente, una mitad se constituye con los ítems pares y la otra con los impares, con el objeto de asegurar que en ambas mitades se incluyan reactivos con un orden de dificultad semejante (lo cual no podría asegurarse si las mitades estuvieran conformadas, respectivamente, por los primeros y los últimos ítems ordenados según su dificultad). Como resultado se obtiene un coeficiente de consistencia interna que ofrece una medida de la confiabilidad de cada mitad. Dicho coeficiente es luego incluido en una fórmula para obtener una estimación de la confiabilidad del test en su totalidad.

La consistencia interna también se calcula a través del coeficiente alfa de Cronbach (1951), que se basa en el promedio de las correlaciones entre los puntajes observados de un grupo de sujetos en los ítems del test. Entre las ventajas de esta medida se encuentra la posibilidad de evaluar cuánto mejoraría (o empejaría) la confiabilidad del test si se excluyera un determinado reactivo. El estudio de la consistencia interna resulta adecuado para estimar la confiabilidad de los *tests de potencia*, que incluyen ítems con distintos grados de dificultad que permiten revelar las diferencias interindividuales en la aptitud o conocimiento específico de interés.

El *método de formas equivalentes, alternativas o paralelas* consiste en administrar dos formas equivalentes de un test a un grupo de sujetos. Para que las dos formas puedan ser consideradas equivalentes, deben compartir sus características formales (cantidad de ítems, nivel de medición de los puntajes obtenidos, etc.) y estadísticas (tener medias y desvíos estándar similares) y los puntajes que arrojan deben mostrar una correlación elevada (Furr & Bacharach, 2014; Olaz, 2008). La correlación entre los puntajes

observados en ambas formas se expresa a través de un coeficiente de equivalencia, que da una medida de la confiabilidad para ambas formas del test. De acuerdo con Olaz (2008), este método puede utilizarse para estimar la estabilidad temporal administrando ambas formas en dos instancias separadas por un intervalo de tiempo. Por otro lado, si ambas formas se administran en una única instancia de evaluación, este método también ofrece una medida de consistencia interna al estudiar el grado en que los ítems de ambas formas miden el constructo de interés. Sin embargo, el coeficiente de equivalencia podría resultar elevado de manera espuria debido al alto grado de asociación entre los puntajes de error, es decir, a efectos de contaminación. Por ejemplo, es probable que las condiciones internas de los sujetos evaluados (motivación, alerta) y las condiciones externas (factores ambientales, evaluador) en una única instancia de evaluación ejerzan la misma influencia en el rendimiento de dichos sujetos en cada una de las formas del test.

Por último, el *método de acuerdo inter-evaluadores* se utiliza para estimar la confiabilidad de aquellos tests en los que el proceso de registro, codificación y calificación de las respuestas de los sujetos depende del criterio del evaluador. El coeficiente a utilizar para estudiar la correlación entre los puntajes asignados por los distintos evaluadores dependerá del nivel de medición (nominal, ordinal, intervalar) utilizado para calificar las respuestas (Olaz, 2008).

La elección del método a seguir depende del tipo de instrumento cuya confiabilidad se va a estimar. Al respecto debe tomarse en consideración si se trata de un instrumento de rendimiento máximo, destinado a obtener el rendimiento mayor del que es capaz el examinado, o bien de ejecución típica, destinado a obtener una información sobre las características que habitualmente lo identifican.

A su vez, dentro de los primeros, es necesario diferenciar los instrumentos de dominio general y específico, según estén destinados a examinar características psicológicas globales o específicas,

respectivamente. En ambos casos, además existen tests de potencia o nivel y de velocidad. En los de potencia se trata de indagar el rendimiento de un examinado sin límite de tiempo en la totalidad de un conjunto de ítems, generalmente ordenados según la dificultad. En cambio, los de velocidad establecen un tiempo límite para responder el mayor número de ítems, que generalmente poseen niveles comparables de dificultad. Dicho de otro modo, en los tests de velocidad, las diferencias interindividuales en el rendimiento dependen de la velocidad de respuesta a los ítems propuestos.

Tal como se afirmó en el capítulo 2, el TECLE Buenos Aires es considerado un test de velocidad. Este tipo de tests se caracterizan por estar conformados por ítems de niveles comparables de dificultad, dado que no se proponen revelar las diferencias en el nivel de dominio de un grupo de sujetos en una capacidad específica, sino las diferencias en la velocidad de procesamiento. Bajo estas consideraciones, no resulta adecuado estudiar la consistencia interna puesto que arrojarían una medida de confiabilidad elevada de manera espuria (Anastasi & Urbina, 1998; Furr & Bacharach, 2014; Strauss, Sherman, Spreen, & Slick, 2006).

Para estos instrumentos, el método más adecuado para estudiar la confiabilidad es el método test-retest. Esto es, la mayor correlación entre los puntajes observados de un mismo grupo de sujetos en distintas instancias de evaluación permite inferir una menor influencia del error de medición, y por lo tanto, una mayor consistencia entre estos puntajes y los puntajes verdaderos.

Furr y Bacharach (2014) señalan que un factor que incide en la interpretación de una medida de la estabilidad temporal de los puntajes observados en un test es el intervalo de tiempo transcurrido entre las instancias de test y de retest. Los intervalos cortos conllevan el riesgo de que el rendimiento en la instancia de retest se vea contaminado por el aprendizaje del material o una disminución en la motivación debida a la evaluación repetida, mientras que en los intervalos largos pueden observarse cambios atribuibles a transformaciones en el constructo que se

pretende medir. Estos autores notan que muchos de los trabajos en los que se evaluó la estabilidad temporal de mediciones de rasgos relativamente estables utilizaron intervalos de 2 a 8 semanas. Otro factor que según estos autores influye en la estabilidad temporal de un conjunto de mediciones es la etapa del desarrollo en la que ocurre el intervalo transcurrido entre las instancias de evaluación.

En este sentido, la *eficacia lectora* es un constructo que se ve afectado por cambios en el curso de la escolarización, siendo los primeros años de la escolarización primaria el escenario de los cambios más rápidos. En esta etapa, las habilidades cognitivas relacionadas con la lectura experimentan el mayor crecimiento y son el blanco de los efectos de la enseñanza. Sin embargo, consideramos que dichos cambios no se observan en el transcurso de períodos cortos; esto es, la eficacia lectora es un constructo relativamente estable.

Por lo expuesto, el objetivo de este estudio fue el de estimar la confiabilidad de los puntajes que arroja el TECLÉ Buenos Aires a través del método test-retest. No obstante, tal como señaláramos en el capítulo 2, consideramos que el test no constituye un test puro de velocidad, sino que la variabilidad en los puntajes que arroja depende de una combinación de factores vinculados a los aspectos específicos y no específicos de la lectura, entre los que se incluye la velocidad de procesamiento.

La hipótesis de trabajo fue que se observará una correlación positiva estadísticamente significativa entre los puntajes observados en ambas instancias de evaluación.

3.1. Metodología

El estudio de confiabilidad seleccionado por las características de este test de *screening* fue su establecimiento a través del método de test-retest.

3.1.1. Participantes

Para este estudio se seleccionó una muestra por conveniencia de los alumnos de 3° a 6° grados de una escuela primaria de gestión estatal de la región noroeste del Gran Buenos Aires.

En la primera instancia de evaluación se examinaron 387 alumnos, y 351 en la segunda. Se excluyeron del análisis final 169 alumnos por las siguientes razones: ausencia en una de las instancias de evaluación, falta de ajuste a la consigna (por ejemplo, comenzar a realizar la prueba de atrás hacia delante o continuar completando el test más allá del tiempo límite), antecedentes de repetición del grado, diagnóstico de trastorno neurológico o del desarrollo, y niños institucionalizados. La muestra final quedó conformada por 218 niños (117 mujeres y 101 varones) de entre 8 y 12 años de edad (56 de 3° grado, 54 de 4°, 72 de 5° y 36 de 6°).

3.1.2. Instrumento

Se utilizó el TECLE Buenos Aires, cuyas características se especificaron en el capítulo 2.

3.1.3. Procedimientos

Previamente a la primera evaluación se obtuvo el consentimiento informado de los padres de los participantes, a los que se les proporcionó información sobre la investigación y se les aseguró la confidencialidad y anonimato de los resultados. La segunda instancia de evaluación (retest) se llevó a cabo luego de un intervalo de 5 semanas de la primera (test). En ambas instancias, se lo administró de manera colectiva dentro del horario escolar, en el aula de cursada habitual para cada grado y sección, y en presencia del docente a cargo del grupo.

Luego de la evaluación particularizada de los protocolos individuales se confeccionó una base de datos con los puntajes observados en ambas instancias de evaluación, que fueron analizados con la versión 17.0 del Paquete Estadístico para las Ciencias Sociales SPSS. Los análisis se

realizaron sobre los puntajes observados de la muestra total y de los grupos según grado escolar. En primer lugar, se analizaron los estadísticos descriptivos correspondientes a las medidas de tendencia central y variabilidad. Luego se realizaron las pruebas de normalidad y de homocedasticidad. Finalmente, se estimaron las correlaciones entre los puntajes observados en ambas instancias de evaluación.

3.2. Resultados

En primer lugar, se informan los resultados de los estadísticos descriptivos correspondientes a las medidas de tendencia central y variabilidad de los puntajes observados en las instancias de test y de retest, para todo el grupo examinado y para cada grupo según grado escolar. A continuación, se reportan los resultados de las pruebas de normalidad y de homocedasticidad, necesarias para seleccionar el procedimiento estadístico adecuado para el estudio de las correlaciones entre los puntajes observados en ambas instancias de evaluación. Finalmente, se proporcionan los coeficientes de correlación encontrados para toda la muestra y para cada grado escolar.

3.2.1. Estadísticos descriptivos

En la tabla 3.1 se presentan la media, el desvío estándar y la varianza de los puntajes observados en las instancias de test y de retest en la muestra total.

Tabla 3.1

Media, desvío estándar y varianza de los puntajes observados en el test y el retest en la muestra total

Aplicaciones	N	Media	DE	Varianza
Test	218	31.1	14.0	195
Retest	218	36.8	14.2	202

DE: desvío estándar

Asimismo se calcularon los mismos estadísticos para los grupos segmentados por los grados escolares considerados (Tabla 3.2).

Tabla 3.2

Media, desvío estándar y varianza de los puntajes observados en el test y el retest por grado escolar

Grado	Instancia	N	Media	DE	Varianza
3°	Test	56	16.3	8.19	67.1
	Retest	56	22.0	8.66	74.9
4°	Test	54	28.8	9.95	99.0
	Retest	54	35.4	11.8	138
5°	Test	72	38.4	9.71	94.3
	Retest	72	42.5	9.13	83.4
6°	Test	36	43.1	12.0	144
	Retest	36	50.2	12.3	151

3.2.2. Pruebas de normalidad y homocedasticidad

La prueba de normalidad de los puntajes observados de toda la muestra (prueba de Shapiro-Wilk) reveló que los mismos no se ajustaron a una distribución normal en ambas instancias de evaluación [test: $W(218) = .99$, $p = .037$; retest: $W(218) = .98$, $p = .013$]. La prueba de Levene mostró homogeneidad de varianzas entre los puntajes observados de toda la muestra en ambas instancias de evaluación [$F(1, 434) = .089$, $p = .77$].

En cuanto a las distribuciones de los puntajes observados en cada grado escolar, la prueba de normalidad (Shapiro-Wilk) reveló que, a excepción de 4° grado en el que los puntajes observados en ambas instancias de evaluación se ajustaron a una distribución normal, en el resto de los grados lo hicieron en sólo una de las instancias [3° grado-test: $W(56) = .94$, $p = .008$; 3° grado-retest: $W(56) = .97$, $p = .14$; 4° grado-test: $W(54) = .98$, $p = .54$; 4° grado-retest: $W(54) = .98$, $p = .46$; 5° grado-test: $W(72) = .96$, $p = .044$; 5° grado-retest: $W(72) = .97$, $p = .062$; 6° grado-test: $W(36) = .97$, $p = .49$; 6° grado-retest: $W(36) = .89$, $p = .002$]. Por otro lado, la prueba de Levene mostró que para todos los grados, la varianza de ambas variables fue homogénea [3° grado: $F(1, 110) = 1.94$, $p = .17$; 4° grado: $F(1, 106) = 1.60$, $p = .21$; 5° grado: $F(1, 142) = .12$, $p = .73$; 6° grado: $F(1, 70) = .001$, $p = .97$].

3.2.3. Coeficiente de confiabilidad test-retest

Con el objeto de estimar la estabilidad temporal de los puntajes observados en la muestra total y habida cuenta de que ninguna de las dos variables se ajustó a una distribución normal aunque mostraron homogeneidad de varianzas, se calculó el coeficiente *tau*-b de Kendall por rangos ordenados (ver apartado A.1 del Anexo A “Análisis no paramétricos”). No obstante, dado que los resultados de esta prueba fueron congruentes con los de la correspondiente prueba paramétrica, a continuación se reportan estos últimos.

Se calculó el coeficiente de correlación *r* de Pearson, que reveló una correlación positiva significativa entre ambas variables [$r(218) = .90, p < .001$]. De acuerdo con la sugerencia de Cohen (1988, 1992) para los coeficientes de correlación, los valores críticos para considerar tamaños del efecto de magnitud pequeña, mediana y grande, respectivamente, deberían ser de $r = .10, = .30$ y $= .50$. En virtud de dicha sugerencia, afirmamos que la fuerza de la asociación entre los puntajes observados en ambas instancias de evaluación fue grande. Planteado en términos más precisos y objetivos, ambas variables compartieron el 80.5% de la varianza ($R^2 = .80$).

Para el estudio de las correlaciones en cada grado escolar, se calculó el coeficiente *tau*-b de Kendall por rangos ordenados ya que, pese a que las varianzas resultaron homogéneas, los puntajes observados no se ajustaron a una distribución normal en todos los casos (ver apartado A.1 del Anexo A “Análisis no paramétricos”). Al igual que en el análisis de los puntajes observados en la muestra completa, los resultados de esta prueba fueron congruentes con los de la correspondiente prueba paramétrica, que se informa a continuación.

Nuevamente, se obtuvieron coeficientes de correlación positivos estadísticamente significativos con un tamaño del efecto de magnitud grande entre los puntajes observados en ambas instancias de evaluación en todos

los grados [3° grado: $r(56) = .73, p < .001$; 4°: $r(54) = .80, p < .001$; 5°: $r(72) = .75, p < .001$; 6°: $r(36) = .93, p < .001$].

3.3. Discusión y conclusiones

En razón de la magnitud del tamaño del efecto de los coeficientes de correlación obtenidos para los puntajes del TECLE Buenos Aires observados en la muestra total y en los grupos según grado escolar, concluimos que los mismos resultaron confiables. Específicamente, los puntajes observados en el test mostraron un alto grado de estabilidad temporal medida a través del método test-retest con un intervalo de 5 semanas entre ambas instancias de evaluación, que arroja un coeficiente de estabilidad $r = .90$.

Capítulo 4

Estudios de validez del TECLE Buenos Aires

Los estudios de validez constituyen otro de los requisitos fundamentales que aparecen entre las directrices de la Comisión Internacional de Tests para el desarrollo y uso de instrumentos (International Test Commission, 2017; Muñiz, Elosua, & Hambleton, 2013). Al respecto señalan la necesidad de proporcionar la evidencia que la respalde cuando se trata de su aplicación a poblaciones diferentes a las que estaba destinada la prueba original.

La validez refiere al grado en que las *interpretaciones* y los *usos propuestos de los puntajes observados* en un test se apoyan en la evidencia empírica y en la teoría (Asociación Americana de Investigación en Educación, Asociación Psicológica Americana, & Consejo Nacional de Medición en Educación, 1999, citado en Furr & Bacharach, 2014). Dicho de otro modo, un test es válido si mide lo que pretende medir y si sirve al propósito para el que es utilizado (Cronbach, 1988). Tradicionalmente se consideran tres tipos de evidencias de validez: de contenido, de criterio y de constructo.

Las evidencias de *validez de contenido* se obtienen a partir de la elaboración o el muestreo de un conjunto de ítems que según el juicio de expertos sea pertinente y representativo del constructo que se pretende medir. Si, adicionalmente, el contenido del test es evaluado como pertinente por sujetos no expertos (por ej., sujetos pertenecientes a la población a la que está dirigido el test, o posibles destinatarios del reporte de los resultados) se obtienen evidencias de validez de *facie* o *aparente*.

Las evidencias de *validez de criterio* refieren al grado de asociación entre los puntajes observados en el test y una variable criterio. En los estudios que se proponen evaluar la eficacia de los puntajes observados en el test para predecir una variable criterio medida en una instancia posterior, se obtienen evidencias de validez predictiva. En cambio, si el objetivo es

utilizar estos puntajes como una alternativa a otra medida que ha demostrado contrastar grupos de sujetos que se diferencian por su nivel de dominio en el atributo de interés, se obtienen evidencias de validez *concurrente*. Generalmente, el test se propone como una alternativa a una medida que demanda más recursos de tiempo y/o de capacitación del examinador, como es el caso de la evaluación basada en pruebas formales o en la experiencia profesional para arribar a un diagnóstico clínico. Las evidencias de validez de criterio, así como las de validez de contenido, también contribuyen a la validez de constructo.

Las evidencias de *validez de constructo* refieren al grado en que los puntajes observados en un test pueden interpretarse como una medida del atributo psicológico que se pretende medir. Este tipo de evidencias de validez es el más relevante en términos de la calidad psicométrica de un instrumento, ya que las interpretaciones de los puntajes observados fundamentan los usos propuestos del test y, en consecuencia, motivan decisiones que impactan en los sujetos evaluados y en la población a la que representan. Los métodos para obtener este tipo de evidencias de validez son variados y todos apuntan a revelar distintos aspectos teóricos del constructo. Entre dichos métodos se incluyen el estudio de la estructura interna del test, el de los cambios en el rendimiento como consecuencia de intervenciones experimentales y/o de los procesos del desarrollo. Cabe una mención especial al análisis de las asociaciones entre los puntajes observados y distintas variables teóricamente relacionadas y no relacionadas, que aportan, respectivamente, evidencias de *validez convergente* y *discriminante*. Este tipo de evidencias de validez es particularmente importante porque pone a prueba las hipótesis referidas a la red nomológica de la que el constructo forma parte.

Una *red nomológica* es un sistema de leyes estadísticas o deterministas que relacionan constructos teóricos entre sí y con sus propiedades observables (Cronbach & Meehl, 1955). Furr y Bacharach (2014) señalaron que existen dos enfoques para los estudios de validez. El

enfoque tradicional es de tipo deductivo, se centra en las evidencias a favor de un conjunto de hipótesis referidas al rendimiento en los tests en relación con una teoría completamente formada, y es más adecuado para contextos aplicados. El enfoque alternativo, en cambio, es de tipo inductivo, enfatiza la exploración de las relaciones entre el constructo medido por el test y potenciales variables relevantes, y resulta más adecuado para contextos de investigación. En ambos enfoques, se estudia el ajuste entre correlaciones esperadas y observadas. La constatación de las correlaciones esperadas provee evidencias a favor de la validez de las interpretaciones teóricas de los puntajes observados en el test. Por otro lado, la falta de ajuste entre las correlaciones esperadas y observadas deben conducir a un examen de los aspectos metodológicos (por ej., el tamaño de la muestra, la selección de los procedimientos estadísticos) o conceptuales del estudio (por ej., la operacionalización de los constructos cuya correlación se pretende estudiar, la precisión en la definición teórica del constructo), y en última instancia, a una revisión crítica de las leyes que componen la red nomológica.

Tal como se mencionó en el capítulo 2 “El Test Colectivo de Eficacia Lectora. Su adaptación”, el TECLE Buenos Aires pretende medir el constructo “eficacia lectora”. Dado que se trata de un test de *screening*, esta herramienta no se propone determinar el grado de dominio que poseen los sujetos evaluados en cada una de las habilidades específicas y no específicas que contribuyen a la lectura, sino ofrecer una estimación global de la capacidad de lectura. No obstante, es necesario tener presente cuáles son los factores que contribuyen a la eficacia lectora a fin de representar la red nomológica de la que el constructo forma parte.

Los aspectos específicos para la lectura que contribuyen a la eficacia lectora son la precisión y la velocidad en el reconocimiento de las palabras escritas (decodificación). Los aspectos no específicos incluyen habilidades lingüísticas generales, sintácticas y semánticas, necesarias para la comprensión de oraciones, y habilidades cognitivas generales que involucran la administración de los recursos de procesamiento para llevar a

cabo la tarea en el tiempo permitido. Dado que el test posee características de los tests de velocidad (administración con tiempo límite y facilidad de los ítems), la velocidad de respuesta o de procesamiento también contribuye al rendimiento en el test, independientemente, en parte, de la contribución de la velocidad en el reconocimiento léxico.

En el presente capítulo, se presentan los resultados de dos estudios: el primero estuvo dirigido a la obtención de evidencias de validez de constructo y el segundo, a la obtención de evidencias de validez de criterio.

4.1. Estudio de validez de constructo

El objetivo de este estudio fue explorar las asociaciones convergentes y discriminantes entre el rendimiento en el TECLE Buenos Aires y otras medidas teóricamente relacionadas.

La hipótesis de trabajo fue que los puntajes observados en el test mostrarán asociaciones más fuertes con las medidas que dependen de las habilidades específicas para la lectura, es decir, con las medidas de decodificación, que con las medidas que dependen de habilidades generales, lingüísticas y cognitivas.

4.1.1. Metodología

Se ajusta a los requerimientos de los estudios de validez de constructo que suponen la correlación de los puntajes del test con el de otras medidas relacionadas con la lectura.

4.1.1.1. Participantes

Para este estudio se conformó una muestra por conveniencia que incluyó a 57 sujetos de 3° a 6° grados de primaria, 26 niñas y 31 varones, provenientes de la Ciudad Autónoma de Buenos Aires y alrededores. Los criterios de selección fueron: no poseer trastornos intelectuales generales, motores o sensoriales no corregidos, diagnóstico de trastornos en la

adquisición del lenguaje o en el aprendizaje escolar, no haber repetido grados escolares, y haber adquirido el español como lengua materna.

4.1.1.2. Instrumentos

Las medidas utilizadas para obtener evidencias de validez convergente y discriminante respecto de las puntuaciones del TECLE Buenos Aires fueron las siguientes:

- La precisión y el tiempo de ejecución en una prueba de *screening* del Test Leer para Comprender.
- La precisión y el tiempo en la lectura en voz alta de palabras y de pseudopalabras y el rendimiento en la prueba de comprensión palabras y frases del test LEE.
- El rendimiento en la prueba de comprensión de vocabulario de la escala Alfa-2.
- El tiempo en la denominación automatizada rápida de dibujos.
- El tiempo de ejecución en las partes 1 y 2 de una prueba de trazado de senderos.

En función de ello se incluyeron los instrumentos que se detallan a continuación.

➤ ***TECLE Buenos Aires***

Referimos al lector al capítulo 2 para una descripción detallada de esta prueba.

➤ ***Prueba de screening sobre el texto expositivo “Un almuerzo prehistórico” extraído del Test Leer para Comprender (Abusamra, Ferreres, Raiter, De Beni, & Cornoldi, 2010).***

El Test Leer para Comprender (en adelante, TLC) es una herramienta exhaustiva para la evaluación de la comprensión de textos, basada en un modelo multicomponente de dicha habilidad académica (De Beni, Cornoldi, Carretti, & Meneghetti, 2003). Fue elaborado en Buenos Aires (Argentina) y

proporciona información sobre sus propiedades psicométricas y los datos normativos correspondientes.

El test está conformado por 11 pruebas de administración individual o colectiva; cada una de ellas arroja una medida del nivel de dominio en una de las 11 áreas consideradas por el modelo, tanto en términos de precisión como del tiempo de ejecución de la prueba. Para cada prueba, los autores presentaron los datos normativos disponibles para alumnos de 5° y 6° grados de educación primaria de la Ciudad Autónoma de Buenos Aires y alrededores, y de 7° grado o 1° año de Escuela Secundaria Básica (ESB), según se trate, respectivamente, de alumnos de la Ciudad Autónoma de Buenos Aires o del Gran Buenos Aires.

Además de las pruebas destinadas a evaluar las áreas del modelo multicomponente, el test incluye cuatro pruebas de *screening*, de administración individual o colectiva, que examinan, cada una, la comprensión de un texto: uno informativo y uno narrativo seleccionados para alumnos de 5° grado, y uno informativo y uno narrativo para alumnos de 6° y 7° grados (o 1° año de ESB). En cada prueba de *screening* se presenta en primer lugar el texto, seguido por 10 preguntas que el alumno debe responder seleccionando la respuesta correcta entre cuatro opciones. Cada respuesta correcta suma un punto, de modo que el puntaje máximo es de 10 puntos. Al igual que en las pruebas destinadas a evaluar las áreas del modelo, en las pruebas de *screening*, el evaluador debe registrar el tiempo de ejecución en minutos, ya que se reportan los datos normativos tanto en términos de precisión como de tiempo de ejecución de la prueba. Para este estudio, las variables derivadas del desempeño en esta prueba fueron el puntaje observado y el tiempo de ejecución en el TLC. La prueba se presenta en el anexo D “Instrumentos administrados en el estudio de validez de constructo”.

➤ ***Pruebas del Test de Lectura y Escritura en Español LEE (Defior Citoler, Fonseca, Gottheil, Aldrey, Jiménez Fernández, Pujals, & Serrano Chica, 2006)***

El Test de Lectura y Escritura en español LEE (en adelante, test LEE) es una batería de pruebas, de administración individual, dirigida a evaluar los procesos implicados en la lectura y la escritura en español. Los autores presentaron información sobre sus propiedades psicométricas de validez y confiabilidad y datos normativos aplicables a alumnos de 1° a 4° grados de educación primaria.

Para este estudio se administraron las pruebas de lectura en voz alta de palabras y de pseudopalabras y de comprensión de palabras y frases.

✓ ***Prueba de lectura en voz alta de palabras.*** Esta prueba está dirigida a evaluar los procesos léxicos y subléxicos de la lectura de palabras, así como la fluidez o velocidad lectora, de modo que ofrece una medida del dominio en los aspectos específicos de la lectura (decodificación).

La tarea consiste en leer una lista de 42 palabras de la manera más precisa y rápida posible. La lista incluye palabras de 2, 3, 4 y 6 sílabas, con y sin estructura silábica compleja, con y sin tilde, con reglas de conversión grafema-fonema de complejidad variable (reglas simples, inconsistentes, dependientes del contexto y el caso de la letra "H"). Aquellas palabras leídas de manera precisa y fluida suman 2 puntos cada una; las palabras leídas de manera precisa, pero silabeada o luego de una autocorrección o pausa prolongada suman 1 punto cada una; las palabras leídas de manera incorrecta o con errores de acentuación no suman puntos. Por lo tanto, el puntaje máximo en precisión es de 84 puntos. Por otro lado, el evaluador debe registrar el tiempo de ejecución de la tarea, contabilizando la cantidad de segundos transcurridos entre el inicio de la primera palabra de la lista y la finalización de la última. Para este estudio, las variables derivadas del rendimiento en esta prueba fueron la precisión y el tiempo en la lectura de palabras.

✓ **Prueba de lectura en voz alta de pseudopalabras.** Esta prueba se propone evaluar la precisión y la velocidad de los procesos subléxicos de la lectura sin el apoyo del conocimiento léxico, es decir que apunta de manera más específica al dominio de las habilidades de segmentación de la cadena de grafemas, al conocimiento de las reglas de conversión grafema-fonema y al dominio de las habilidades de ensamblaje fonológico.

La tarea consiste en leer 42 pseudopalabras de la manera más precisa y rápida posible. Al igual que para las palabras, en la lista se incluyen pseudopalabras de longitud y complejidad ortográfica semejante. La puntuación sigue los mismos criterios que para las palabras, de modo que el puntaje máximo en precisión es de 84 puntos. Asimismo, el evaluador debe contabilizar el tiempo en segundos transcurridos entre el inicio de la primer pseudopalabra de la lista y la finalización de la última. Para este estudio, las variables derivadas del rendimiento en esta prueba fueron la precisión y el tiempo en la lectura de pseudopalabras.

✓ **Prueba de comprensión de palabras y frases.** Esta prueba está dirigida a evaluar los procesos morfosintácticos involucrados en la comprensión de palabras y de frases. Está compuesta por las siguientes 4 subpruebas: frases, familia de palabras, preguntas, y completar oraciones. La suma de los puntajes parciales observados en las mismas da el puntaje total para esta prueba, cuyo rango de puntuación es de 0 a 43 puntos. Para este estudio, la variable derivada del rendimiento en esta prueba fue el puntaje observado en la comprensión de palabras y frases.

La subprueba de frases está compuesta por 5 ítems, que suman cada uno 1 punto. El primer y segundo ítems consisten en dos oraciones escritas, una en voz activa y otra en voz pasiva, que el niño debe leer y comprender para poder responder a una pregunta referida al agente de la acción. Los ítems 3 a 5 consisten en instrucciones escritas que el niño debe leer y ejecutar.

La subprueba de familia de palabras apunta a evaluar el conocimiento morfológico de las palabras. Está compuesta por 5 ítems (1 de ensayo y 4 de prueba) que consisten cada uno en una palabra de referencia y 7 palabras, de las cuales 4 pertenecen a la familia de la palabra de referencia (comparten la raíz). La tarea consiste en señalar aquéllas pertenecientes a la familia, y se asigna 1 punto por cada palabra señalada y no señalada correctamente, de modo que el puntaje máximo es de 28 puntos.

La subprueba de preguntas se propone evaluar los procesos sintácticos que intervienen en la comprensión de oraciones. Consiste en 2 oraciones escritas que contienen las mismas palabras, pero en distinto orden, y la tarea consiste en responder 3 preguntas por cada oración. Cada respuesta correcta suma 1 punto; por lo tanto, el puntaje máximo es igual a 6 puntos.

Por último, el objetivo de la subprueba de completar es evaluar los procesos de coherencia semántica y cohesión sintáctica que participan en la comprensión de oraciones. Incluye 4 oraciones escritas incompletas seguidas de tres opciones escritas como finales posibles. La tarea consiste en seleccionar el final correcto para cada oración incompleta y cada respuesta correcta suma 1 punto.

➤ ***Prueba de comprensión de vocabulario de la escala Alfa-2 (Feldman, 2009)***

Esta prueba se administró con el objeto de explorar la asociación entre el rendimiento en el TECLE Buenos Aires y el nivel de vocabulario. El tipo de conocimiento involucrado en esta prueba constituye una habilidad lingüística general, no específica de la lectura, pero necesaria para la comprensión de las oraciones y la selección de la palabra correcta a fin de poder ejecutar la tarea.

En ella se presentan 28 láminas que incluyen, cada una, cuatro dibujos. Junto con cada lámina, el evaluador emite una palabra, y la tarea consiste en señalar el dibujo correcto. Cada respuesta correcta suma 1 punto, de

modo que el puntaje máximo equivale a 28 puntos, que se transforman en un puntaje de edad madurativa. La prueba ofrece una tabla de edades madurativas de 4 a 17 años. Sin embargo, para este estudio, la variable derivada del rendimiento en esta prueba fue el puntaje observado en la comprensión de vocabulario.

➤ ***Prueba de denominación automatizada rápida de dibujos del test DAR (Morello García, China, & Ferreres, 2015)***

El Test de denominación automatizada rápida (DAR) es una herramienta recientemente desarrollada en el medio local que se propone medir la velocidad de acceso a la etiqueta léxica. Siguiendo la estructura del test RAN (*rapid automatized naming*) diseñado por Denckla y Rudel (1974), el test DAR está conformado por 4 láminas en las que se presentan 5 elementos familiares repetidos 10 veces en un orden pseudoaleatorio para su denominación rápida. Cada lámina evalúa la denominación de un tipo de elemento: letras, dígitos, dibujos y colores. El puntaje obtenido está dado por el tiempo en segundos transcurridos entre el inicio de la denominación del primer elemento y la finalización del último.

Previamente a la administración de cada lámina, el test incluye una prueba de familiaridad de los elementos a través de la denominación por confrontación de cada uno de ellos, seguida de un ensayo de práctica de la denominación automatizada rápida de los 5 elementos de prueba repetidos 3 veces y presentados en un orden pseudoaleatorio. Finalmente, se da comienzo al ensayo de prueba.

Para este estudio se utilizó la lámina de dibujos ya que se consideró que la recuperación léxica de los nombres de los dibujos comparte con la lectura de frases y de textos la característica de no limitarse a un campo semántico.

Esta prueba se administró con el objeto de explorar la asociación entre el rendimiento en el TECLE Buenos Aires y una medida de la velocidad de acceso a la etiqueta léxica. La variable derivada del rendimiento en esta

prueba fue el tiempo en la denominación automatizada rápida de dibujos. La lámina correspondiente se presenta en el anexo D.

➤ **Prueba de trazado de senderos**

Con el objeto de explorar la asociación entre el rendimiento en el TECLE Buenos Aires y una medida general de funcionamiento ejecutivo y velocidad de procesamiento, se diseñó una prueba ad hoc de trazado de senderos en colores (China, no publicado).

La prueba está inspirada en el instrumento Children's Color Trails (Williams, Rickert, Hogan, Zolten, Satz, D'Elia, Asarnow, Zaucha, & Light, 1995), no disponible en el medio local. El test Children's Color Trails fue desarrollado con el fin de minimizar los sesgos culturales que impone el Trail Making Test (Partington & Leiter, 1949, citado en Strauss, Sherman & Spreen, 2006), más específicamente su parte B. En esta parte del test, el examinado debe unir en el orden correcto los números del 1 al 13 y las letras de la A a la L, distribuidos de manera aleatoria en una página, alternando los miembros de estas categorías (es decir, en la secuencia 1, A, 2, B, ..., 12, L, 13). De acuerdo con los autores del test Children's Color Trails (Williams et al., 1995), la segunda parte del mismo, a diferencia de la parte B del Trail Making Test, puede ser utilizada en niños no familiarizados con el alfabeto romano, no alfabetizados o con trastornos en el aprendizaje de la lectura. Al igual que el Trail Making Test, consta de dos partes. En la primera, el examinado debe unir en el orden correcto los números 1 al 15; los números impares se presentan dentro de círculos rosados y los pares dentro de círculos amarillos. En la segunda parte, a partir del número 2, cada número se presenta dos veces, dentro de un círculo rosado y dentro de uno amarillo, y el niño debe unir en el orden correcto los números del 1 al 15, alternando los colores (es decir, en la secuencia 1 rosado, 2 amarillo, 3 rosado, 4 amarillo, ..., 15 rosado). Ambas partes del test arrojan una medida de atención, búsqueda visual, habilidades de secuenciación y velocidad de procesamiento, pero la segunda parte evalúa además la flexibilidad mental,

atención dividida, o inhibición de la interferencia, es decir, funcionamiento ejecutivo.

La prueba de trazado de senderos utilizada para este estudio fue diseñada siguiendo los lineamientos generales del test Children's Color Trails, que se describen a continuación. La prueba consta de dos partes: la ejecución de la primera parte no se ve afectada por la interferencia del color, mientras que la segunda, sí. Ambas partes están precedidas por un ensayo de práctica para verificar la comprensión de la consigna, con un número reducido de elementos (los números del 1 al 8). En ambas partes, el niño debe unir 15 elementos (los números del 1 al 15), de los cuales están señalados el elemento inicial y el final, y el trazado de los senderos no resulta en la intersección de líneas. Finalmente, ambos tests arrojan dos medidas: el tiempo de ejecución en segundos de cada parte. Por otro lado, la prueba desarrollada para este estudio se diferencia del test original en la disposición de los elementos en la hoja y en los colores empleados: se utilizaron círculos azules en lugar de rosados y círculos rojos en lugar de amarillos. Además, en el presente estudio no se contabilizaron medidas propuestas en el test original como el número de errores y de facilitaciones por parte del evaluador, ni el índice de interferencia dado por la siguiente fórmula: $(\text{tiempo en la parte 2} - \text{tiempo en la parte 1}) / \text{tiempo en la parte 1}$. Por lo tanto, para este estudio, las variables derivadas del rendimiento en esta prueba fueron el tiempo en la parte 1 y el tiempo en la parte 2 del trazado de senderos. En el anexo D se presentan las láminas correspondientes.

4.1.1.3. Procedimientos

Previamente a la evaluación de cada niño, se obtuvo el consentimiento informado por escrito de uno de los padres, se les proporcionó información sobre la investigación, asegurándoles la confidencialidad y anonimato de los datos.

A continuación, se administraron las pruebas en el mismo orden para todos los participantes según el siguiente detalle: el TECLE Buenos Aires, las pruebas de lectura en voz alta del test LEE, el *screening* del TLC, la prueba de comprensión de palabras y frases del test LEE, la prueba de comprensión de vocabulario de la escala Alfa-2, la denominación automatizada rápida de dibujos, y las partes 1 y 2 de la prueba de trazado de senderos. Las pruebas se administraron en una única sesión individual de 30 a 60 minutos de duración y ninguno de los participantes requirió interrumpir la evaluación.

Finalmente, se confeccionó una base de datos con los desempeños observados, que fueron analizados con la versión 17.0 del Paquete Estadístico para las Ciencias Sociales SPSS. Se analizaron los estadísticos descriptivos correspondientes a las medidas de tendencia central y variabilidad del rendimiento en el TECLE Buenos Aires y demás variables de interés. Seguidamente, se realizaron las pruebas de normalidad con el fin de seleccionar el procedimiento estadístico adecuado para el estudio de las correlaciones entre el desempeño en el TECLE Buenos Aires y en las otras medidas. Por último, se calcularon los coeficientes de correlación.

4.1.2. Resultados

4.1.2.1. Estadísticos descriptivos

En la tabla 4.1 se presentan la media, el desvío estándar y la varianza de los puntajes observados en el TECLE Buenos Aires y de las 11 variables consideradas: la precisión y el tiempo de ejecución en la prueba de *screening* del TLC, la precisión y el tiempo en la lectura en voz alta de palabras y de pseudopalabras del test LEE, el desempeño en la comprensión de palabras y frases del mismo test, el rendimiento en la prueba de comprensión de vocabulario del test Alfa-2, el tiempo en la denominación automatizada rápida de dibujos, y el tiempo en la ejecución de cada una de las partes 1 y 2 de la prueba de trazado de senderos.

Tabla 4.1

Media, desvío estándar y varianza de los puntajes observados en el TECLE Buenos Aires y de medidas teóricamente relacionadas

Medidas	N	Media	DE	Varianza
TECLE	57	43.0	11.0	120
TLC	57	8.40	1.44	2.07
T-TLC	57	498	270	72767
LVA pal	57	72.0	8.36	69.9
T-LVA pal	57	45.3	25.6	657
LVA pseudo	57	61.0	11.5	132
T-LVA pseudo	57	62.6	28.6	820
Comp LEE	57	39.0	2.91	8.45
Alfa-2	57	19.5	5.37	28.8
DAR	57	41.2	9.54	91.1
Send 1	57	20.1	7.13	50.9
Send 2	57	45.5	13.5	181

DE: desvío estándar; TECLE: TECLE Buenos Aires; TLC: precisión en la prueba de *screening* del TLC; T-TLC: tiempo en la prueba de *screening* del TLC (segs.); LVA pal: precisión en la prueba de lectura en voz alta de palabras del test LEE; T-LVA pal: tiempo en la prueba lectura en voz alta de palabras del test LEE (segs.); LVA pseudo: precisión en la prueba de lectura en voz alta de pseudopalabras del test LEE; T-LVA pseudo: tiempo en la prueba de lectura en voz alta de pseudopalabras del test LEE (segs.); Comp LEE: rendimiento en la prueba de comprensión de palabras y frases del test LEE; Alfa-2: rendimiento en la prueba de comprensión de vocabulario de la escala Alfa-2; DAR: tiempo en la denominación automatizada rápida de dibujos (segs.); Send 1: tiempo en la parte 1 del test de trazado de senderos (segs.); Send 2: tiempo en la parte 2 del test de trazado de senderos (segs.).

4.1.2.2. Prueba de normalidad de las distribuciones

La prueba de normalidad (Shapiro-Wilk) aplicada a las 12 variables consideradas reveló que solo 4 mostraron distribuciones que se ajustaron a una distribución normal: el rendimiento en el TECLE Buenos Aires, la precisión en la lectura en voz alta de pseudopalabras, y el tiempo de ejecución de las partes 1 y 2 de la prueba de trazado de senderos [TECLE: $W(57) = .98$, $p = .661$; TLC: $W(57) = .86$, $p < .001$; T-TLC: $W(57) = .76$, $p < .001$; LVA pal: $W(57) = .95$, $p = .013$; T-LVA pal: $W(57) = .68$, $p < .001$; LVA pseudo: $W(57) = .96$, $p = .074$; T-LVA pseudo: $W(57) = .75$, $p < .001$; Comp. LEE: $W(57) = .90$, $p < .001$; Alfa-2: $W(57) = .91$, $p < .001$; DAR: $W(57) = .87$, $p < .001$; Send 1: $W(57) = .96$, $p = .073$; Send 2: $W(57) = .97$, $p = .13$].

4.1.2.3. Análisis correlacionales entre los puntajes observados en el TECLE Buenos Aires y otras medidas relacionadas

En virtud de los resultados obtenidos en la prueba de normalidad reportados en el apartado precedente, se utilizó el coeficiente de correlación de Spearman, rho (ρ), para analizar las correlaciones bivariadas entre el rendimiento en el TECLE Buenos Aires de la muestra completa y las medidas basadas en el número de aciertos, por un lado, y las del tiempo de ejecución, por otro (Tabla 4.2).

Como puede verse, todas las correlaciones alcanzaron significación estadística. Sin embargo, es necesario realizar algunas distinciones analíticas. Se observaron diferencias en la fuerza de las asociaciones entre el rendimiento en el TECLE Buenos Aires y las variables relacionadas.

Las variables que mostraron asociaciones fuertes con el rendimiento en el TECLE Buenos Aires fueron, en orden decreciente, el tiempo en la lectura en voz alta de palabras, el tiempo en la prueba de *screening* de comprensión de textos, la precisión en la lectura en voz alta de palabras, el tiempo en la lectura en voz alta de pseudopalabras, y la precisión en esta última.

El tiempo en la denominación automatizada rápida de dibujos, la precisión en la prueba de *screening* de comprensión de textos, el tiempo en la parte 2 de la prueba de trazado de senderos (con interferencia) y el rendimiento en la comprensión de palabras y frases, mostraron, en orden decreciente, asociaciones moderadas. Finalmente, el rendimiento en el TECLE Buenos Aires mostró las asociaciones más débiles con el tiempo en la parte 1 de la prueba de trazado de senderos (sin interferencia) y el rendimiento en comprensión de vocabulario.

Tabla 4.2

Correlaciones (coeficientes rho de Spearman) entre el rendimiento en el TECLE Buenos Aires y en 11 medidas de interés de la muestra completa

Medidas	Medida	N	Coefficiente rho	Sig.
Medidas basadas en el número de aciertos	TLC	57	.51**	< .001
	LVA pal	57	.72**	< .001
	LVA pseudo	57	.68**	< .001
	Comp LEE	57	.42**	.001
	Alfa-2	57	.32*	.016
Medidas del tiempo de ejecución	T-TLC	57	-.74**	< .001
	T-LVA pal	57	-.80**	< .001
	T-LVA pseudo	57	-.74**	< .001
	DAR	57	-.59**	< .001
	Send 1	57	-.35**	.007
	Send 2	57	-.43**	.001

** . La correlación es significativa en el nivel de .01.

* . La correlación es significativa en el nivel de .05.

TLC: precisión en la prueba de *screening* del TLC; LVA pal: precisión en la prueba de lectura en voz alta de palabras del test LEE; LVA pseudo: precisión en la prueba de lectura en voz alta de pseudopalabras del test LEE; Comp LEE: rendimiento en la prueba de comprensión de palabras y frases del test LEE; Alfa-2: rendimiento en la prueba de comprensión de vocabulario de la escala Alfa-2; T-TLC: tiempo en la prueba de *screening* del TLC (segs.); T-LVA pal: tiempo en la prueba de lectura en voz alta de palabras del test LEE (segs.); T-LVA pseudo: tiempo en la prueba de lectura en voz alta de pseudopalabras del test LEE (segs.); DAR: tiempo en la denominación automatizada rápida de dibujos (segs.); Send 1: tiempo en la parte 1 de la prueba de trazado de senderos (segs.); Send 2: tiempo en la parte 2 de la prueba de trazado de senderos (segs.).

Adicionalmente, con el objeto de explorar las posibles diferencias en las correlaciones estudiadas en función del grado escolar, se segmentó la muestra en dos grupos: uno conformado por los participantes de 3° y 4° grados, y otro por los de 5° y 6°. En la tabla 4.3 y 4.4 se presentan, respectivamente, las correlaciones bivariadas entre el rendimiento en el TECLE Buenos Aires y las medidas basadas en el número de aciertos, por un lado, y las del tiempo de ejecución, por otro, en ambos grupos de participantes.

Al segmentar la muestra en dos grupos, se observa que las correlaciones entre el rendimiento en el TECLE Buenos Aires y la precisión y el tiempo en las pruebas de lectura en voz alta de palabras y de pseudopalabras, el tiempo en la prueba de *screening* de comprensión de

textos y el tiempo en la denominación automatizada rápida alcanzaron significación estadística en ambos grupos, mientras que la correlación entre el rendimiento en el TECLE Buenos Aires y la precisión en la prueba de *screening* de comprensión de textos fue estadísticamente significativa sólo en el grupo de 3°-4° grados (Tablas 4.3 y 4.4).

Tabla 4.3

Correlaciones (coeficientes rho de Spearman) entre el rendimiento en el TECLE Buenos Aires y en las medidas basadas en el número de aciertos en los grupos de 3°-4° y 5°-6° grados

Grupo	Medida	N	Coefficiente rho	Sig.
3°-4°	TLC	28	.46 [*]	.015
	LVA pal	28	.78 ^{**}	< .001
	LVA pseudo	28	.66 ^{**}	< .001
	Comp LEE	28	.26	.18
	Alfa-2	28	.17	.39
5°-6°	TLC	29	.33	.078
	LVA pal	29	.50 ^{**}	.006
	LVA pseudo	29	.51 ^{**}	.005
	Comp LEE	29	.20	.31
	Alfa-2	29	.029	.88

** . La correlación es significativa en el nivel de .01.

* . La correlación es significativa en el nivel de .05.

TLC: precisión en la prueba de *screening* del TLC; LVA pal: precisión en la prueba de lectura en voz alta de palabras del test LEE; LVA pseudo: precisión en la prueba de lectura en voz alta de pseudopalabras del test LEE; Comp LEE: rendimiento en la prueba de comprensión de palabras y frases del test LEE; Alfa-2: rendimiento en la prueba de comprensión de vocabulario de la escala Alfa-2

Por otro lado, las correlaciones entre el rendimiento en el TECLE Buenos Aires y en las pruebas de comprensión de palabras y de frases, comprensión de vocabulario y trazado de senderos sin y con interferencia no alcanzaron significación estadística en ninguno de los grupos (Tablas 4.3 y 4.4).

Finalmente, los coeficientes de correlación entre el rendimiento en el TECLE Buenos Aires y el resto de las variables fueron de mayor magnitud en el grupo de 3°-4° grados que en el grupo de 5°-6°, tanto los que alcanzaron significación estadística como los que no lo hicieron.

Tabla 4.4

Correlaciones (coeficientes rho de Spearman) entre el rendimiento en el TECLE Buenos Aires y en las medidas del tiempo de ejecución en los grupos de 3°-4° y 5°-6° grados

Grupo	Medida	N	Coefficiente rho	Sig.
3°-4°	T-TLC	28	-.79**	< .001
	T-LVA pal	28	-.83**	< .001
	T-LVA pseudo	28	-.74**	< .001
	DAR	28	-.54**	.003
	Send 1	28	-.23	.24
	Send 2	28	-.26	.18
	5°-6°	T-TLC	29	-.55**
T-LVA pal		29	-.75**	< .001
T-LVA pseudo		29	-.71**	< .001
DAR		29	-.45*	.013
Send 1		29	-.20	.30
Send 2		29	-.25	.19

** . La correlación es significativa en el nivel de .01.

* . La correlación es significativa en el nivel de .05.

T-TLC: tiempo en la prueba de *screening* del TLC (segs.); T-LVA pal: tiempo en la prueba de lectura en voz alta de palabras del test LEE (segs.); T-LVA pseudo: tiempo en la prueba de lectura en voz alta de pseudopalabras del test LEE (segs.); DAR: tiempo en la denominación automatizada rápida de dibujos (segs.); Send 1: tiempo en la parte 1 de la prueba de trazado de senderos (segs.); Send 2: tiempo en la parte 2 de la prueba de trazado de senderos (segs.).

4.1.3. Discusión y conclusiones del estudio de validez de constructo

Las asociaciones observadas en la muestra completa entre el rendimiento en el TECLE Buenos Aires y las variables estudiadas apoyan la hipótesis de que la eficacia lectora es un constructo de naturaleza componencial que depende de habilidades específicas para la lectura, esto es, la precisión y la fluidez en la decodificación, así como de habilidades lingüísticas generales de comprensión y velocidad en la recuperación léxica, y habilidades cognitivas generales de velocidad de procesamiento y funcionamiento ejecutivo. Asimismo, aportan evidencia a favor de que el test posee características de un test de velocidad.

En la tabla 4.5, se presentan las medidas estudiadas ordenadas en función de la magnitud de la correlación hallada con el rendimiento en el TECLE Buenos Aires y agrupadas de acuerdo con los procesos que comparten. Aquéllas que dependen de las habilidades específicas para la

lectura, es decir, de la precisión y la velocidad en la decodificación mostraron las asociaciones más fuertes con el rendimiento en el TECLE Buenos Aires. Del resto de las variables, las que involucran habilidades vinculadas más estrechamente con la eficacia lectora desde una perspectiva teórica, como la comprensión de material escrito (precisión en comprensión de textos y rendimiento en comprensión de frases y palabras escritas), la velocidad en la recuperación léxica (tiempo en denominación automatizada rápida de dibujos) y el funcionamiento ejecutivo (tiempo en el trazado de senderos con interferencia), mostraron asociaciones más fuertes que las que dependen de la comprensión del lenguaje oral (rendimiento en comprensión de vocabulario) y la velocidad de procesamiento (tiempo en el trazado de senderos sin interferencia).

Por otro lado, la segmentación de la muestra, que resultó en dos grupos más homogéneos, reveló que las únicas medidas que mantienen una correlación estadísticamente significativa con el rendimiento en el TECLE Buenos Aires, con un tamaño del efecto de magnitud mediana a grande, son las que operacionalizan el constructo de las habilidades de decodificación.

En virtud de los resultados obtenidos, cabe concluir que pese a la naturaleza componencial del constructo “eficacia lectora”, el puntaje obtenido en el TECLE Buenos Aires parece depender *principalmente* de las habilidades de decodificación. Esto podría deberse a que los procesos más relevantes para la selección rápida y precisa de la palabra correcta son el análisis y el reconocimiento ortográfico dadas las características de los distractores: son ortográficamente similares al blanco, dos son pseudopalabras, y la mayoría de las palabras no está semánticamente relacionada con el blanco, no comparten la categoría gramatical, o bien, no concuerdan en género y/o número.

Tabla 4.5

Correlaciones (coeficientes rho de Spearman) entre el rendimiento de la muestra completa en el TECLE Buenos Aires y el resto de las medidas, presentadas en orden decreciente, y agrupadas en términos de los procesos compartidos

Medida	Estadístico	Procesos compartidos
T-LVA pal	-.80	
T-TLC	-.74	Decodificación: precisión y velocidad en el reconocimiento de palabras escritas
T-LVA pseudo	-.74	
LVA pal	.72	
LVA pseudo	.68	
DAR	-.59	
TLC	.51	Velocidad en la recuperación léxica, procesamiento de material escrito, funcionamiento ejecutivo
Send 2 (con interferencia)	-.43	
Comp LEE	.42	
Send 1 (sin interferencia)	-.35	Velocidad de procesamiento, comprensión de vocabulario
Alfa-2	.32	

T-LVA pal: tiempo en la prueba de lectura en voz alta de palabras del test LEE; T-TLC: tiempo en la prueba de *screening* del TLC; LVA pal: precisión en la prueba de lectura en voz alta de palabras del test LEE; T-LVA pseudo: tiempo en la prueba de lectura en voz alta de pseudopalabras del test LEE; LVA pseudo: precisión en la prueba de lectura en voz alta de pseudopalabras del test LEE; DAR: tiempo en la denominación automatizada rápida de dibujos; TLC: precisión en la prueba de *screening* del TLC; Send 2: tiempo en la parte 2 de la prueba de trazado de senderos; Comp LEE: rendimiento en la prueba de comprensión de palabras y frases del test LEE; Send 1: tiempo en la parte 1 de la prueba de trazado de senderos; Alfa-2: rendimiento en la prueba de comprensión de vocabulario de la escala Alfa-2.

Finalmente, al analizar por un lado la muestra completa, y por otro, grupos más pequeños y homogéneos, se observaron diferencias en el conjunto de asociaciones que alcanzaron significación estadística, así como en los tamaños del efecto. Estas diferencias pueden haber sido la manifestación de los cambios en la contribución de los procesos componentes a la eficacia lectora a lo largo de la escolarización, o bien pueden haber estado influidas por las diferencias en los tamaños muestrales. Un estudio en grupos homogéneos de mayor tamaño podría contribuir a determinar el rol de las habilidades no específicas para la lectura en el rendimiento en el TECLE Buenos Aires.

4.2. Estudio de validez de criterio

El objetivo de este estudio fue estudiar la diferencia entre los puntajes observados en el test en dos grupos contrastados de sujetos: un grupo clínico conformado por sujetos con diagnóstico de trastornos en el aprendizaje de la lectura, de la escritura, de las matemáticas y/o trastornos del lenguaje, y un grupo control que incluyó a sujetos sin estas dificultades. Es decir que la variable criterio estuvo definida por el diagnóstico de trastornos en el aprendizaje y/o el lenguaje. La inclusión de participantes con trastornos en el aprendizaje de la escritura, de las matemáticas y trastornos del lenguaje se basó en la asociación frecuente entre estas dificultades y las que afectan al aprendizaje de la lectura (Fletcher, Lyon, Fuchs, & Barnes, 2007).

La hipótesis de trabajo fue que los puntajes observados en el grupo clínico serán significativamente inferiores a los observados en el grupo control.

4.2.1. Metodología

Se ajusta a los requerimientos de los estudios de validez de criterio. Se utilizó la prueba *t* para muestras independientes que suponen la comparación de los puntajes del test con el de otra medida criterio, en este caso entre las puntuaciones específicas de lectura obtenidas por un grupo clínico y uno control (no clínico), tal como se señalara precedentemente.

4.2.1.1. Participantes

Se conformaron dos muestras por conveniencia, cada una conformada por 28 participantes de 2° a 6° grados, en adelante, grupo clínico y grupo control. Los criterios de inclusión para ambos grupos fueron: no poseer trastornos intelectuales generales, motores o sensoriales no corregidos, haber adquirido el español como lengua materna y no haber repetido grados escolares.

Los sujetos del grupo clínico tenían diagnóstico de trastornos en el aprendizaje de la lectura, de la escritura, de las matemáticas y/o trastornos del lenguaje, que en el momento de la evaluación estaban siendo tratados por una profesional fonoaudióloga o psicopedagoga.

4.2.1.2. Instrumento

➤ **TECLE Buenos Aires.** Referimos al lector al capítulo 2 para una descripción detallada de esta prueba.

4.2.1.3. Procedimientos

Con el objeto de conformar el grupo clínico, se convocó a profesionales fonoaudiólogas y psicopedagogas de la Ciudad Autónoma de Buenos Aires y alrededores que estuvieran desempeñándose en el diagnóstico y el tratamiento de los trastornos del aprendizaje y del lenguaje, sea en un ámbito institucional o particular. Se explicitaron los objetivos del estudio y se expusieron los criterios de selección de los niños. Una vez contactados los participantes, se obtuvo el consentimiento informado por escrito de los padres, se les proporcionó información sobre la investigación, asegurándoles la confidencialidad y anonimato de los datos y luego se procedió a la evaluación.

El grupo control se conformó a través de la selección de los puntajes observados en el TECLE Buenos Aires de los participantes de la muestra para la confección de los baremos. Para cada participante del grupo clínico, se reunió un conjunto de sujetos control del mismo sexo, edad, grado escolar y nivel de oportunidades educativas. Del conjunto de candidatos reunido para cada participante del grupo clínico se seleccionó un sujeto al azar. Para tres niños del grupo clínico que compartían las características mencionadas se halló un único sujeto control, por lo que dos participantes del grupo clínico fueron excluidos al azar. Del mismo modo, se excluyó a otro participante del grupo clínico que compartía las características mencionadas con otros tres niños del mismo grupo, debido a que se hallaron tres sujetos del grupo control.

Para el grupo clínico, el TECLE Buenos Aires fue administrado individualmente por la tesista en el consultorio habitual de atención del niño y previamente a la sesión de tratamiento. En el grupo control, la administración se realizó en los establecimientos educativos correspondientes.

Se confeccionó una base de datos con los puntajes observados en ambos grupos, que fueron analizados con la versión 17.0 del Paquete Estadístico para las Ciencias Sociales SPSS. En primer lugar, se analizaron los estadísticos descriptivos correspondientes a las medidas de tendencia central y variabilidad del rendimiento en el TECLE Buenos Aires de cada grupo. Luego se realizaron las pruebas de normalidad y homocedasticidad con el fin de seleccionar el procedimiento estadístico adecuado para la comparación de las medias de ambos grupos. Finalmente, se estimó la diferencia entre las medias de ambos grupos así como el tamaño del efecto.

4.2.2. Resultados

4.2.2.1. Estadísticos descriptivos

La tabla 4.6 muestra la media, el desvío estándar y la varianza de las puntuaciones observadas en el TECLE Buenos Aires en ambos grupos.

Tabla 4.6

Media, desvío estándar y varianza de las puntuaciones observadas en el TECLE Buenos Aires en ambos grupos

Grupos	N	Media	DE	Varianza
Clínico	28	28.3	13.6	185
Control	28	35.1	11.1	123

DE: desvío estándar

4.2.2.2. Pruebas de normalidad y homocedasticidad

La prueba de normalidad (Shapiro-Wilk) reveló que los puntajes observados en el TECLE Buenos Aires en ambos grupos se ajustaron a una distribución normal [grupo clínico: $W(28) = .96$, $p = .28$; grupo control: $W(28) = .95$, $p = .24$]. La prueba de Levene mostró igualdad de varianzas entre ambos grupos [$F(54) = 3.26$, $p = .076$].

4.2.2.3. Análisis de las diferencias entre las puntuaciones

Una prueba t para muestras independientes reveló una diferencia estadísticamente significativa entre las medias de dichos puntajes [$t(54) = -2.07, p = .044$], con un tamaño del efecto mediano ($d = .55$) (Cohen, 1988).

4.2.3. Discusión y conclusiones del estudio de validez de criterio

Es de señalar que la diferencia hallada entre las medias de los puntajes obtenidos en el TECLE Buenos Aires por el grupo clínico y el grupo control aportan evidencias de validez de criterio.

Sin embargo, la utilización de este tipo de evidencia para la interpretación de los puntajes observados en el test requiere de una reflexión acerca de la variable criterio utilizada en este estudio. Dicha variable criterio fue el diagnóstico de trastornos en el aprendizaje de la lectura, de la escritura, de las matemáticas, y/o trastornos del lenguaje, debido a la asociación frecuente entre estos trastornos (Fletcher, Lyon, Fuchs, & Barnes, 2007).

Por un lado, es posible que dicha asociación obedezca al déficit en el procesamiento en un componente (o más) de la arquitectura funcional para la lectura que también es reclutado para tareas en otro dominio. Por ejemplo, el conocimiento ortográfico está involucrado no sólo en tareas de reconocimiento de palabras escritas, sino también en tareas de escritura y en la comprensión de problemas matemáticos presentados bajo la modalidad escrita.

Por otro lado, también es posible que en algunos casos los procesos neurobiológicos del desarrollo resulten en la conformación atípica de una variedad de regiones cerebrales necesarias para la implementación del procesamiento en más de un dominio académico.

Por lo tanto, la evidencia de validez de criterio aportada por el presente estudio no justifica la utilización del test en reemplazo del procedimiento diagnóstico de los trastornos mencionados. En este sentido, es necesario

además tomar las precauciones del caso a la hora de distinguir los indicadores diagnósticos que proporciona una prueba, aunque resulte válida y confiable, y el diagnóstico propiamente dicho que debe ser realizado por un profesional competente, a la luz de diversas fuentes de información.

Por lo expuesto, la evidencia de validez de criterio hallada en este estudio justifica el uso del TECLE Buenos Aires como una medida indicadora de la existencia de riesgo de desarrollar o presentar trastornos en el aprendizaje de la lectura. El diagnóstico preciso de dichos trastornos requiere de la utilización de una batería estandarizada y referida a normas conformada por pruebas administradas individualmente, dirigidas a determinar el grado de funcionamiento de los componentes específicos para la lectura.

4.3. Discusión y conclusiones generales de los estudios de validez

Los resultados de los estudios de validez de constructo y de criterio presentados en este capítulo constituyen evidencias de que los puntajes obtenidos en el TECLE Buenos Aires pueden interpretarse como una medida global de “eficacia lectora” controlada principalmente por las habilidades de decodificación.

Sin embargo, es necesario notar que la generalización de esta conclusión es limitada dado el tamaño de las muestras utilizadas así como el rango de categorías de las variables grado escolar, edad y nivel de oportunidades educativas. No obstante, estos resultados podrían contribuir a la realización de meta-análisis dirigidos a incrementar el grado de generalización de la validez.

Por otro lado, sería recomendable llevar a cabo un estudio longitudinal de validez predictiva que se proponga medir el grado de asociación entre el rendimiento en el TECLE Buenos Aires en los primeros años de escuela primaria y el desempeño en lectura en años posteriores, utilizando una batería de pruebas de lectura estandarizadas y con baremos locales. La

obtención de índices de sensibilidad y especificidad del TECLE Buenos Aires a través de este estudio permitiría determinar la utilidad del test en la detección temprana y oportuna de trastornos del aprendizaje de la lectura.

Finalmente, existe otro aspecto de la validez no siempre explicitado, o aún tenido en cuenta, en el momento de argumentar a favor o en contra de las interpretaciones y los usos propuestos de las puntuaciones de los tests. El tema de las evidencias de validez tomando en cuenta las consecuencias de la evaluación fue planteado hace décadas por Cronbach (1971, 1982, 1988; citado en Lane, 2014), y aún hoy continúa el debate acerca de su pertinencia para determinar la calidad psicométrica de un instrumento. Refiere al estudio de la red de inferencias elaboradas a partir de los resultados observados, que se conectan de manera lógica con las conclusiones, las decisiones y las consecuencias de dichas decisiones. Para ello, no sólo es necesario evaluar las consecuencias de las interpretaciones y usos que un test propone para los puntajes observados, sino también las consecuencias de los que no propone. Las decisiones incluyen, entre otras, la clasificación de los individuos como pertenecientes a una u otra categoría (con o sin dificultades en el aprendizaje) y la implementación de modificaciones en los programas de enseñanza dirigidos a mejorar el desempeño de los alumnos. Las consecuencias negativas no contempladas por los diseñadores y/o los usuarios de los tests incluyen, por un lado, los efectos adversos asociados al “estigma” de una etiqueta diagnóstica, y por otro, la exclusión del acceso a prácticas de enseñanza más controladas de alumnos sin un diagnóstico.

En el capítulo 6 “Discusión y conclusiones de la tesis”, se vuelve sobre este punto al discutir las interpretaciones y usos de los baremos generales y los segmentados por nivel de oportunidades educativas.

Capítulo 5

Los datos normativos del TECLE Buenos Aires para la Ciudad Autónoma de Buenos Aires y alrededores. Justificación estadística

Los tests psicológicos son instrumentos que permiten estimar el grado en que las personas poseen determinados atributos psicológicos variables e inobservables. Están compuestos por una muestra de ítems a los que los sujetos examinados responden poniendo en juego el conocimiento y los procesos psicológicos que hipotéticamente los tests pretenden medir. En el caso del TECLE Buenos Aires, a partir de la calificación de las respuestas a cada uno de los 64 ítems se obtiene un puntaje bruto o directo.

Sin embargo, los puntajes directos son inherentemente arbitrarios y ambiguos ya que dependen de la cantidad de ítems del test y del rango de calificaciones posibles que admiten algunas pruebas para cada respuesta. Es decir que los puntajes directos en sí mismos no permiten inferir el nivel de conocimiento o el grado de dominio en la habilidad de interés. Por lo tanto, su interpretación es un proceso necesario en la utilización de los tests (Furr & Bacharach, 2014; Olaz & Tornimbeni, 2008). Cuando se trata de instrumentos que miden el tiempo de ejecución en una tarea, el puntaje directo se expresa en unidades de tiempo (comúnmente minutos, segundos o milisegundos), pero resultan igualmente poco informativos.

Los puntajes de los tests psicológicos pueden utilizarse con dos propósitos: determinar la posición de un sujeto examinado en relación con un grupo de referencia, o bien indagar si dicho sujeto posee o no determinada característica. Así, el uso propuesto de un test determina cómo deben interpretarse los puntajes que arroja: en el primer caso, la interpretación de los puntajes está referida a normas, mientras que en el segundo, está referida a un criterio.

En la interpretación de los puntajes referida a normas se dispone de información acerca de la distribución de los datos en un grupo de referencia, que se resume en las medidas de tendencia central y de variabilidad o dispersión de los datos. Para ello, existe una serie de procedimientos por los cuales los puntajes directos que arrojan los tests se transforman en puntajes derivados que permiten determinar la posición relativa de los sujetos examinados en la distribución de puntajes del grupo de referencia.

Los puntajes derivados más frecuentemente utilizados son los puntajes estándar, los percentiles y los rangos percentiles. Los primeros se utilizan cuando los puntajes directos siguen una distribución normal, se expresan en unidades de desvío estándar e involucran una transformación lineal de los puntajes directos, es decir que la distancia entre dos puntajes estándar es proporcional a la distancia entre los puntajes directos de los que se derivan. Cuando la distribución de los puntajes directos no se ajusta a una distribución normal, pueden utilizarse los percentiles y los rangos percentiles. Los percentiles se refieren a los puntajes directos por debajo de los cuales se encuentra un porcentaje dado de observaciones; por otro lado, los rangos percentiles se refieren al porcentaje de observaciones con valores inferiores al puntaje directo que se desea interpretar.

Ambos involucran una transformación no lineal de los puntajes directos. Esto es, la distancia entre dos percentiles no refleja la distancia entre sus dos puntajes directos, así como la distancia entre dos rangos percentiles no refleja la distancia entre los porcentajes. Esto se debe a que ningún atributo psicológico se distribuye de manera monótona en la población sino que una alta proporción de individuos obtienen valores cercanos a un valor central, mientras que unos pocos corresponden a puntajes extremos.

Para las personas no expertas en procedimientos estadísticos, los percentiles y rangos percentiles ofrecen una interpretación más fácilmente comprensible que los puntajes estándar y es por ello que muchos autores optan por presentar las normas de sus tests en términos de tales medida de posición (Olaz & Tornimbeni, 2008).

Un aspecto fundamental en la interpretación referida a normas de los puntajes directos observados en un test es el requerimiento de que la muestra de referencia sea representativa de la población a la que el test está dirigido, para lo cual se recomienda que la misma sea probabilística y de un tamaño adecuado. Sin embargo, en el presente estudio, varias razones impidieron que la muestra de referencia cumpla con dicho requisito (ver el apartado 5.2.1 “La muestra de tipificación” más abajo).

Los objetivos de este estudio fueron, en primer lugar, confeccionar los baremos según grado escolar. En segundo lugar, explorar, por un lado, el efecto sobre el rendimiento en el test de las variables independientes “grado escolar”, “edad en años” y “sexo” de los niños; por el otro, el de las variables “nivel de oportunidades educativas”, “tipo de gestión” y “región geográfica” de las escuelas a las que asistían, de modo tal de confeccionar los baremos en función de aquéllas que indicaran las mayores diferencias entre los grupos.

5.1. Metodología

Como en los otros estudios, corresponde a una metodología instrumental específica para la elaboración/adaptación de instrumentos de evaluación psicológica, destinada en este caso a la obtención de las puntuaciones del test, su transformación en percentiles y rangos percentiles para la confección de los baremos correspondientes. Además, interesaba examinar la incidencia en el desempeño en la prueba (variable dependiente) de cada una de las variables independientes consideradas: grado escolar, edad en años y sexo de los participantes, y nivel de oportunidades educativas (alto, medio y bajo), tipo de gestión (estatal y privada) y región geográfica de las escuelas (Ciudad Autónoma de Buenos Aires y sur, oeste y norte del Gran Buenos Aires). En función de ello, se aplicaron pruebas paramétricas (prueba t , ANOVA y prueba de Games-Howell) y no paramétricas (H de Kruskal-Wallis y U de Mann-Whitney) para analizar las diferencias entre las puntuaciones medias de dos o más grupos o variables, previo análisis de la normalidad de las distribuciones (prueba de Shapiro-

Wilk) y homogeneidad de las varianzas (prueba de Levene). Teniendo en cuenta que no resulta suficiente identificar la presencia de un cierto efecto, sino que se requiere, además, “determinar su magnitud o tamaño” (Ledesma, Macbeth, & Cortada de Kohan, 2008, p. 426), se procedió a evaluar el tamaño del efecto (η^2_p -eta cuadrado parcial- y d de Cohen) para poder establecerlo, con el propósito de contar con elementos de juicio para la elaboración de baremos generales y parciales, así como para las recomendaciones para su uso.

5.1.1. La muestra de tipificación

Para este estudio se conformó una muestra por conveniencia, en razón de las dificultades encontradas para la conformación de una muestra probabilística. El principal impedimento fue la dificultad para obtener las autorizaciones de los funcionarios de algunos distritos escolares o regiones educativas para llevar a cabo la evaluación en las escuelas. En algunos casos, no se obtuvo una respuesta a la presentación escrita del proyecto, y en otros, la respuesta fue desfavorable. Un segundo impedimento fue que, en algunas escuelas, en su mayoría de gestión privada, no se obtuvo el consentimiento informado por escrito de una proporción de los padres de los participantes. Este hecho puede haber generado una posible fuente de sesgo en los puntajes observados en dichas escuelas, aunque no se dispone de la información necesaria para tal afirmación. Por último, no se contó con información acerca de la distribución en la población de la variable nivel de oportunidades educativas de las escuelas o de sus indicadores.

La muestra inicial para la obtención de los baremos estuvo conformada por 3614 participantes. De ella se descartaron 517 sujetos de acuerdo con uno o más de los siguientes criterios de exclusión: tener diagnóstico de algún trastorno neurológico o del desarrollo, estar en proceso de integración escolar, haber repetido grados escolares, tener una edad no acorde al grado escolar, estar institucionalizado, haber adquirido una lengua materna diferente al español, ser bilingüe, no haber respetado parte de la consigna en cuanto a completar el test de manera individual, dentro del tiempo límite,

desde el ítem 1 en adelante y marcando sólo una opción por ítem, no haber completado la página de datos personales (edad, sexo, grado), exhibir un patrón de respuestas estereotipado, haber omitido completar los ítems de una página. De esta manera, la muestra final quedó conformada por 3097 sujetos.

Nivel de oportunidades educativas. En este estudio, así como en el estudio para la validez de criterio, y en consonancia con los objetivos propuestos, los participantes incluidos en la muestra fueron caracterizados según las variables grado escolar, edad en años y sexo, y tipo de gestión, región geográfica y nivel de oportunidades educativas de las escuelas a las que asistían. La inclusión de esta última variable tiene por objeto operacionalizar el constructo de la contribución de los factores ambientales en el desarrollo típico y atípico de las habilidades de lectura, en el marco de la idea de la “división del trabajo” planteado por Byrne (2005) y desarrollada en el capítulo 1 “Marco teórico y antecedentes”.

Cabe mencionar que el nivel de oportunidades educativas fue considerado en otras investigaciones realizadas en el país. Ferreres, Abusamra, Squillace, Fernández Liporace, Cartoceti, & Sampedro (2010), analizaron el efecto de dicha variable sobre el rendimiento de niños de los grados 5°, 6° y 7° de educación primaria (o 1° año de ESB) en un test de comprensión de textos. El nivel de oportunidades educativas fue definido en términos del nivel socio-económico predominante de la población escolar y de una serie de indicadores referidos a las instituciones, como los porcentajes de ausentismo, repitencia y deserción escolar, el equipamiento y la infraestructura y el dictado de actividades extracurriculares. Los autores encontraron que esta variable impactó significativamente en la comprensión de textos, con un rendimiento menor en los niños que asistían a escuelas con bajo nivel de oportunidades educativas. Estos hallazgos fueron tenidos en cuenta en el momento de considerar su inclusión en los estudios de la presente tesis.

En otros trabajos se ha considerado la incidencia de las características del contexto alfabetizador hogareño en familias de la pobreza urbana y en familias no pobres, en relación con el futuro desempeño escolar de sus hijos (Piacente, Marder, & Resches, 2008). Los resultados muestran las experiencias diferenciales en las que participan los niños durante los años preescolares, que no siempre son compensadas en la enseñanza formal a través de oportunidades educativas apropiadas. En el mismo sentido, otros trabajos (Piacente, 2012) han señalado la imposibilidad de identificar las competencias del lector/escritor si se carece de “información sobre las características textuales, los propósitos de las actividades de lectura y escritura y las estrategias de enseñanza y el contexto en el que se llevan a cabo” (p. 21).

Bajo esas consideraciones, las escuelas a las que concurrían los niños de la muestra fueron categorizadas según tres niveles de oportunidades educativas: bajo, medio y alto. Sin embargo, es necesario aclarar que debido a la dificultad para obtener información proveniente de las autoridades de las escuelas acerca de indicadores referidos a las instituciones a su cargo, el proceso de categorización se basó, por un lado, en información parcial acerca de la infraestructura, el equipamiento y el dictado de actividades extracurriculares, y por otro lado, en la impresión y el juicio subjetivos de la tesista acerca del nivel socio-económico-cultural de la población escolar que asistía a dichas instituciones.

En la tabla 5.1 se muestra la distribución de las escuelas por tipo de gestión, nivel de oportunidades educativas y región geográfica.

Las tablas 5.2, 5.3, 5.4 y 5.5 muestran para cada grado escolar, la distribución de los alumnos según el sexo, la región geográfica, el tipo de gestión y el nivel de oportunidades educativas de las escuelas.

Tabla 5.1

Distribución de las escuelas por tipo de gestión, región geográfica y nivel de oportunidades educativas

Gestión	NOE	Sur	Oeste	Norte	CABA
Estatal	Bajo	3	2	2	1
	Medio	2	3	1	1
	Alto	-	-	-	-
Privada	Bajo	-	-	-	-
	Medio	3	-	-	3
	Alto	-	-	2	1

NOE: Nivel de oportunidades educativas; Sur: sur del Gran Buenos Aires; Oeste: oeste del Gran Buenos Aires; Norte: norte del Gran Buenos Aires; CABA: Ciudad Autónoma de Buenos Aires.

Tabla 5.2

Distribución de los participantes según sexo para cada grado escolar

Grado		Sexo		Total
		M	F	
2°	N (%)	321 (53.3)	281 (46.7)	602 (100)
3°	N (%)	320 (49.8)	323 (50.2)	643 (100)
4°	N (%)	330 (46.9)	373 (53.1)	703 (100)
5°	N (%)	251 (50.4)	247 (49.6)	498 (100)
6°	N (%)	330 (50.7)	321 (49.3)	651 (100)
Total	N (%)	1552 (50.1)	1545 (49.9)	3097 (100)

M: masculino; F: femenino.

Tabla 5.3

Distribución de los participantes según región geográfica de la escuela para cada grado escolar

Grado		Región				Total
		Sur	Oeste	Norte	CABA	
2°	N (%)	194 (32.2)	164 (27.2)	189 (31.4)	55 (9.14)	602 (100)
3°	N (%)	255 (39.7)	168 (26.1)	128 (19.9)	92 (14.3)	643 (100)
4°	N (%)	203 (28.9)	173 (24.6)	221 (31.4)	106 (15.1)	703 (100)
5°	N (%)	198 (39.8)	113 (22.7)	72 (14.5)	115 (23.1)	498 (100)
6°	N (%)	222 (34.1)	169 (26.0)	172 (26.4)	88 (13.5)	651 (100)
Total	N (%)	1072 (34.6)	787 (25.4)	782 (25.3)	456 (14.7)	3097 (100)

Sur: sur del Gran Buenos Aires; Oeste: oeste del Gran Buenos Aires; Norte: norte del Gran Buenos Aires; CABA: Ciudad Autónoma de Buenos Aires.

Tabla 5.4

Distribución de los participantes según tipo de gestión de la escuela para cada grado escolar

Grado	Gestión		Total
	Estatal	Privada	
2° N (%)	362 (60.1)	240 (39.9)	602 (100)
3° N (%)	455 (70.8)	188 (29.2)	643 (100)
4° N (%)	456 (64.9)	247 (35.1)	703 (100)
5° N (%)	319 (64.1)	179 (35.9)	498 (100)
6° N (%)	431 (66.2)	220 (33.8)	651 (100)
Total N (%)	2023 (65.4)	1074 (34.6)	3097 (100)

Tabla 5.5

Distribución de los participantes según nivel de oportunidades educativas de la escuela para cada grado escolar

Grado	Nivel de oportunidades educativas			Total
	Bajo	Medio	Alto	
2° N (%)	133 (22.1)	352 (58.5)	117 (19.4)	602 (100)
3° N (%)	202 (31.4)	405 (63.0)	36 (5.6)	643 (100)
4° N (%)	199 (28.3)	386 (54.9)	118 (16.8)	703 (100)
5° N (%)	140 (28.1)	337 (67.7)	21 (4.2)	498 (100)
6° N (%)	173 (26.6)	390 (59.9)	88 (13.5)	651 (100)
Total N (%)	847 (27.3)	1870 (60.4)	380 (12.3)	3097 (100)

5.1.2. Instrumento

Se utilizó el TECLE Buenos Aires (ver capítulo 2 “El Test Colectivo de Eficacia Lectora. Su adaptación” para mayor detalle).

5.1.3. Procedimientos

En primer lugar, con el objeto de conformar una muestra de tamaño adecuado se estableció contacto con funcionarios a cargo de varios distritos escolares, regiones educativas y establecimientos educativos de nivel primario, de gestión estatal y privada de la Ciudad de Buenos Aires y alrededores, a quienes se les presentó el test y el proyecto de obtención de los baremos.

Previamente a la evaluación, se obtuvo el consentimiento informado por escrito de los padres de los participantes y se les informó sobre la investigación, garantizando el anonimato y confidencialidad de los datos.

El test fue administrado de manera colectiva por la tesista, con la colaboración de profesionales psicólogos especialmente capacitados, dentro del horario escolar, en el aula de cursada habitual para cada grado y sección, y en presencia del docente a cargo del grupo.

Luego de la evaluación de los protocolos individuales se confeccionó una matriz de datos con los puntajes directos obtenidos en el test, que fueron analizados con la versión 17.0 del Paquete Estadístico para las Ciencias Sociales SPSS.

En primer lugar, se analizaron los estadísticos descriptivos correspondientes a las medidas de tendencia central y variabilidad de los puntajes observados en los grupos según el grado escolar, la edad en años y el sexo de los niños y el nivel de oportunidades educativas de las escuelas.

Para estudiar el efecto del tipo de gestión se analizó el rendimiento del grupo que asistía a escuelas con nivel de oportunidades educativas medio, dado que este grupo incluyó escuelas de gestión estatal y de gestión privada mientras que el grupo que asistía a escuelas con nivel alto no incluyó escuelas de gestión estatal así como el grupo con nivel bajo no incluyó escuelas de gestión privada.

Para estudiar el efecto de la región geográfica se analizó el rendimiento del grupo que asistía a escuelas de gestión estatal ya que el mismo incluyó escuelas de las cuatro regiones: sur, oeste y norte del Gran Buenos Aires y Ciudad Autónoma de Buenos Aires.

Luego se aplicaron las pruebas de normalidad y homocedasticidad a las distribuciones de los puntajes observados en los distintos grupos con el objeto de seleccionar el procedimiento estadístico adecuado para el análisis de las diferencias. A continuación, se comparó el rendimiento entre los

distintos grupos. En las comparaciones en las que era recomendable aplicar pruebas no paramétricas, se realizaron adicionalmente las pruebas paramétricas correspondientes, toda vez que los resultados orientaran ese procedimiento.

Finalmente, se confeccionaron los baremos sobre la base del rendimiento en el test de los participantes incluidos a partir de la transformación de los puntajes directos observados en percentiles y rangos percentiles correspondientes a la muestra completa segmentada de acuerdo al siguiente detalle.

- Según el grado escolar (baremos generales).
- Según el grado escolar y el nivel de oportunidades educativas (baremos según nivel de oportunidades educativas).

5.2. Resultados

En primer lugar, se informan los estadísticos descriptivos correspondientes a las medidas de tendencia central y variabilidad de los puntajes observados en la muestra completa segmentada según el grado escolar, los resultados de las pruebas de normalidad y homocedasticidad aplicadas a las distribuciones de estos puntajes para la selección de los estadísticos adecuados para la comparación entre los grupos y los resultados de los análisis de las diferencias.

A continuación, se informan los resultados de estos análisis aplicados a cada grupo según la edad y el sexo de los participantes, y según el nivel de oportunidades educativas de las escuelas a las que asistían.

Seguidamente, se informan los resultados de los análisis aplicados al grupo con nivel de oportunidades educativas medio segmentado según el tipo de gestión y los aplicados a los grupos de gestión estatal segmentados según la región.

Las tablas con los baremos generales correspondientes a cada grado escolar y los correspondientes a cada grado escolar segmentado según el nivel de oportunidades educativas se presentan en el Anexo B “Baremos del TECLE Buenos Aires para la Ciudad Autónoma de Buenos Aires y alrededores”.

5.2.1. Desempeño en la prueba según grado escolar

5.2.1.1. Estadísticos descriptivos

Dado que se espera una progresión en el desempeño de la prueba a medida que se avanza en la escolaridad, se calcularon los estadísticos descriptivos correspondientes a la media, el desvío estándar y la varianza de los puntajes observados en la muestra completa segmentada según el grado escolar (Tabla 5.6).

Tabla 5.6

Media, desvío estándar y varianza de los puntajes observados en los grupos según el grado escolar

Grado	N	Media	DE	Varianza
2°	602	15.6	8.65	74.7
3°	643	22.5	11.2	125
4°	703	29.6	10.7	115
5°	498	35.8	11.6	136
6°	651	39.4	11.9	141

DE: desvío estándar

5.2.1.2. Pruebas de normalidad y homocedasticidad

Para la selección de los estadísticos a utilizar en el análisis de las diferencias según grado escolar se aplicó la prueba de normalidad a los puntajes observados en cada grado escolar (Shapiro-Wilk). La misma reveló que no se ajustaron a una distribución normal [2° grado: $W(602) = .97, p < .001$; 3° grado: $W(643) = .98, p < .001$; 4° grado: $W(703) = .99, p = .013$; 5° grado: $W(498) = .99, p = .004$; 6° grado: $W(651) = .99, p = .001$].

A su vez, la prueba de Levene mostró que las varianzas entre los puntajes observados en los grupos no fue homogénea [$F(4, 3092) = 18.5; p < .001$].

5.2.1.3. Análisis de las diferencias entre las puntuaciones

En virtud de los resultados obtenidos en las pruebas de normalidad y homocedasticidad, se aplicó la prueba H de Kruskal-Wallis para comparar el rendimiento entre los grupos según el grado escolar. Sin embargo, dado que los resultados de esta prueba fueron congruentes con los de los análisis paramétricos correspondientes, a continuación se reportan los resultados de las pruebas de ANOVA y Games-Howell, mientras que los resultados de la prueba H de Kruskal-Wallis se informan en el apartado A.2.1., que aparecen en el Anexo A “Análisis no paramétricos”.

La prueba de ANOVA reveló un efecto estadísticamente significativo del grado escolar en los puntajes observados entre todos los grados [$F(4, 3092) = 482, p < .001, \eta^2_p = .38$]. Según la sugerencia de Cohen (1988, 1992; Richardson, 2011), los valores de referencia de η^2_p (eta cuadrado parcial) para un tamaño del efecto pequeño, mediano y grande son, respectivamente, .010, .059 y .14. Por lo tanto, el tamaño del efecto encontrado es grande. La tabla 5.7 muestra el resultado de las comparaciones *post hoc* utilizando la prueba de Games-Howell, en donde se observa que las diferencias entre las medias de los puntajes entre todos los grados alcanzaron significación estadística.

Tabla 5.7

Diferencias entre las medias de los puntajes observados en los grupos según el grado escolar (prueba de Games-Howell)

(I) Grado	(J) Grado	Diferencia de medias (I-J)	Error estándar	Sig.	Intervalo de confianza de 95%	
					Límite inferior	Límite superior
2°	3°	-6.90*	.56	< .001	-8.44	-5.35
	4°	-14.0*	.54	< .001	-15.5	-12.5
	5°	-20.2*	.63	< .001	-21.9	-18.5
	6°	-23.8*	.58	< .001	-25.3	-22.2
3°	2°	6.90*	.56	< .001	5.35	8.44
	4°	-7.10*	.60	< .001	-8.74	-5.47
	5°	-13.3*	.68	< .001	-15.2	-11.4
	6°	-16.9*	.64	< .001	-18.6	-15.1
4°	2°	14.0*	.54	< .001	12.5	15.5
	3°	7.10*	.60	< .001	5.47	8.74
	5°	-6.21*	.66	< .001	-8.01	-4.40
	6°	-9.76*	.62	< .001	-11.4	-8.08
5°	2°	20.2*	.63	< .001	18.5	21.9
	3°	13.3*	.68	< .001	11.4	15.2
	4°	6.21*	.66	< .001	4.40	8.01
	6°	-3.55*	.70	< .001	-5.46	-1.64
6°	2°	23.8*	.58	< .001	22.2	25.3
	3°	16.9*	.64	< .001	15.1	18.6
	4°	9.76*	.62	< .001	8.08	11.4
	5°	3.55*	.70	< .001	1.64	5.46

*. La diferencia entre las medias es significativa en el nivel de .05.

5.2.2. Desempeño en la prueba según edad

5.2.2.1. Estadísticos descriptivos

También en este caso se espera una progresión del desempeño a medida que se avanza en edad; en la tabla 5.8 se presentan la media, el desvío estándar y la varianza de los puntajes observados en la muestra completa segmentada según la edad en años.

Tabla 5.8

Media, desvío estándar y varianza de los puntajes observados en los grupos según la edad en años

Edad	N	Media	DE	Varianza
7	361	16.4	8.65	74.8
8	673	19.7	10.9	120
9	644	27.3	11.3	128
10	592	33.3	11.6	136
11	569	38.5	12.3	151
12	258	37.9	11.8	140

DE: desvío estándar

5.2.2.2. Pruebas de normalidad y homocedasticidad

La prueba de normalidad aplicada a los puntajes observados en cada grupo según la edad (Shapiro-Wilk) reveló que se ajustaron a una distribución normal para los grupos de 10 y 12 años [7 años: $W(361) = .97$, $p < .001$; 8 años: $W(673) = .97$, $p < .001$; 9 años: $W(644) = .99$, $p = .002$; 10 años: $W(592) = 1.00$, $p = .11$; 11 años: $W(569) = .99$, $p = .001$; 12 años: $W(258) = .99$, $p = .11$]. Por otro lado, la prueba de Levene reveló que las varianzas entre los puntajes observados en los grupos no fue homogénea [$F(5, 3091) = 10.7$; $p < .001$].

5.2.2.3. Análisis de las diferencias entre las puntuaciones

El análisis de las diferencias observadas entre los distintos grupos de edad se realizó siguiendo el mismo procedimiento que para el análisis de las diferencias observadas en función del grado escolar. En el apartado A.2.2. del Anexo A “Análisis no paramétricos” se informan los resultados de la prueba H de Kruskal-Wallis.

Una prueba de ANOVA reveló una diferencia estadísticamente significativa con un tamaño del efecto grande de la edad sobre los puntajes observados [$F(5, 3091) = 316$, $p < .001$, $\eta^2_p = .34$]. La tabla 5.10 muestra el resultado de las comparaciones *post hoc* utilizando la prueba de Games-Howell, en donde se observa que las diferencias entre las medias de los puntajes observados en los grupos por edad alcanzaron significación

estadística en todas las comparaciones, a excepción de la diferencia entre los grupos de 11 y 12 años (Tabla 5.9).

Tabla 5.9

Diferencias entre las medias de los puntajes observados en los grupos según la edad (prueba de Games-Howell)

(I) Edad	(J) Edad	Diferencia de medias (I-J)	Error estándar	Sig.	Intervalo de confianza de 95%	
					Límite inferior	Límite superior
7	8	-3.30 [*]	.62	< .001	-5.08	-1.53
	9	-10.9 [*]	.64	< .001	-12.8	-9.11
	10	-16.9 [*]	.66	< .001	-18.8	-15.1
	11	-22.2 [*]	.69	< .001	-24.2	-20.2
	12	-21.5 [*]	.87	< .001	-24.0	-19.0
8	7	3.30 [*]	.62	< .001	1.53	5.08
	9	-7.63 [*]	.61	< .001	-9.38	-5.88
	10	-13.6 [*]	.64	< .001	-15.5	-11.8
	11	-18.9 [*]	.67	< .001	-20.8	-17.0
	12	-18.2 [*]	.85	< .001	-20.6	-15.8
9	7	10.9 [*]	.64	< .001	9.11	12.8
	8	7.63 [*]	.61	< .001	5.88	9.38
	10	-6.01 [*]	.65	< .001	-7.88	-4.14
	11	-11.3 [*]	.68	< .001	-13.2	-9.31
	12	-10.6 [*]	.86	< .001	-13.0	-8.12
10	7	16.9 [*]	.66	< .001	15.1	18.8
	8	13.6 [*]	.64	< .001	11.8	15.5
	9	6.01 [*]	.65	< .001	4.14	7.88
	11	-5.25 [*]	.70	< .001	-7.25	-3.24
	12	-4.57 [*]	.88	< .001	-7.09	-2.06
11	7	22.2 [*]	.69	< .001	20.2	24.2
	8	18.9 [*]	.67	< .001	17.0	20.8
	9	11.3 [*]	.68	< .001	9.31	13.2
	10	5.25 [*]	.70	< .001	3.24	7.25
	12	.67	.90	.98	-1.90	3.24
12	7	21.5 [*]	.87	< .001	19.0	24.0
	8	18.2 [*]	.85	< .001	15.8	20.6
	9	10.6 [*]	.86	< .001	8.12	13.0
	10	4.57 [*]	.88	< .001	2.06	7.09
	11	-.67	.90	.98	-3.24	1.90

*. La diferencia entre las medias es significativa en el nivel de .05.

5.2.3. Desempeño en la prueba según grado escolar versus edad

Dada la posibilidad que los resultados encontrados según la edad sobre los puntajes observados hubiera sido una manifestación del grado escolar, y viceversa, se comparó el rendimiento entre los grupos por edad para cada grado escolar y entre los grupos por grado escolar para las edades de 8, 9, 10 y 11 años (debido a que la evaluación se llevó a cabo entre los meses de septiembre y noviembre, todos los niños de 7 años estaban cursando 2° grado, y todos los de 12 años estaban cursando 6° grado).

5.2.3.1. Estadísticos descriptivos según grado escolar y edad

En la tabla 5.10 se presentan la media, el desvío estándar y la varianza de los puntajes observados en la muestra completa segmentada según el grado escolar y la edad.

Tabla 5.10

Media, desvío estándar y varianza de los puntajes observados en los grupos según el grado escolar y la edad

Grado	Edad	N	Media	DE	Varianza
2°	7	361	16.4	8.65	74.8
	8	241	14.5	8.53	72.8
3°	8	432	22.5	11.1	123
	9	211	22.4	11.4	130
4°	9	433	29.7	10.5	111
	10	270	29.5	11.1	123
5°	10	322	36.5	11.2	125
	11	176	34.6	12.4	154
6°	11	393	40.3	11.8	140
	12	258	37.9	11.8	140

DE: desvío estándar

5.2.3.2. Pruebas de normalidad

La prueba de normalidad (Shapiro-Wilk) aplicada a los puntajes observados en cada grupo según el grado escolar y la edad reveló que la mayoría no se ajustó a una distribución normal [2° grado-7 años: $W(361) = .97$, $p < .001$; 2° grado-8 años: $W(241) = .96$, $p < .001$; 3° grado-8 años: $W(432) = .98$, $p < .001$; 3° grado-9 años: $W(211) = .98$, $p = .002$; 4° grado-9

años: $W(433) = .99, p = .003$; 4° grado-10 años: $W(270) = 1.00, p = .68$; 5° grado-10 años: $W(322) = .99, p = .066$; 5° grado-11 años: $W(176) = .99, p = .067$; 6° grado-11 años: $W(393) = .99, p = .013$; 6° grado-12 años: $W(258) = .99, p = .11$].

En virtud de estos resultados, en el apartado A.2.3 del Anexo A “Análisis no paramétricos” se informan los que corresponden a las pruebas *U* de Mann-Whitney utilizadas para analizar los efectos de la edad sobre el rendimiento en el test en cada grado escolar y los efectos del grado escolar sobre el rendimiento en los grupos de 8, 9, 10 y 11 años. No obstante, dado que la prueba de Levene reveló homogeneidad de varianzas para las comparaciones de interés, se aplicaron adicionalmente análisis paramétricos. A continuación se informan los resultados de estos análisis, que fueron en la misma dirección que los de las pruebas no paramétricas correspondientes.

5.2.3.3. Análisis de las diferencias entre las puntuaciones según grado escolar y edad

5.2.3.3.1. Prueba de homocedasticidad y análisis de las diferencias entre las puntuaciones según la edad para cada grado escolar

La prueba de Levene reveló homogeneidad de varianzas entre los grupos por edad para todos los grados [2° grado: $F(1, 600) = .056, p = .81$; 3° grado: $F(1, 641) = .041, p = .84$; 4° grado: $F(1, 701) = .36, p = .55$; 5° grado: $F(1, 496) = 1.26, p = .26$; 6° grado: $F(1, 649) = .017, p = .90$].

Una serie de pruebas *t* para muestras independientes revelaron que las diferencias alcanzaron significación estadística entre los grupos de 7 y 8 años para 2° grado y entre los grupos de 11 y 12 años para 6° grado [2° grado: $t(600) = 2.61, p = .009, d = .22$; 6° grado: $t(649) = 2.60, p = .010, d = .20$]. El tamaño del efecto de la edad en estas comparaciones sería pequeño, dada la sugerencia de Cohen (1988, 1992; Richardson, 2011) acerca de valores de referencia de *d* de .20, .50 y .80 para tamaños de efecto pequeño, mediano y grande, respectivamente.

En los grados restantes, las diferencias en el rendimiento entre los grupos por edad no alcanzaron significación estadística [3° grado: $t(641) = .15$, $p = .88$; 4° grado: $t(701) = .21$, $p = .83$; 5° grado: $t(496) = 1.78$, $p = .075$].

5.2.3.3.2. Prueba de homocedasticidad y análisis de la diferencias entre las puntuaciones según grado escolar para los grupos de 8, 9, 10 y 11 años

La prueba de Levene reveló homogeneidad de varianzas entre los grupos por grado escolar en todas las edades analizadas, a excepción de la comparación de los grupos de 8 años por lo que se ajustaron los grados de libertad de 671 a 606 [8 años: $F(1, 606) = 27.6$, $p < .001$; 9 años: $F(1, 642) = 1.86$, $p = .17$; 10 años: $F(1, 590) = .37$, $p = .85$; 11 años: $F(1, 567) = .048$, $p = .83$].

En contraposición con los resultados obtenidos en los análisis realizados entre los grupos por edad para cada grado escolar, todas las comparaciones del rendimiento entre los grupos por grado escolar en niños de la misma edad revelaron diferencias estadísticamente significativas con tamaños del efecto pequeño a mediano [8 años: $t(606) = -10.5$, $p < .001$, $d = .13$; 9 años: $t(642) = -8.00$, $p < .001$, $d = .67$; 10 años: $t(590) = .7.63$, $p < .001$, $d = .63$; 11 años: $t(567) = -5.31$, $p < .001$, $d = .47$].

5.2.4. Desempeño en la prueba según sexo

5.2.4.1. Estadísticos descriptivos

Interesaba observar el desempeño según se tratara de niñas o niños, sobre la base de no esperar diferencias.

En la tabla 5.11 se presentan la media, el desvío estándar y la varianza de los puntajes observados en cada grupo según el sexo.

Tabla 5.11

Media, desvío estándar y varianza de los puntajes observados en cada grupo según el sexo

Sexo	N	Media	DE	Varianza
F	1545	28.8	13.9	194
M	1552	28.1	13.7	188

DE: desvío estándar; M: masculino; F: femenino

5.2.4.2. Pruebas de normalidad y homocedasticidad

La prueba de normalidad (Shapiro-Wilk) aplicada a los puntajes observados en los grupos según el sexo reveló que no se ajustaron a una distribución normal [niñas: $W(1545) = .99$, $p < .001$; niños: $W(1552) = .98$, $p < .001$]. Por otro lado, la prueba de Levene mostró homogeneidad de varianzas entre los puntajes observados en ambos grupos [$F(3095) = .030$, $p = .86$].

5.2.4.3. Análisis de las diferencias entre las puntuaciones

Teniendo en cuenta los resultados de las pruebas de normalidad, se aplicó la prueba U de Mann-Whitney para comparar el rendimiento de las niñas y los niños (ver el apartado A.2.4 del Anexo “Análisis no paramétricos”). No obstante, aquí se reportan los resultados de la prueba paramétrica correspondiente, que fueron congruentes con los de la no paramétrica.

Una prueba t para muestras independientes reveló que no existieron diferencias estadísticamente significativas entre los puntajes observados en los dos grupos [$t(3095) = -1.33$, $p = .18$].

5.2.5. Desempeño en la prueba según nivel de oportunidades educativas

5.2.5.1. Estadísticos descriptivos

En la tabla 5.12 se presentan la media, el desvío estándar y la varianza de los puntajes observados en cada grupo según el nivel de oportunidades educativas.

Tabla 5.12

Media, desvío estándar y varianza de los puntajes observados en cada grupo según el nivel de oportunidades educativas

NOE	N	Media	DE	Varianza
BA	847	21.4	12.4	154
ME	1870	30.6	13.3	176
AL	380	33.4	13.9	193

DE: desvío estándar; NOE: nivel de oportunidades educativas; BA: bajo; ME: medio; AL: alto

5.2.5.2. Pruebas de normalidad y homocedasticidad

Las pruebas de normalidad (Shapiro-Wilk) revelaron que los puntajes observados no se ajustaron a una distribución normal en ninguno de los grupos [NOE bajo: $W(847) = .97, p < .001$; NOE medio: $W(1870) = .99, p < .001$; NOE alto: $W(380) = .98, p < .001$]. Por otro lado, la prueba de Levene mostró que las varianzas entre los puntajes observados en los tres grupos no fueron homogéneas [$F(2, 3094) = 3.81, p = .022$].

5.2.5.3. Análisis de las diferencias entre las puntuaciones

En vista de los resultados de las pruebas de normalidad, se aplicó la prueba H de Kruskal-Wallis para comparar el rendimiento entre los grupos con distinto nivel de oportunidades educativas en cada grado escolar (ver el apartado A.2.5 del Anexo A “Análisis no paramétricos”). Ahora bien, dado que los resultados de esta prueba fueron congruentes con los de la prueba de ANOVA, a continuación se reportan estos últimos.

La prueba de ANOVA mostró que el nivel de oportunidades educativas tuvo un efecto estadísticamente significativo sobre el rendimiento, con un tamaño del efecto mediano a grande [$F(2, 3094) = 174, p < .001, \eta^2_p = .10$]. La tabla 5.13 muestra los resultados de las comparaciones *post hoc* utilizando la prueba de Games-Howell, en donde se observa que las diferencias entre todos los grupos alcanzaron significación estadística.

Tabla 5.13

Diferencias entre las medias de los puntajes observados entre los grupos según el nivel de oportunidades educativas (prueba de Games-Howell)

(I) NOE	(J) NOE	Diferencia de medias (I-J)	Error estándar	Sig.	Intervalo de confianza de 95%	
					Límite inferior	Límite superior
BA	ME	-9.20 [*]	.53	< .001	-10.4	-7.96
	AL	-12.0 [*]	.83	< .001	-13.9	-10.0
ME	BA	9.20 [*]	.53	< .001	7.96	10.4
	AL	-2.77 [*]	.78	.001	-4.59	-.95
AL	BA	12.0 [*]	.83	< .001	10.0	13.9
	ME	2.77 [*]	.78	.001	.95	4.59

*. La media de la diferencia es significativa en el nivel de .05.

NOE: nivel de oportunidades educativas; BA: bajo; ME: medio; AL: alto.

5.2.6. Desempeño en la prueba según el tipo de gestión

El estudio del desempeño según el tipo de gestión se llevó a cabo en los grupos que asistían a escuelas con nivel de oportunidades educativas medio, dado que este nivel incluyó escuelas de gestión estatal y de gestión privada. En cambio, no fue posible incluir niños que concurrieran a escuelas de gestión privada con nivel de oportunidades educativas bajo ni que asistieran a escuelas de gestión estatal con nivel de oportunidades educativas alto.

5.2.6.1. Estadísticos descriptivos

En la tabla 5.14 se muestran la media, el desvío estándar y la varianza de los puntajes observados en el grupo que asistía a escuelas con nivel de oportunidades educativas medio segmentado según el tipo de gestión.

Tabla 5.14

Media, desvío estándar y varianza de los puntajes observados en el grupo que asistía a escuelas con NOE medio segmentado según el tipo de gestión

Gestión	N	Media	DE	Varianza
Estatal	1176	29.5	13.0	170
Privada	694	32.5	13.5	181

5.2.6.2. Pruebas de normalidad y homocedasticidad

La prueba de normalidad (Shapiro-Wilk) reveló que los puntajes observados no se ajustaron a una distribución normal [gestión estatal: $W(1176) = .99, p < .001$; gestión privada: $W(694) = .98, p < .001$]. En cambio, la prueba de Levene mostró homogeneidad de varianzas entre ambos grupos [$F(1868) = 1.05, p = .31$].

5.2.6.3. Análisis de las diferencias entre las puntuaciones

En conformidad con los resultados obtenidos en la prueba de normalidad, se aplicó la prueba U de Mann-Whitney para estudiar las diferencias en el rendimiento de los grupos de gestión estatal y privada que asistían a escuelas con nivel de oportunidades educativas medio (ver el apartado A.2.6 del Anexo A “Análisis no paramétricos”). Sin embargo, puesto que los resultados de esta prueba fueron congruentes con los de la correspondiente prueba paramétrica, a continuación se reportan estos últimos.

Una prueba t para muestras independientes mostró una diferencia estadísticamente significativa con un tamaño del efecto pequeño del tipo de gestión sobre el rendimiento [$t(1868) = -4.77, p < .001, d = .23$].

5.2.7. Desempeño en la prueba según región geográfica

El estudio del desempeño en la prueba según la región geográfica se llevó a cabo en el grupo que asistía a escuelas de gestión estatal, distribuidas en cuatro regiones.

5.2.7.1. Estadísticos descriptivos

La tabla 5.15 muestra la media, el desvío estándar y la varianza de los puntajes observados en el grupo que asistía a escuelas de gestión estatal segmentado según la región.

Tabla 5.15

Media, desvío estándar y varianza de los puntajes observados en el grupo que asistía a escuelas de gestión estatal segmentado según la región

Región	N	Media	DE	Varianza
SUR	577	25.9	13.9	194
OESTE	787	26.2	13.0	169
NORTE	438	24.6	13.6	185
CABA	221	29.6	12.3	150

DE: desvío estándar; SUR: sur del Gran Buenos Aires; OES: oeste del Gran Buenos Aires; NOR: norte del Gran Buenos Aires; CABA: Ciudad Autónoma de Buenos Aires

5.2.7.2. Pruebas de normalidad y homocedasticidad

Las pruebas de normalidad (Shapiro-Wilk) revelaron que sólo los puntajes observados en el grupo de la Ciudad Autónoma de Buenos Aires se ajustaron a una distribución normal [sur del Gran Buenos Aires (en adelante, GBA): $W(577) = .98$, $p < .001$; oeste del GBA: $W(787) = .98$, $p < .001$; norte del GBA: $W(438) = .98$, $p < .001$; CABA: $W(221) = .99$, $p = .11$]. La prueba de Levene mostró que las varianzas de los puntajes observados entre los cuatro grupos por región no fue homogénea [$F(3, 2019) = 3.30$, $p = .020$].

5.2.7.3. Análisis de las diferencias entre las puntuaciones

De acuerdo con los resultados de las pruebas de normalidad y homocedasticidad, se comparó el rendimiento de los grupos que asistían a escuelas de gestión estatal entre las distintas regiones utilizando la prueba H de Kruskal-Wallis (ver el apartado A.2.7 del Anexo A “Análisis no paramétricos”). Sin embargo, puesto que los resultados de esta prueba fueron congruentes con los de la correspondiente prueba paramétrica, a continuación se reportan estos últimos.

Una prueba de ANOVA reveló diferencias estadísticamente significativas con un tamaño del efecto pequeño entre los rendimientos de los grupos según la región [$F(3, 2019) = 7.12$, $p < .001$, $\eta^2_p = .010$]. La tabla 5.16 muestra las comparaciones *post hoc* utilizando la prueba de Games-Howell que revelan un rendimiento significativamente mejor del grupo de

CABA en relación con el resto de los grupos, mientras que las diferencias de rendimiento entre estos últimos no alcanzó significación estadística.

Tabla 5.16

Diferencias entre las medias de los puntajes observados en el grupo que asistía a escuelas de gestión estatal segmentado según la región (prueba de Games-Howell)

(I) Región	(J) Región	Diferencia de medias (I-J)	Error estándar	Sig.	Intervalo de confianza de 95%	
					Límite inferior	Límite superior
SUR	OES	-.26	.74	.98	-2.17	1.65
	NOR	1.36	.40	.40	-.88	3.60
	CABA	-3.70 [*]	1.01	.002	-6.30	-1.10
OES	SUR	.26	.74	.98	-1.65	2.17
	NOR	1.62	.80	.18	-.44	3.67
	CABA	-3.44 [*]	.95	.002	-5.88	-1.00
NOR	SUR	-1.36	.87	.40	-3.60	.88
	OES	-1.62	.80	.18	-3.67	.44
	CABA	-5.06 [*]	1.05	< .001	-7.77	-2.35
CABA	SUR	3.70 [*]	1.01	.002	1.10	6.30
	OES	3.44 [*]	.95	.002	1.00	5.88
	NOR	5.06 [*]	1.05	< .001	2.35	7.77

*. La media de la diferencia es significativa en el nivel de .05.

SUR: sur del Gran Buenos Aires; OES: oeste del Gran Buenos Aires; NOR: norte del Gran Buenos Aires; CABA: Ciudad Autónoma de Buenos Aires.

5.2.8. Resumen de los resultados del estudio del desempeño en la prueba según grado escolar, edad, sexo, nivel de oportunidades educativas, tipo de gestión y región geográfica

Los resultados de los análisis realizados ponen de manifiesto las diferencias en las puntuaciones obtenidas de acuerdo con las distintas variables independientes incluidas, de modo tal que puede arribarse a un análisis de conjunto que permita señalar los efectos diferenciales en el desempeño, que se resumen en la tabla 5.17.

Tabla 5.17

Tamaño del efecto de las variables independientes sobre el desempeño en el TECLA Buenos Aires

Variables independientes	Tamaño del efecto	Magnitud	Estadístico	
			Paramétrico	No paramétrico
Grado	$\eta^2_p = .38$	grande	ANOVA	Kruskal-Wallis
Edad	$\eta^2_p = .34$	grande	ANOVA	Kruskal-Wallis
Edad (mismo grado)				
7 y 8 años (2º grado)	$d = .22$	pequeño	<i>t</i>	Mann-Whitney
11 y 12 años (6º grado)	$d = .20$	pequeño	<i>t</i>	Mann-Whitney
Grado (misma edad)				
2º y 3º grados (8 años)	$d = .13$	pequeño	<i>t</i>	Mann-Whitney
3º y 4º grados (9 años)	$d = .67$	mediano	<i>t</i>	Mann-Whitney
4º y 5º grados (10 años)	$d = .63$	mediano	<i>t</i>	Mann-Whitney
5º y 6º grados (11 años)	$d = .47$	pequeño	<i>t</i>	Mann-Whitney
Sexo: sin diferencias significativas				
NOE	$\eta^2_p = .10$	mediano	ANOVA	Kruskal-Wallis
Tipo gestión	$d = .23$	pequeño	<i>t</i>	Mann-Whitney
Región geográfica	$\eta^2_p = .010$	pequeño	ANOVA	Kruskal-Wallis

NOE: nivel de oportunidades educativas

En la tabla se observa el tamaño del efecto según la variable de la que se trate, que conduce al siguiente resumen de los principales hallazgos.

(a) El *grado escolar* tuvo un efecto estadísticamente significativo de tamaño grande, con un mejor rendimiento en los grados más avanzados ($\eta^2_p = .38$).

(b) El *nivel de oportunidades educativas* mostró un efecto estadísticamente significativo de tamaño mediano a grande, con un mejor rendimiento de los grupos de nivel alto ($\eta^2_p = .10$).

(c) El *tipo de gestión* tuvo un efecto estadísticamente significativo de tamaño pequeño con un mejor rendimiento del grupo que asistía a escuelas de gestión privada ($d = .23$).

(d) La *región* tuvo un efecto estadísticamente significativo de tamaño pequeño, con un mejor rendimiento del grupo de la Ciudad Autónoma de Buenos Aires ($\eta^2_p = .010$).

En virtud de estos resultados, se confeccionaron dos tipos de datos normativos:

1) Los baremos generales para la población de 2° a 6° grados de educación primaria de la Ciudad Autónoma de Buenos Aires y alrededores.

2) Los baremos para la misma población segmentada según el nivel de oportunidades educativas.

Las tablas correspondientes se presentan en el Anexo B “Baremos del TECLE Buenos Aires para la Ciudad Autónoma de Buenos Aires y alrededores”.

En el capítulo 6 “Discusión y conclusiones de la tesis” se ofrecen consideraciones para el uso de cada uno de ellos.

Capítulo 6

Discusión y conclusiones de la tesis

El objetivo central de la presente tesis fue determinar si el TECLE Buenos Aires constituye una herramienta de *screening* válida y confiable para evaluar la eficacia lectora en niños de 2° a 6° grados de educación primaria. Por otro lado, su uso eficaz en el medio local requiere contar con datos normativos apropiados, de modo tal que se han elaborado los baremos correspondientes.

De acuerdo con las recomendaciones propuestas por la Comisión Internacional de Tests para la traducción y adaptación de tests ya existentes (International Test Commission, 2017; Muñiz, Elosua, & Hambleton, 2013), se llevaron a cabo tres estudios: 1) estudio de confiabilidad, 2) estudio de validez, que contempla la validez de constructo y de criterio, y 3) estudio para la elaboración de datos normativos.

En primer lugar, se presenta una síntesis de los principales resultados de estos tres estudios. A continuación, se discute la interpretación y los usos propuestos para los puntajes observados en el test, en el marco del enfoque de la validación basado en el argumento. Seguidamente, se explicitan las limitaciones de la presente tesis. Por último, se ofrecen las conclusiones a las que se ha arribado.

6.1. Síntesis de los resultados

El estudio de *confiabilidad* mostró un alto grado de estabilidad temporal entre los puntajes observados en el test en dos instancias de evaluación separadas por un intervalo de 5 semanas, con un coeficiente de estabilidad $r = .90$.

El estudio de *validez de constructo* reveló que los puntajes observados en el TECLE Buenos Aires estuvieron más fuertemente asociados con las medidas de los aspectos específicos de la lectura, esto es, la *decodificación*

—que involucra la precisión y la velocidad en el reconocimiento ortográfico—, que con las que dependen de habilidades no específicas de la lectura, como la comprensión del lenguaje, la denominación automatizada rápida, el funcionamiento ejecutivo y la velocidad de procesamiento.

Por otro lado, el estudio de *validez de criterio* mostró una diferencia estadísticamente significativa con un tamaño del efecto mediano entre la media de los puntajes observados en el test en un grupo clínico de sujetos con diagnóstico de trastornos en el aprendizaje de la lectura, de la escritura, de las matemáticas y/o trastornos del lenguaje y la de los puntajes observados en un grupo control con desarrollo típico.

Finalmente, el estudio del efecto de un conjunto de variables sobre los puntajes observados en el test permitió identificar al nivel de oportunidades educativas como la variable que, junto con el grado escolar, explicó mejor la variación en los puntajes.

Como consecuencia de tales resultados, se confeccionaron dos tipos de datos normativos: los baremos generales para niños de 2° a 6° grados de la Ciudad Autónoma de Buenos Aires y alrededores y los baremos segmentados según el nivel de oportunidades educativas para cada grado escolar, que se ofrecen en el Anexo B “Baremos del TECLE Buenos Aires para la Ciudad Autónoma de Buenos Aires y alrededores”.

6.2. Discusión general

El objetivo de toda evaluación psicológica es obtener información acerca de algún constructo o atributo no observable en un sujeto evaluado, a partir de la valoración surgida de su operacionalización, relativa al rendimiento o conducta incluidos. La información obtenida motiva decisiones que afectan la vida de dicho sujeto y/o del grupo al que pertenece. Así, el procedimiento de evaluación involucra una cadena de afirmaciones o reglas (asunciones y/o inferencias) que conducen a la interpretación de los puntajes observados en términos de algún atributo no observable y a la implementación de acciones dirigidas a generar efectos sobre dicho atributo

que redunden en consecuencias positivas para el sujeto evaluado y/o el grupo al que pertenece.

El enfoque de la validación basado en el argumento (Cronbach, 1988; Cronbach & Meehl, 1955; Kane, 2013; Lane, 2014) ofrece un marco general para la validación de un procedimiento de evaluación, aun en ausencia de una teoría acabada del constructo de interés. Propone explicitar las asunciones e inferencias en dos tipos de argumentos: un argumento para la *interpretación* de los puntajes observados y un argumento para su *uso*.

El argumento para la *interpretación* explicita las afirmaciones acerca de la manifestación del constructo y su validación requiere del tipo de evidencia investigada en los estudios de validez de constructo y de criterio.

El argumento para el *uso* explicita los objetivos de las decisiones basadas en la interpretación. En el ámbito educativo, los usos incluyen la planificación o modificación de la enseñanza, el monitoreo del aprendizaje, la promoción o acreditación de cursos, y la ubicación de alumnos en programas de enseñanza tutorada. La validación del argumento de uso requiere, a su vez, de un estudio de las consecuencias de las acciones implementadas. La validez de la interpretación de los puntajes no justifica necesariamente cualquier acción motivada por dicha interpretación. Del mismo modo, la falta de validez de una decisión debido al hallazgo de consecuencias negativas inadmisibles en un contexto particular no invalida la interpretación de los puntajes que motivó dicha decisión, a menos que dichas consecuencias obedezcan a defectos no advertidos en el test o en el procedimiento de evaluación.

En cuanto a la presente tesis, si bien la validación del TECLE Buenos Aires no se desarrolló en el marco de la validación basada en el argumento, consideramos enriquecedor discutir los resultados presentados desde la perspectiva de un argumento de interpretación y un argumento de uso de los puntajes que el test arroja.

6.2.1. Interpretación de los puntajes observados en el TECLE Buenos Aires. Una medida de eficacia lectora

En un artículo acerca de la validez de las interpretaciones y los usos de los puntajes obtenidos en los tests, Kane (2013) sugirió incluir tres tipos de inferencias en el argumento de interpretación: de puntuación, de generalización y de extrapolación.

Las inferencias de puntuación refieren a los criterios de asignación de puntos a las conductas observadas. En el caso del TECLE Buenos Aires, se asigna un punto si el sujeto evaluado respondió al ítem seleccionando la palabra que completa coherentemente la oración dentro del tiempo límite establecido. Por el contrario, si luego de finalizado ese tiempo, el sujeto continuó completando ítems, se le asignan 0 puntos a dichos ítems, aun cuando su respuesta fuera correcta. Esto ejemplifica, nuevamente, la necesidad de una interpretación que no solo considere el puntaje, sino que tenga en cuenta circunstancias diferenciales de respuesta, ya que obviamente no es lo mismo acertar un ítem fuera del tiempo permitido que no poder hacerlo en absoluto.

Las inferencias de generalización refieren a un conjunto mayor de conductas, de las cuales las observadas en una instancia de evaluación constituyen una muestra representativa. Aquí se incluye el rendimiento en el test en otros contextos y condiciones de observación, o el rendimiento en una forma paralela del test. La validación de las inferencias de generalización involucra la búsqueda de evidencias de confiabilidad de los puntajes observados. El alto índice de confiabilidad obtenido para los puntajes observados en el TECLE a través del método test-retest permiten validar la generalización del rendimiento en una instancia de evaluación al rendimiento observado en otra instancia.

Las inferencias de extrapolación refieren a conductas teóricamente relacionadas con la conducta evaluada por el test, es decir que ponen en juego al mismo constructo subyacente y a la red nomológica de la que forma parte, e incluyen el rendimiento en otras tareas. En cuanto al TECLE Buenos

Aires, las inferencias de extrapolación incluyen la precisión y el tiempo observados en la lectura en voz alta de palabras y pseudopalabras, o el tiempo en la comprensión de textos.

Los autores de la versión original del TECLE (Marín & Carrillo, 1999, citado en Cuadro, Costa, Trías, & Ponce de León, 2009, y en Marín, Cuadro, & Pagán, 2007) propusieron una interpretación de los puntajes arrojados por el test en términos de “eficacia lectora”. De acuerdo con nuestros resultados, la eficacia lectora depende de las habilidades de decodificación, específicas para la lectura, que consisten en la precisión y la velocidad en el reconocimiento ortográfico, y en menor medida, como ya se señalara, de habilidades generales lingüísticas (semánticas y sintácticas) y cognitivas (funcionamiento ejecutivo y velocidad de procesamiento).

Asimismo, es de señalar, por una parte, que no existe una teoría acabada del constructo, ya que el término “eficacia” remite al rendimiento en una tarea. Por la otra, que el rendimiento en el TECLE Buenos Aires depende no sólo de las habilidades y estrategias idiosincrásicas del sujeto evaluado sino también de las características de los ítems, la modalidad de respuesta solicitada y las condiciones de evaluación.

No obstante, es factible enunciar de manera explícita las inferencias que constituyen el argumento de interpretación de los puntajes observados. Los resultados obtenidos en los estudios de validez de constructo y de criterio permiten validar la interpretación de tales puntajes como indicadores principalmente de las habilidades específicas, aunque también, pero en menor medida, de las habilidades generales lingüísticas y cognitivas.

Al respecto, Kane (2013) afirmó que la validación de un argumento de interpretación no requiere necesariamente de una explicación detallada de los procesos involucrados en la ejecución de una tarea. En cambio, es apropiado y factible formular inferencias que asocien el constructo de interés con las regularidades observadas en el rendimiento. Con todo, un cuarto tipo de inferencias más ambiciosas en el proceso de validación son las

inferencias causales. Este tipo de inferencias requieren en cambio, de una teoría consolidada que vincule el constructo subyacente con el rendimiento observado y pueden ser validadas empíricamente a través de la comparación de los efectos de la presencia del constructo sobre el rendimiento con los de la ausencia del mismo. En cuanto al TECLE Buenos Aires, las hipótesis acerca de las diferencias entre el desempeño de un grupo de sujetos con dificultades en el aprendizaje y el de otro grupo control de sujetos con desarrollo típico constituyen una inferencia causal para la cual se halló evidencias de validez en el estudio referido al criterio.

6.2.2. Uso propuesto de los puntajes observados en el TECLE Buenos Aires. Detección del riesgo de desarrollar dificultades en el aprendizaje de la lectura. Evaluación de la calidad de la enseñanza de la lectura

Kane (2013) señaló que la validación de los usos de los puntajes observados en un test requiere del estudio de las consecuencias de las acciones basadas en la interpretación de dichos puntajes, que incluya no solo los beneficios pretendidos, sino también los efectos adversos, siendo el objetivo del procedimiento de evaluación la obtención de consecuencias positivas para los sujetos evaluados y/o los grupos a los que pertenecen, a un costo aceptable y con consecuencias negativas limitadas.

En el caso del TECLE Buenos Aires, el uso propuesto de los puntajes es el de obtener rápidamente una *medida global de la eficacia lectora* en los alumnos. El objetivo de este uso es el de ajustar oportunamente los métodos y estrategias de enseñanza dirigidos a alcanzar los niveles de precisión y velocidad en la decodificación necesarios para lograr una buena comprensión lectora. Se inscribe en la extensa literatura sobre la utilidad de los tests de *screening* para una detección oportuna y una intervención eficaz.

Ahora bien, el estudio del efecto de un conjunto de variables sobre el desempeño en el test justificó la confección de dos tipos de datos normativos: baremos en función del grado escolar y en función del grado escolar y el nivel de oportunidades educativas. Al respecto, es necesario

advertir acerca de los posibles efectos adversos asociados al uso de un tipo de baremo u otro.

En primer lugar, asignarle a un niño un puntaje referido a normas como una medida global de su habilidad lectora requiere considerar que, tal como se señaló en el apartado 1.2 dedicado al aprendizaje de la lectura en los sistemas alfabéticos, dicha habilidad no sólo depende de las características del niño sino también de factores ambientales. Algunos de estos factores son compartidos en la muestra y en la población, como el grado de transparencia/opacidad del sistema alfabético y la complejidad silábica de la lengua oral, mientras que los hábitos de lectura en el entorno de crianza, y el método de enseñanza de la lectura varían de un sujeto a otro, de un aula a otra, de una escuela a otra y de un distrito escolar a otro. De modo que los puntajes observados en el test no deberían ser utilizados como indicadores diagnósticos de un déficit cognitivo de naturaleza constitucional. En cambio, es apropiado el uso del test como indicador de un riesgo de desarrollar o presentar dificultades en el aprendizaje de la lectura. En tal caso, deben extremarse las precauciones para asegurar que la presencia de dicho riesgo derive en una decisión de ajustar la enseñanza de la lectura con la implementación de métodos y estrategias probadamente eficaces. De esta manera, se espera obtener beneficios para el curso del aprendizaje en aquellos niños detectados, así como en el conjunto de los niños.

En segundo lugar, el uso de los baremos generales para cada grado escolar puede dar lugar a una alta tasa de detecciones del riesgo en niños que provienen de entornos de crianza de mayor vulnerabilidad y asisten a escuelas con bajo nivel de oportunidades educativas, al compararlos con un grupo heterogéneo conformado por sujetos provenientes de diferentes entornos socio-económico-culturales. Si en estos casos, la escuela/el sistema educativo no asume su rol en la prevención de las dificultades, es decir no ajusta las prácticas de enseñanza de la lectura con métodos probadamente eficaces, una consecuencia posible es la derivación innecesaria de un número considerable de niños a una instancia de

evaluación más detallada de las habilidades cognitivas en el ámbito clínico, fuera del ámbito escolar. De esta manera, se genera una secuencia de posibles efectos adversos que inicia con la sobrecarga de las familias en cuanto al tiempo y/o el dinero destinado a la búsqueda de un diagnóstico y tratamiento adecuados, la demora en la conformación de un léxico ortográfico, la pérdida de tiempo de exposición a la lectura, la imposibilidad de lograr niveles de fluidez adecuados para la comprensión lectora, e incluso el fracaso escolar.

En los niños provenientes de entornos de crianza favorables y que asisten a escuelas con alto nivel de oportunidades educativas, en cambio, el uso de los baremos generales puede dar lugar a una baja tasa de detecciones del riesgo. En estos casos, los efectos adversos pueden afectar a los niños que poseen déficits que ameritan una evaluación diagnóstica y un tratamiento, pero que no son derivados debido a que el puntaje que obtienen en el test no es lo suficientemente bajo en comparación con los de su grupo heterogéneo de referencia. Como es de esperar, la secuencia de posibles efectos adversos puede ser semejante a la de los niños provenientes de ambientes desfavorecidos: demora en el inicio de un diagnóstico y tratamiento adecuados, retraso en la conformación de un léxico ortográfico, pérdida de tiempo de exposición a la lectura, imposibilidad de lograr niveles de fluidez necesarios para la comprensión lectora y fracaso escolar.

En tercer lugar, el uso de los baremos para cada grado escolar segmentados en función del nivel de oportunidades educativas permite comparar a los niños con un grupo de referencia más homogéneo que los representa en mayor medida que un grupo más heterogéneo. Sin embargo, un posible efecto adverso de este uso de los puntajes, particularmente en los grupos de bajo nivel de oportunidades educativas, es que directamente evade la cuestión de la contribución general del entorno y específica de la escuela al bajo rendimiento, y por lo tanto, del rol del sistema educativo en la prevención de las dificultades.

Por lo expuesto, consideramos que el uso de los puntajes observados en el test para la detección del riesgo de presentar o desarrollar dificultades en el aprendizaje de la lectura debería basarse en la interpretación referida a los baremos segmentados en función del grado escolar y el nivel de oportunidades educativas. Complementariamente, recomendamos ubicar el desempeño de cada niño en relación con el de sus compañeros de curso.

Por otro lado, el uso de los puntajes para la evaluación de la calidad de las prácticas de enseñanza de la lectura debería basarse en la interpretación referida a los baremos generales segmentados en función del grado escolar.

En ambos casos, la validación del argumento de uso requiere considerar la utilidad de cada interpretación de los puntajes en función del logro de los objetivos planteados previamente a la evaluación y de la relación entre los beneficios obtenidos y los posibles efectos adversos.

Por último, Kane (2013) plantea que la responsabilidad de la evaluación de los usos de los puntajes de un test es compartida entre los diseñadores del test y los usuarios del mismo. Por un lado, los diseñadores del test proponen una interpretación y uno o varios usos de los puntajes. Por otro lado, los usuarios del test, es decir los profesionales de los ámbitos educativo y clínico que interpretan los puntajes observados en las instancias reales de evaluación y toman decisiones informadas por la interpretación, están en una mejor posición para valorar los beneficios y los efectos adversos de esas decisiones en los distintos contextos de aplicación. Por lo tanto, deberían asumir la mayor responsabilidad.

6.3. Limitaciones de la tesis

La principal limitación de la presente tesis afecta a la representatividad de las muestras conformadas para los estudios de confiabilidad, de validez y para la confección de los baremos, dada la imposibilidad de aplicar técnicas de muestreo probabilístico. En primer lugar, esto se debió a las dificultades para obtener una entrevista y/o la autorización de una parte de las autoridades distritales y de los establecimientos consultados y, en menor

medida, el consentimiento informado por escrito de los padres de los alumnos. En segundo lugar, no se contó con los datos demográficos acerca de la distribución de las variables de nivel de oportunidades educativas, tipo de gestión y región geográfica.

Esta limitación es particularmente relevante para la confección de los datos normativos. No obstante, es posible que el tamaño de las muestras reunidas para cada grupo según el grado escolar y el nivel de oportunidades educativas haya mitigado el efecto de dicha limitación (N= 88 a 405), a excepción de los tamaños alcanzados para 3° y 5° grados con alto nivel de oportunidades educativas (N = 36 y N = 21, respectivamente).

6.4. Proyección para futuras investigaciones

Los resultados encontrados, más allá de su interés intrínseco, nos permiten plantear, en primer lugar, que es necesario realizar estudios de validez de las consecuencias del uso propuesto de los puntajes observados. En segundo lugar, los estudios de validez predictiva dirigidos a determinar la sensibilidad y la especificidad del TECLE Buenos Aires como medida predictora del futuro desempeño en lectura y escritura y de los eventuales trastornos en el aprendizaje que puedan aparecer, sin duda, contribuirán al logro del objetivo general de la presente tesis.

6.5. Conclusiones

A la luz de los resultados obtenidos en los tres estudios que integran la presente tesis, concluimos que:

- 1) Los puntajes observados en el TECLE Buenos Aires muestran un alto grado de estabilidad temporal, lo cual nos permite afirmar que la influencia del error de medición fue aceptablemente baja.
- 2) Los puntajes observados en el TECLE Buenos Aires pueden interpretarse principalmente como una manifestación de las habilidades de decodificación, esto es, precisión y velocidad en el

reconocimiento ortográfico. Efectivamente, los estudios de validez, aún con las limitaciones que se señalaran, adicionan evidencia empírica acerca del alcance del constructo “eficacia lectora”. En los resultados encontrados aparece con nitidez que dicho constructo depende, en mayor medida de tales habilidades y en menor medida, de habilidades generales lingüísticas (semánticas y sintácticas) y cognitivas (funcionamiento ejecutivo y velocidad de procesamiento). Esta circunstancia pone de manifiesto la importancia de la interpretación adecuada de los puntajes en relación con los conocimientos y habilidades subyacentes, a la hora de diseñar estrategias de enseñanza adecuadas.

- 3) La variación en los puntajes observados en el TECLE Buenos Aires en sujetos de 2° a 6° grados con desarrollo típico estuvieron influidas principalmente por el grado escolar y el nivel de oportunidades educativas.
- 4) En función de ello, la disponibilidad de baremos segmentados en función del grado escolar para toda la muestra y de los particularizados en función del grado escolar y el nivel de oportunidades educativas, resulta particularmente útil para una interpretación adecuada de los puntajes. Para ello se proporcionan recomendaciones acerca de la pertinencia y objetivos del uso de uno u otro de ellos.

En conclusión, el TECLE Buenos Aires constituye, por una parte, una herramienta válida y confiable para la identificación del riesgo de desarrollar o presentar dificultades en el aprendizaje de la lectura. Por la otra, se presenta como una herramienta confiable para evaluar la efectividad de las prácticas de enseñanza de la lectura. Se adiciona a la literatura creciente sobre los instrumentos de *screening* que posibilitan la obtención de indicadores para la detección temprana y la prevención y/o intervención oportuna destinada a optimizar los procesos de aprendizaje del lenguaje escrito.

Referencias

- Abusamra, V., Ferreres, A., Raiter, A., De Beni, R., & Cornoldi, C. (2010). *Test Leer para Comprender TLC: evaluación de la comprensión de textos*. Buenos Aires: Paidós.
- Anastasi, A. & Urbina, S. (1998). *Tests Psicológicos*. México: Prentice Hall.
- Andrade, O. V. C. A., Andrade, P. E., & Capellini, S. A. (2015). Collective screening tools for early identification of dyslexia. *Frontiers in Psychology, 5*, 1-13. DOI: 10.3389/fpsyg.2014.01581
- Asociación Americana de Psiquiatría (1995). *DSM-IV. Manual diagnóstico y estadístico de los trastornos mentales*. Barcelona: Masson.
- Asociación Americana de Psiquiatría (2013). *Guía de consulta de los criterios diagnósticos del DSM 5*. Arlington, VA: Asociación Americana de Psiquiatría.
- Ato, M., López, J. J., & Benavente, A. (2013). Un sistema de clasificación de los diseños de investigación en psicología. *Anales de Psicología, 29*(3), 1038-1059. DOI: 10.6018/analesps.29.3.178511
- Byrne, B. (2005). Theories of learning to read. En M. J. Snowling & C. Hulme (Eds.), *The Science of Reading: A Handbook* (pp. 104-119). Oxford, UK: Blackwell. DOI: 10.1002/9780470757642.ch6
- Carillo, M. & Alegría, J. (2009). Mecanismos de identificación de palabras en niños disléxicos en español: ¿existen subtipos?. *Ciencias Psicológicas, III*(2), 135-152. Recuperado de <http://www.scielo.edu.uy/pdf/cp/v3n2/v3n2a03.pdf>

- Carrillo, M. S., Alegría, J., & Calvo, A. R. (2008). Subtipos de dislexia en una ortografía transparente. *Actas del XXIV Congreso Internacional de AESLA. Aprendizaje de lenguas, uso del lenguaje y modelación cognitiva: perspectivas aplicadas entre disciplinas*, 819-825. Recuperado de <http://www.um.es/lacell/aesla/contenido/pdf/8/carrill2.pdf>
- Carrillo, M. S., Alegría, J., & Marín, J. (2013). On the acquisition of some basic word spelling mechanisms in a deep (French) and a shallow (Spanish) system. *Reading and Writing*, 26(6), 799-819. DOI: 10.1007/s11145-012-9391-6
- Clemens, N. H., Shapiro, E. S. & Thoemmes, F. (2011). Improving the efficacy of first grade reading screening: An investigation of word identification fluency with other early literacy indicators. *School Psychology Quarterly*, 26, 231–244. DOI: 10.1037/a0025173
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd edition). Hillsdale, NJ: Lawrence Erlbaum Associates. Recuperado de <http://www.utstat.toronto.edu/~brunner/oldclass/378f16/readings/Cohen Power.pdf>
- Cohen, J. (1992). A power primer. *Psychological Bulletin*, 112(1), 155-159. DOI: 10.1037/0033-2909.112.1.155
- Coltheart, M. (2005). Modeling reading: The dual-route approach. En M. J. Snowling & C. Hulme (Eds.), *The Science of Reading: A Handbook* (pp. 6-23). Oxford, UK: Blackwell. DOI: 10.1002/9780470757642.ch1
- Coltheart, M., Rastle, K., Perry, C., Langdon, R., & Ziegler, J. (2001). DRC: A dual route cascaded model of visual word recognition and reading aloud. *Psychological Review*, 108(1), 204-256. DOI: 10.1037//0033-295X.108.1.204

- Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, 16(3), 297-334. DOI: 10.1007/BF02310555
- Cronbach, L. J. (1988). Five perspectives on validity argument. En H. Wainer & H. Braun (Eds.), *Test validity* (pp. 3-17), Hillsdale, NJ: Erlbaum.
- Cronbach, L. J. & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin*, 52(4), 281-302. DOI: 10.1037/h0040957
- Cuadro, A., Costa, D. Trías, D. & Ponce de León, P. (2009). *Evaluación del nivel lector. Manual técnico del test de eficacia lectora (TECLE) de J. Marín y M. Carrillo*. Montevideo: Prensa Médica Latinoamericana.
- Cuetos, F., Rodríguez, B., Ruano, E., & Arribas, D. (2009). *PROLEC-R. Batería de Evaluación de los Procesos Lectores, Revisada*. Madrid: TEA Ediciones.
- De Beni, R., Cornoldi, C., Carretti, B., & Meneghetti, B. (2003). *Nuova Guida alla Comprensione del Testo. Volume 1*. Trento, Italia: Erickson.
- Defior, S. & Serrano, F. (2011). Procesos fonológicos explícitos e implícitos, lectura y dislexia. *Revista Neuropsicología, Neuropsiquiatría y Neurociencias*, 11(1), 79-94. Recuperado de <https://dialnet.unirioja.es/servlet/articulo?codigo=3640858>
- Defior Citoler, S. (1996). *Las dificultades del aprendizaje: Un enfoque cognitivo*. Málaga: Aljibe.
- Defior Citoler, S., Fonseca, L., Gottheil, B., Aldrey, A., Jiménez Fernández, G., Pujals, G. R., & Serrano Chica, F. D. (2006). *LEE Test de lectura y escritura en español*. Buenos Aires: Paidós.
- Démonet, J.-F., Taylor, M. J., & Chaix, Y. (2004). Developmental dyslexia. *Lancet*, 303, 1451-1460. DOI: 10.1016/S0140-6736(04)16106-0

Denckla, M. B. & Rudel, R. G. (1974). Rapid "automatized" naming of pictured objects, colors, letters and numbers by normal children. *Cortex*, 10, 186-202. DOI: 10.1016/S0010-9452(74)80009-2

Dirección Nacional de Información y Evaluación de la Calidad Educativa. Ministerio de Educación de la Nación. (2007). Estudio Nacional de Evaluación y Consideraciones Conceptuales. Educación Primaria. Educación Secundaria. Operativo Nacional de Evaluación 2007. Recuperado de <http://www.bnm.me.gov.ar/giga1/documentos/EL005285.pdf>

Dirección Nacional de Información y Evaluación de la Calidad Educativa. Ministerio de Educación de la Nación. (2010). Pautas Pedagógicas para Elaborar Evaluaciones Diagnósticas. Ciencias Sociales. Lengua. 3° Año de la Educación Primaria. Recuperado de <http://www.bnm.me.gov.ar/giga1/documentos/EL003248.pdf>

Dirección Nacional de Información y Evaluación de la Calidad Educativa. Ministerio de Educación de la Nación. (2010). Recomendaciones Metodológicas para la Enseñanza. Lengua. Educación Primaria-ONE 2007. Pruebas de 3° año y 6° año Primaria. Recuperado de <http://repositorio.educacion.gov.ar:8080/dspace/bitstream/handle/123456789/109673/RM-LENGUA-primaria.pdf>

Ehri, L. C. (1999). Phases of development in learning to read words. En J. Oakhill & R. Beard (Eds.), *Reading development and the teaching of reading. A psychological perspective* (pp. 79-108). Oxford, Inglaterra: Blackwell.

Elliott, J. G. & Grigorenko, E. L. (2014). *Cambridge studies in cognitive and perceptual development. The dyslexia debate*. Nueva York, EEUU: Cambridge University Press.

Feldman, J. (2009). *Escala Alfa-2 y Gamma para la evaluación del lenguaje en niños*. Buenos Aires: CEC Publicaciones.

- Ferreres, A., Abusamra, V., Casajús, A., & China, N. (2011). Adaptación y estudio preliminar de un test breve para evaluar la eficacia lectora (TECLE). *Neuropsicología Latinoamericana*, 3(1), 1-7. DOI: 10.5579/rnl.2011.0040
- Ferreres, A., Abusamra, V., Squillace, M., Fernández Liporace, M., Cartoceti, R., & Sampedro, B. (2010). Análisis psicométricos, datos normativos y otros resultados del test Leer para Comprender. En V. Abusamra, A. Ferreres, A. Raiter, R. De Beni, & C. Cornoldi (Au.), *Test Leer para Comprender TLC: evaluación de la comprensión de textos* (pp. 131-154). Buenos Aires: Paidós.
- Fletcher, J. M., Lyon, G. R., Fuchs, L. S., & Barnes, M. A. (2007). *Learning Disabilities: From Identification to Intervention*. New York, NY: The Guildford Press.
- Fletcher, J. M., Simos, P. G., Papanicolaou, A. C., & Denton, C. (2004). Neuroimaging in reading research. En N. Duke & M. Mallette (Eds.), *Literacy research methods* (pp. 252–286). New York: Guilford Press.
- Frith, U. (1985). Beneath the surface of developmental dyslexia. En K. Patterson, J. Marshall, & M. Coltheart (Eds.), *Surface Dyslexia, Neuropsychological and Cognitive Studies of Phonological Reading* (pp. 301-330). Londres: Erlbaum.
- Frith, U. (1995). Dyslexia: Can we have a shared theoretical framework? *Educational and Child Psychology*, 12, 6-17.
- Frost, R. (2005). Orthographic systems and skilled word recognition process in reading. En M. J. Snowling & C. Hulme (Eds.), *The Science of Reading: A Handbook* (pp. 272-295). Oxford, UK: Blackwell. DOI: 10.1002/9780470757642.ch15

- Fuchs, D. & Fuchs, L. S. (2006). Introduction to response to intervention: What, why, and how valid is it? *Reading Research Quarterly*, 41, 93-99. DOI: 10.1598/RRQ.41.1.4
- Furr, R. M. & Bacharach, V. R. (2014). *Psychometrics: An Introduction* (2nd edition). Thousand Oaks, CA, US: SAGE Publications, Inc.
- Georgiou, G. K., Parrila, R., & Papadopoulos, T. C. (2008). Predictors of word decoding and reading fluency across languages varying in orthographic consistency. *Journal of Educational Psychology*, 100(3), 566-580. DOI: 10.1037/0022-0663.100.3.566
- Hanley, J. R. (2005). Learning to read in Chinese. En M. J. Snowling & C. Hulme (Eds.), *The Science of Reading: A Handbook* (pp. 316-335). Oxford, UK: Blackwell. DOI: 10.1002/9780470757642.ch17
- Hoover, W. A. & Gough, P. B. (1990). The simple view of reading. *Reading and Writing: An Interdisciplinary Journal*, 2, 127-160. DOI: 10.1007/BF00401799
- International Test Commission (2017). *The ITC Guidelines for Translating and Adapting Tests (Second Edition)*. Recuperado de https://www.intestcom.org/files/guideline_test_adaptation_2ed.pdf
- Jiménez, J. E. (2012). Dislexia en español. Prevalencia e indicadores cognitivos, culturales, familiares y biológicos. Madrid: Pirámide.
- Jiménez, J. E., Rodríguez, C., & Ramírez, G. (2009). Spanish developmental dyslexia: Prevalence, cognitive profile, and home literacy experiences. *Journal of Experimental Child Psychology*, 103, 167-185. DOI: 10.1016/j.jecp.2009.02.004
- Kane, M. T. (2013). Validating the interpretations and uses of test scores. *Journal of Educational Measurement*, 50(1), 1-73. DOI: 10.1111/jedm.12000

- Kirk, S. A., & Bateman, B. (1962). Diagnosis and remediation of learning disabilities. *Exceptional Children, 29*, 73-78.
- Lane, S. (2014). Validity evidence based on testing consequences. *Psicothema, 26*(1), 127-135. DOI: 10.7334/psicothema2013.258
- Ledesma, R., Macbeth, G., & Cortada De Kohan, N. (2008). Tamaño del Efecto: Revisión teórica y aplicaciones con el Sistema Estadístico Vista. *Revista Latinoamericana de Psicología, 40*(3), 425-439.
- Lobrot, M. (1973). *Lire avec épreuves pour évaluer la capacité de lecture (D-OR-LEC)*. Paris: Editions ESF.
- Lyon, G. R. (1995). Toward a definition of dyslexia. *Annals of Dyslexia, 45*, 3-27. DOI: 10.1007/BF02648210
- Lyon, G. R., Shaywitz, S. E., & Shaywitz, B. A. (2003). A definition of dyslexia. *Annals of Dyslexia, 53*, 1-14. DOI: 10.1007/s11881-003-0001-9
- Marín, J., & Carrillo, M. (1999). *Test colectivo de eficacia lectora (TECLE)*. Manuscrito no publicado. Departamento de Psicología Básica y Metodología. Universidad de Murcia.
- Marín, J., Cuadro, A., & Pagán, A. (2007). Léxico ortográfico y competencia lectora. *Ciencias Psicológicas, 1*(1), 15-26. Recuperado de <http://www.redalyc.org/pdf/4595/459545423003.pdf>
- Mehler, J., Dommergues, J. I., Frauenfelder, U., & Seguí, J. (1981). The syllable's role in speech segmentation. *Journal of Verbal Learning and Verbal Behavior, 20*, 298-305. DOI: 10.1016/S0022-5371(81)90450-3
- Moats, C. L. (2010). *Speech to print. Language essentials for teachers. Second edition*. Baltimore: Paul Brookes Publishing.

- Morais, J., Bertelson, P., Cary, L., & Alegría, J. (1986). Literacy training and speech segmentation. *Cognition*, 24, 45-64. DOI: 10.1016/0010-0277(86)90004-1
- Morello García, F., China, N., & Ferreres, A. (2015). *Test de Denominación Automatizada Rápida (test DAR)*. Manuscrito en preparación.
- Muñiz, J. Elosua, P, & Hambleton, R. K. (2013). Directrices para la traducción y adaptación de los tests: segunda edición. *Psicothema*, 25(2), 151-157. DOI: 10.7334/psicothema2013.24.
- Olaz, F. (2008). Confiabilidad. En S. Tornimbeni, E. Pérez, & F. Olaz (Au.), *Introducción a la psicometría* (pp. 71-99). Buenos Aires: Paidós.
- Olaz, F. & Tornimbeni, S. (2008). Interpretación de puntuaciones. En S. Tornimbeni, E. Pérez, & F. Olaz (Au.), *Introducción a la psicometría* (pp. 137-159). Buenos Aires: Paidós.
- Perfetti, C. (2007). Reading ability: Lexical quality to comprehension. *Scientific Studies of Reading*, 11(4), 357-383. DOI: 10.1080/10888430701530730
- Petscher, Y. & Kim, Y.-S. (2011). Efficiency of predicting risk in word reading using fewer, easier letters. *Assessment for Effective Intervention*, 37(1), 17–25. DOI: 10.1177/1534508411407761
- Piacente, T. (2012). Alfabetización inicial y alfabetización académica. Investigación y evaluación de los procesos implicados en el aprendizaje del lenguaje escrito. *Revista Iberoamericana de Diagnóstico y Evaluación – e Avaliação Psicológica*, 33, 9-30.
- Piacente, T., Marder, S., & Resches, M. (2008). *Condiciones de la Familia y el niño para la alfabetización. Evaluación de Impacto del Plan Más Vida*. Buenos Aires: Comisión de Investigaciones Científicas de la Provincia de Buenos Aires – CIC.

- Real Academia Española & Asociación de Academias de la Lengua Española (2010). *Ortografía de la lengua española*. Madrid: Espasa.
- Restori, A. F., Katz, G. S., & Lee, H. B. (2009). A critique of the IQ / achievement discrepancy model for identifying specific learning disabilities. *Europe's Journal of Psychology*, 4, 128-145. DOI: 10.5964/ejop.v5i4.244
- Richardson, J. T. E. (2011). Eta squared and partial eta squared as measures of effect size in educational research. *Educational Research Review*, 6, 135-147. DOI: 10.1016/j.edurev.2010.12.001
- Sampson, G. (1997). *Sistemas de escritura. Análisis lingüístico*. Barcelona: Gedisa
- Scarborough, H. S. (2001). Connecting early language and literacy to later reading (dis)abilities: Evidence, theory, and practice. En S. Neuman & D. Dickinson (Eds.), *Handbook for research in early literacy* (pp. 97-110). New York: Guilford Press.
- Secretaría de Evaluación Educativa. Ministerio de Educación y Deportes. Presidencia de la Nación (2017). *Aprender 2016. Análisis de desempeños por capacidades y contenidos. Nivel primario*. Recuperado de www.bnm.me.gov.ar/giga1/documentos/EL005596.pdf
- Seymour, P. H. K., Aro, M., & Erskine, J. M. (2003). Foundation literacy acquisition in European orthographies. *British Journal of Psychology*, 94, 143-174. DOI: 10.1348/000712603321661859
- Share, D. L. (1995). Phonological recoding and self-teaching: *sine qua non* of reading acquisition. *Cognition*, 55, 151-218. DOI: 10.1016/0010-0277(94)00645-2

- Shaywitz, S. E., Fletcher, J. M., Holahan, J. M., Shneider, A. E., Marchione, K. E., Stuebing, K. K., Francis, D. J., Pugh, K. R., & Shaywitz, B. A. (1999). Persistence of dyslexia: The Connecticut Longitudinal Study at adolescence. *Pediatrics*, *104*, 1351-1359. DOI: 10.1542/peds.104.6.1351
- Shaywitz, S. E., Morris, R., & Shaywitz, B. A. (2008). The education of dyslexic children from childhood to young adulthood. *Annual Review of Psychology*, *59*, 451-475. DOI: 10.1146/annurev.psych.59.103006.093633
- Snow, C. E., Burns, M. & Griffin, P. (Eds.) (1998). *Preventing reading difficulties in young children: Committee on Prevention of Reading Difficulties in Young Children, Commission on Behavioral and Social Sciences and Education*. National Research Council, Washington, DC: National Academy Press. DOI: 10.17226/6023.
- Snow, C. E. & Juel, C. (2005). Teaching children to read: What do we know about how to do it? En M. J. Snowling & C. Hulme (Eds.), *The Science of Reading: A Handbook* (pp. 501-520). Oxford, UK: Blackwell. DOI: 10.1002/9780470757642.ch26
- Snowling, M. J. (2000). The phonological representations hypothesis. En M. J. Snowling (Au.), *Dyslexia* (Second edition, pp. 29-61). Oxford, UK: Blackwell.
- Stage, S. A., Abbott, R. D., Jenkins, J. R., & Berninger, V. W. (2003). Predicting response to early reading intervention from verbal IQ, reading-related language abilities, attention ratings, and verbal IQ-word reading discrepancy: Failure to validate the discrepancy method. *Journal of Learning Disabilities*, *36*, 24-33. DOI: 10.1177/00222194030360010401

- Stanovich, K. E. (1993). Romance and reality. *The Reading Teacher*, 47(4), 280-291. Recuperado de http://www.keithstanovich.com/Site/Research_on_Reading_files/RdTch93.pdf
- Strauss, E., Sherman, E. M. S., Spreen, O., & Slick, D. J. (2006). Psychometrics in neuropsychological assessment. En E. Strauss, E. M. S. Sherman, & O. Spreen (Au.), *A Compendium of Neuropsychological Tests: Administration, Norms, and Commentary* (3rd edition, pp. 3-43). New York: Oxford University Press.
- Strauss, E., Sherman, E. M. S., & Spreen, O. (2006). Attention. En E. Strauss, E. M. S. Sherman, & O. Spreen (Au.), *A Compendium of Neuropsychological Tests: Administration, Norms, and Commentary* (3rd edition, pp. 546-677). New York: Oxford University Press.
- Torgesen, J. & Mathes, P. (2000). *Assesment and instruction in phonological awareness*. Florida Department of Education. Florida State University.
- Van Orden, G. C. & Kloos, H. (2005). The question of phonology and reading. En M. J. Snowling & C. Hulme (Eds.), *The Science of Reading: A Handbook* (pp. 61-78). Oxford, UK: Blackwell. DOI: 10.1002/9780470757642.ch4
- Vaughn, S., Linan-Thompson, S., & Hickman, P. (2003). Response to treatment as a means of identifying students with reading/learning disabilities. *Exceptional Children*, 69, 391–409. DOI: 10.1177/001440290306900401
- Vellutino, F. R. & Fletcher, J. M. (2005). Developmental dyslexia. En M. J. Snowling & C. Hulme (Eds.), *The Science of Reading: A Handbook* (pp. 362-378). Oxford, UK: Blackwell. 10.1002/9780470757642.ch19

- Vellutino, F. R., Scanlon, D. M., & Jaccard, J. (2003). Toward distinguishing between cognitive and experiential deficits as primary sources of difficulty in learning to read: A two-year follow-up to difficult to remediate and readily remediated poor readers. En B. R. Foorman (Ed.), *Preventing and remediating reading difficulties* (pp. 73–120). Baltimore: York Press.
- Vellutino, F. R., Scanlon, D. M., Zhang, H., & Schatschneider, C. (2008). Using response to kindergarten and first grade intervention to identify children at-risk for long-term reading difficulties. *Reading and Writing: An Interdisciplinary Journal*, 21(4), 437–480. DOI: 10.1007/s11145-007-9098-2
- Venezky, R. L. (1999). *The American Way of Spelling: The Structure and Origins of American English Orthography*. Nueva York: The Guilford Press.
- Williams, J., Rickert, V., Hogan, J., Zolten, A. J., Satz, P., D'Elia, L. F., Asarnow, R. F., Zaucha, K., & Light, R. (1995). Children's Color Trails. *Archives of Clinical Neuropsychology*, 10(3), 211-223. DOI: 10.1016/0887-6177(94)00041-N
- Wilson, S. B. & Lonigan, C. J. (2010). Identifying preschool children at risk of later reading difficulties: Evaluation of two emergent literacy screening tools. *Journal of Learning Disabilities*, 43(1), 62–76. DOI: 10.1177/0022219409345007
- Ziegler, J. C., & Goswami, U. (2005). Reading acquisition, developmental dyslexia, and skilled reading across languages: A psycholinguistic grain size theory. *Psychological Bulletin*, 131, 3-29. DOI: 10.1037/0033-2909.131.1.3

Anexo A Análisis no paramétricos

En el presente anexo, se informan los resultados de las pruebas no paramétricas realizadas en los estudios en los que la distribución de los datos no se ajustó a una distribución normal. Los resultados de estas pruebas fueron congruentes con los de las pruebas paramétricas correspondientes, los cuales se informan en los respectivos capítulos.

A.1. Estudio de confiabilidad test-retest

La mediana y el rango intercuartílico de los puntajes observados fueron de 31 y 20 en la instancia de test, y de 37 y 19 en la instancia de retest, respectivamente.

La tabla A.1. muestra la mediana y el rango intercuartílico de los puntajes observados en cada grupo según el grado escolar.

Tabla A.1

Mediana y rango intercuartílico de los puntajes observados en el test y el retest por grado escolar

Grado	Instancia	N	Mediana	RIC
3	Test	56	15.5	8
	Retest	56	22.5	14.5
4	Test	54	29	12
	Retest	54	35.5	15
5	Test	72	37	13
	Retest	72	41	12.5
6	Test	36	44	12.5
	Retest	36	51.5	14.5

RIC: rango intercuartílico

La prueba *tau-b* de Kendall aplicada a los puntajes observados en ambas instancias de evaluación en la muestra completa reveló una correlación positiva estadísticamente significativa [$\tau_b(218) = .76, p < .001$]. Teniendo en cuenta que el coeficiente *tau-b* de Kendall resulta típicamente de menor magnitud que el coeficiente *r* de Pearson, y la sugerencia de

Cohen (1988, 1992) acerca de los valores críticos de $r = .10$, $r = .30$ y $r = .50$ como indicadores de tamaños del efecto pequeño, mediano y grande, respectivamente, consideramos que los valores obtenidos de τ_u -b indicaron que la fuerza de las asociaciones fue de tamaño grande.

La misma prueba aplicada a los puntajes observados en ambas instancias de evaluación en cada grado arrojó correlaciones positivas estadísticamente significativas con tamaños del efecto grandes [3° grado: $\tau_b(56) = .58$, $p < .001$; 4° grado: $\tau_b(54) = .70$, $p < .001$; 5° grado: $\tau_b(72) = .54$, $p < .001$; 6° grado: $\tau_b(36) = .81$, $p < .001$].

A.2. Estudio del efecto de variables sobre el rendimiento en el test en la muestra para la obtención de los baremos

A.2.1. Efecto del grado escolar sobre el rendimiento en el test

La tabla A.2 muestra la mediana y el rango intercuartílico de los puntajes observados en los grupos según el grado escolar.

Tabla A.2

Mediana y rango intercuartílico de los puntajes observados en los grupos según el grado escolar

Grado	N	Mediana	RIC
2°	602	14	12
3°	643	21	16
4°	703	30	15
5°	498	35	15
6°	651	39	17

RIC: rango intercuartílico

La prueba H de Kruskal-Wallis aplicada a los puntajes observados en los grupos según el grado escolar reveló diferencias estadísticamente significativas [$\chi^2(4) = 1210$, $p < .001$] con un promedio de rangos de 698 para 2° grado, 1162 para 3°, 1646 para 4°, 2038 para 5° y 2239 para 6°.

A.2.2. Efecto de la edad sobre el rendimiento en el test

La tabla A.3 muestra la mediana y el rango intercuartílico de los puntajes observados en los grupos según la edad.

Tabla A.3

Mediana y rango intercuartílico de los puntajes observados en los grupos según la edad en años

Edad	N	Mediana	RIC
7	361	15	11
8	673	18	16
9	644	27	16.5
10	592	33	15.5
11	569	38	17
12	258	37.5	17

RIC: rango intercuartílico

La prueba H de Kruskal-Wallis aplicada a los puntajes observados en los grupos según la edad reveló diferencias estadísticamente significativas [$\chi^2(5) = 1066, p < .001$] con un promedio de rangos de 744 para el grupo de 7 años, 974 para el de 8 años, 1487 para el de 9 años, 1882 para el de 10 años, 2187 para el de 11 años y 2258 para el de 12 años.

A.2.3. Efecto del grado escolar versus la edad sobre el rendimiento en el test

Una serie de pruebas U de Mann-Whitney aplicadas a los puntajes observados en los dos grupos según la edad para cada grado escolar reveló diferencias estadísticamente significativas para los grupos de 2° y 6° grados (2° grado: $U = 37699, Z = -2.78, p = .005$; 3° grado: $U = 45203, Z = -.17, p = .87$; 4° grado: $U = 58382, Z = -.028, p = .98$; 5° grado: $U = 25976, Z = -1.54, p = .12$; 6° grado: $U = 44904, Z = -2.47, p = .014$).

Las pruebas U de Mann-Whitney aplicadas a los puntajes observados en los dos grupos según el grado escolar para los grupos de niños de la misma edad reveló diferencias estadísticamente significativas para todas las comparaciones (8 años: $U = 29792, Z = -9.21, p < .001$; 9 años: $U = 28920,$

$Z = -7.57, p < .001$; 10 años: $U = 28694, Z = -7.13, p < .001$; 11 años: $U = 25500, Z = -5.01, p < .001$).

En la tabla A.4 se muestran la mediana, el rango intercuartílico y los promedios de rango de los puntajes observados en los grupos según el grado escolar y la edad.

Tabla A.4

Mediana y rango intercuartílico de los puntajes observados en los grupos según el grado escolar y la edad

Grado	Edad	N	Mediana	RIC	PR edad	PR grado
2°	7	361	15	11	318	
	8	241	13	11	277	245
3°	8	432	21	16.5	323	388
	9	211	21	16.5	320	243
4°	9	433	30	15	352	361
	10	270	29	15	353	242
5°	10	322	35	16	257	342
	11	176	34	16	236	233
6°	11	393	40	17	341	308
	12	258	37.5	17	303	

RIC: rango intercuartílico; PR edad: promedio de rangos de los grupos según la edad para cada grado escolar; PR grado: promedio de rangos de los grupos según el grado escolar para cada edad (8, 9, 10 y 11 años).

A.2.4. Efecto del sexo sobre el rendimiento en el test

La prueba U de Mann-Whitney aplicada a los puntajes observados en las niñas y los niños reveló que la diferencia no alcanzó significación estadística ($U = 1166678, Z = -1.30, p = .19$).

La mediana, el rango intercuartílico y los promedios de rangos de los puntajes observados fueron de 29, 20 y 1570 en las niñas, y de 28, 21 y 1528 en los niños, respectivamente.

A.2.5. Efecto del nivel de oportunidades educativas (NOE) sobre el rendimiento en el test

La mediana y el rango intercuartílico de los puntajes observados fueron de 20 y 19 en el grupo con NOE bajo, de 31 y 20 en el grupo con NOE medio, y de 34 y 22 en el grupo con NOE alto, respectivamente.

La prueba H de Kruskal-Wallis aplicada a los grupos con distinto nivel de oportunidades educativas reveló un efecto estadísticamente significativo de esta variable sobre el rendimiento [$\chi^2(2) = 304, p < .001$], con un promedio de rangos de 1101 para el grupo de NOE bajo, 1689 para el de NOE medio y 1858 para el de NOE alto.

A.2.6. Efecto del tipo de gestión sobre el rendimiento en el test

La mediana y el rango intercuartílico de los puntajes observados fueron de 30 y 19 en el grupo que asistía a escuelas de gestión estatal, y de 32 y 21 en el grupo que asistía a escuelas de gestión privada.

La prueba U de Mann-Whitney aplicada a los puntajes observados en el grupo de NOE medio segmentado según el tipo de gestión mostró una diferencia estadísticamente significativa entre los niños que asistían a escuelas de gestión estatal y los que asistían a escuelas de gestión privada ($U = 359166, Z = -4.34, p < .001$), con un promedio de rangos de 894 para los primeros y 1006 para los segundos.

A.2.7. Efecto de la región geográfica sobre el rendimiento en el test

La tabla A.5 muestra la mediana y el rango intercuartílico de los puntajes observados en el grupo que asistía a escuelas de gestión estatal segmentado según la región.

Tabla A.5

Mediana y rango intercuartílico de los puntajes observados en el grupo que asistía a escuelas de gestión estatal segmentado según la región

Región	N	Mediana	RIC
SUR	577	26	11
OES	787	26	20
NOR	438	23	20
CABA	221	29	16

RIC: rango intercuartílico; SUR: sur del Gran Buenos Aires; OES: oeste del Gran Buenos Aires; NOR: norte del Gran Buenos Aires; CABA: Ciudad Autónoma de Buenos Aires

La prueba H de Kruskal-Wallis mostró un efecto estadísticamente significativo de la región sobre el rendimiento [$\chi^2(3) = 22.53, p < .001$], con

un promedio de rangos de 999 para el sur del GBA, 1018 para el oeste del GBA, 939 para el norte del GBA y 1166 para la CABA.

Anexo B

Baremos del TECLE Buenos Aires para la Ciudad Autónoma de Buenos Aires y alrededores

Los baremos del test se confeccionaron a partir de la transformación de los puntajes directos observados en los distintos grupos en percentiles y rangos percentiles.

En este Anexo se presentan, en primer lugar, las tablas con las medidas de posición, el rango intercuartílico y los baremos en rangos percentiles correspondientes a los puntajes directos observados en los grupos según el grado escolar (tablas B.1, B.2, B.3, B.4, B.5 y B.6).

A continuación, se presentan los datos para cada uno de estos grupos segmentados según el nivel de oportunidades educativas (tablas B.7, B.8, B.9, B.10, B.11 y B.12).

La inclusión del percentil 16 (P16) como puntaje de alerta (tablas B.1 y B.7) está justificada por la noción de que en las variables cuyos valores siguen una distribución normal dicho puntaje se ubica a un desvío estándar por debajo de la media, de acuerdo con el siguiente razonamiento: (i) la media divide al conjunto de valores en dos mitades y por lo tanto coincide con la mediana (P50); (ii) los valores que se ubican entre un desvío estándar por debajo y un desvío estándar por encima de la media constituyen aproximadamente el 68% de las observaciones; por lo tanto, (iii) el puntaje que se ubica a un desvío estándar por debajo de la media supera a un porcentaje de observaciones igual al 50% menos el 34% inferior.

Por otro lado, un criterio de alerta más exigente sería considerar el puntaje correspondiente al percentil 25 (P25). Estudios futuros destinados a obtener evidencias de validez predictiva podrían contribuir a determinar el valor predictivo de ambos criterios en términos de su sensibilidad y especificidad para la identificación de dificultades en el aprendizaje de la lectura.

B.1. Medidas de posición, rango intercuartílico y baremos en rangos percentiles para los grupos según grado escolar

Tabla B.1

Medidas de posición (P16, P25, P50, P75) y rango intercuartílico de los puntajes directos observados en cada grupo según grado escolar

Grado (N)	P16 (puntaje de alerta)	P25	P50	P75	Rango intercuartílico
2° (602)	7	9	14	21	12
3° (643)	11	14	21	30	16
4° (703)	18	22	30	37	15
5° (498)	24	29	35	44	15
6° (651)	27	31	39	48	17

Tabla B.2*Rangos percentiles de los puntajes directos observados en 2° grado*

Puntaje	Frecuencia	Rango percentil	Puntaje	Frecuencia	Rango percentil
0	4	< 1	19	24	69
1	10	1	20	15	73
2	8	3	21	17	75
3	19	5	22	13	78
4	11	8	23	9	80
5	16	10	24	15	82
6	17	13	25	13	84
7	21	16	26	16	86
8	17	19	27	10	89
9	31	23	28	9	90
10	18	27	29	11	92
11	34	31	30	6	93
12	41	38	31	9	94
13	34	44	32	8	96
14	22	48	33 – 34	7	97
15	32	53	35 – 37	6	98
16	24	57	38 – 40	5	99
17	20	61	44 – 47	3	> 99
18	28	65	-	-	-

Tabla B.3*Rangos percentiles de los puntajes directos observados en 3° grado.*

Puntaje	Frecuencia	Rango percentil	Puntaje	Frecuencia	Rango percentil
0 – 1	5	< 1	25	17	60
2	4	1	26	23	63
3 – 4	11	2	27	8	66
5	6	4	28	15	67
6	8	5	29	14	70
7	12	6	30	29	73
8	16	8	31	16	76
9	17	11	32	17	79
10	14	13	33	22	82
11	17	16	34	9	84
12	21	19	35	12	86
13	27	22	36	12	88
14	21	26	37	11	90
15	16	29	38	9	91
16	24	32	39	3	92
17	29	36	40	6	93
18	20	40	41	9	94
19	30	44	42 – 43	6	95
20	17	47	44 – 45	6	96
21	15	50	46	4	97
22	13	52	47 – 49	9	98
23	20	55	50 – 51	6	99
24	16	57	52 – 60	3	> 99

Tabla B.4*Rangos percentiles de los puntajes directos observados en 4° grado*

Puntaje	Frecuencia	Rango percentil	Puntaje	Frecuencia	Rango percentil
0 – 2	3	< 1	31	30	55
6 – 8	8	1	32	26	59
9 – 10	10	2	33	32	63
11	6	3	34	19	66
12	8	4	35	25	70
13	11	6	36	16	72
14	11	7	37	18	75
15	15	9	38	15	77
16	10	11	39	21	80
17	20	13	40	15	82
18	15	15	41	17	85
19	16	18	42	10	86
20	16	20	43	11	88
21	25	23	44	11	90
22	14	26	45	6	91
23	31	29	46	11	92
24	32	33	47	17	94
25	13	36	48	11	96
26	19	39	49 – 50	8	97
27	25	42	51 – 52	4	98
28	14	45	53 – 55	9	99
29	29	48	57 – 62	2	> 99
30	20	51	-	-	-

Tabla B.5*Rangos percentiles de los puntajes directos observados en 5° grado*

Puntaje	Frecuencia	Rango percentil	Puntaje	Frecuencia	Rango percentil
6	1	< 1	36	21	55
7 – 10	7	1	37	9	58
11 – 12	6	2	38	15	60
13 – 14	4	3	39	9	63
15	6	4	40	9	65
16	3	5	41	8	66
17	6	6	42	12	68
18	11	8	43	15	71
19	2	9	44	13	74
20	8	10	45	11	76
21	3	11	46	12	79
22	11	12	47	12	81
23	8	14	48	9	83
24	9	16	49	8	85
25	2	17	50	15	87
26	10	18	51	8	89
27	13	21	52	10	91
28	13	23	53	6	93
29	13	26	54	4	94
30	10	28	55	10	95
31	18	31	56 – 57	9	97
32	27	36	58 – 59	3	98
33	28	41	60 – 62	4	99
34	28	47	63 – 64	3	> 99
35	17	51	--	-	-

Tabla B.6*Rangos percentiles de los puntajes directos observados en 6° grado*

Puntaje	Frecuencia	Rango percentil	Puntaje	Frecuencia	Rango percentil
5 – 7	4	< 1	40	23	52
9 – 12	6	1	41	17	55
13- 16	7	2	42	11	57
17 – 18	9	3	43	14	59
19	8	5	44	18	62
20	10	6	45	31	65
21	4	7	46	20	69
22	8	8	47	16	72
23	10	9	48	23	75
24	7	11	49	15	78
25	7	12	50	18	80
26	12	13	51	14	83
27	16	15	52	9	85
28	14	18	53	13	86
29	12	20	54	11	88
30	16	22	55	10	90
31	16	24	56	8	91
32	15	27	57	11	93
33	18	29	58	8	94
34	21	32	59	8	95
35	27	36	60	7	96
36	23	40	61	5	97
37	23	43	62	4	98
38	18	46	63	6	99
39	16	49	64	4	> 99

B.2. Medidas de posición, rango intercuartílico y baremos en rangos percentiles para los grupos según el grado escolar y el nivel de oportunidades educativas

Tabla B.7

Medidas de posición (P16, P25, P50, P75) y rango intercuartílico de los puntajes directos observados en cada grupo según grado escolar y nivel de oportunidades educativas.

Grado (N)	NOE (N)	P16 (puntaje de alerta)	P25	P50	P75	Rango Intercuartílico
2° (602)	Bajo (133)	3	3	8	13	10
	Medio (352)	9	11	15	21	10
	Alto (117)	12	13	19	25	12
3° (643)	Bajo (202)	7	8	13	18	10
	Medio (405)	15	18	25	32	14
	Alto (36)	24	29	33	30	11
4° (703)	Bajo (199)	13	15	22	29	14
	Medio (386)	21	24	31	38	14
	Alto (118)	23	28	36	43	15
5° (498)	Bajo (140)	16	18	30	34	16
	Medio (337)	28	32	38	47	15
	Alto (21)	33	34	42	45	11
6° (651)	Bajo (173)	20	24	32	38	14
	Medio (390)	30	34	41	49	15
	Alto (88)	37	40	48	55	15

NOE: nivel de oportunidades educativas.

Tabla B.8

Rangos percentiles de los puntajes directos observados en los grupos de 2° grado con NOE bajo, medio y alto

Puntaje	NOE bajo	NOE medio	NOE alto	Puntaje	NOE bajo	NOE medio	NOE alto
0	1	-	-	21	91	75	59
1	6	< 1	-	22	91	77	64
2	12	1	-	23	91	79	68
3	20	1	-	24	92	81	72
4	28	2	< 1	25	94	83	75
5	32	4	2	26	96	85	79
6	37	7	2	27	96	88	82
7	43	10	3	28	97	90	83
8	49	12	5	29	97	91	86
9	55	16	7	30	98	93	88
10	60	21	7	31	99	94	90
11	64	26	11	32	> 99	96	92
12	70	32	17	33	-	97	92
13	76	38	25	34	-	97	94
14	79	43	29	35	-	98	94
15	80	49	34	36	-	98	95
16	82	55	38	37	-	99	95
17	84	59	41	38	-	99	96
18	87	64	45	39	-	99	97
19	88	69	50	40 – 43	-	99	99
20	89	72	55	44	-	> 99	> 99

NOE: nivel de oportunidades educativas

Tabla B.9

Rangos percentiles de los puntajes directos observados en los grupos de 3° grado con NOE bajo, medio y alto

Puntaje	NOE bajo	NOE medio	NOE alto	Puntaje	NOE bajo	NOE medio	NOE alto
0	< 1	-	-	27	89	58	15
1	1	-	-	28	90	60	19
2	3	-	-	29	91	63	24
3	5	-	-	30	92	67	29
4	7	< 1	-	31	93	72	35
5	10	1	-	32	94	75	40
6	13	1	-	33	94	79	49
7	17	1	-	34	95	82	54
8	22	2	-	35	96	84	57
9	28	3	-	36	96	86	61
10	32	5	-	37	97	88	67
11	36	7	-	38	98	90	71
12	41	9	-	39	98	91	74
13	47	12	-	40	98	91	76
14	53	15	-	41	98	93	76
15	59	17	-	42	98	95	76
16	63	19	-	43	99	95	76
17	68	23	3	44	99	96	76
18	74	26	3	45	99	96	79
19	77	30	3	46	99	97	83
20	80	35	7	47	99	98	89
21	82	37	10	48	99	98	93
22	85	39	10	49	99	98	96
23	86	43	12	50	99	99	99
24	86	47	15	51	99	> 99	-
25	87	50	15	52	> 99	-	-
26	88	55	15	-	-	-	-

NOE: nivel de oportunidades educativas

Tabla B.10

Rangos percentiles de los puntajes directos observados en los grupos de 4° grado con NOE bajo, medio y alto

Puntaje	NOE bajo	NOE medio	NOE alto	Puntaje	NOE bajo	NOE medio	NOE alto
0	< 1	-	-	29	76	40	25
1	1	-	-	30	79	44	27
2	1	-	-	31	83	48	29
3	1	-	-	32	84	53	33
4	1	-	-	33	86	58	38
5	1	-	-	34	88	63	42
6	2	-	-	35	90	66	45
7	3	-	-	36	91	70	50
8	4	-	< 1	37	92	72	54
9	5	-	1	38	93	75	58
10	7	-	2	39	93	78	63
11	10	-	2	40	94	81	65
12	13	< 1	3	41	96	84	68
13	16	1	4	42	98	85	72
14	20	2	4	43	98	86	76
15	25	2	5	44	99	88	80
16	30	3	5	45	99	89	82
17	33	5	6	46	99	90	86
18	36	7	6	47	99	92	91
19	40	9	7	48	> 99	95	93
20	43	11	8	49	-	96	94
21	47	14	10	50	-	97	95
22	51	17	11	51	-	97	96
23	54	20	14	52	-	98	97
24	59	24	18	53 – 54	-	98	98
25	63	27	21	55 – 56	-	99	99
26	66	30	22	57	-	99	> 99
27	71	32	24	58 – 61	-	99	-
28	74	36	25	62	-	> 99	-

NOE: nivel de oportunidades educativas

Tabla B.11

Rangos percentiles de los puntajes directos observados en los grupos de 5° grado con NOE bajo, medio y alto

Puntaje	NOE bajo	NOE medio	NOE alto	Puntaje	NOE bajo	NOE medio	NOE alto
6	< 1	-	-	36	80	46	36
7	2	-	-	37	83	49	36
8	5	-	-	38	85	51	36
9	5	-	-	39	86	55	36
10	5	-	-	40	86	57	43
11	6	-	-	41	88	59	43
12	8	-	-	42	88	61	50
13	10	-	-	43	89	64	57
14	11	< 1	-	44	91	67	69
15	13	1	-	45	92	70	69
16	16	1	-	46	94	72	79
17	18	1	-	47	94	75	83
18	23	2	-	48	94	78	83
19	23	3	-	49	95	80	88
20	26	4	-	50	95	83	88
21	27	5	-	51	95	87	88
22	30	6	2	52	96	89	95
23	34	7	2	53	97	91	-
24	37	8	2	54	97	92	-
25	39	9	2	55	97	94	-
26	41	10	2	56	99	95	-
27	44	12	2	57	> 99	96	-
28	46	15	2	58	-	97	-
29	49	18	2	59	-	98	-
30	51	20	2	60	-	98	-
31	56	22	2	61	-	99	-
32	61	27	7	62	-	99	-
33	67	32	14	63	-	99	-
34	73	37	21	64	-	> 99	-
35	77	42	29	-	-	-	-

NOE: nivel de oportunidades educativas

Tabla B.12

Rangos percentiles de los puntajes directos observados en los grupos de 6° grado con NOE bajo, medio y alto

Puntaje	NOE bajo	NOE medio	NOE alto	Puntaje	NOE bajo	NOE medio	NOE alto
6	< 1	-	-	36	68	33	13
7	1	-	-	37	71	37	15
8	1	-	-	38	74	40	18
9	1	-	-	39	77	43	21
10	2	< 1	-	40	79	46	24
11	3	1	-	41	81	50	28
12	3	1	-	42	83	52	30
13	4	1	-	43	85	54	31
14	5	1	-	44	86	57	33
15	5	1	-	45	89	61	37
16	6	1	-	46	91	66	41
17	7	1	-	47	92	69	44
18	9	1	-	48	94	73	48
19	12	2	-	49	95	76	51
20	16	3	-	50	96	79	54
21	18	4	-	51	98	82	59
22	20	4	1	52	98	84	62
23	22	5	1	53	98	85	67
24	24	7	1	54	98	87	73
25	26	8	1	55	98	89	77
26	30	8	2	56	98	90	80
27	34	10	3	57	99	92	83
28	38	12	3	58	99	94	86
29	40	14	3	59	99	95	89
30	42	17	4	60	99	96	92
31	45	19	5	61	99	97	94
32	50	21	6	62	99	98	96
33	54	23	8	63	> 99	99	98
34	57	26	10	64	-	> 99	99
35	63	29	10	-	-	-	-

NOE: nivel de oportunidades educativas

Anexo C

El TECLE Buenos Aires. Protocolo de respuestas y normas de aplicación y calificación

En el presente anexo se presenta en primer lugar el protocolo de respuestas del TECLE Buenos Aires. A continuación, se detallan las normas de aplicación y calificación. Finalmente, se ofrece una planilla destinada a reunir los datos del desempeño en el test de los alumnos de un curso con el objeto de facilitar la interpretación de los puntajes observados.

C.1. TECLE Buenos Aires: protocolo de respuestas

DATOS DEL ALUMNO		
Nombre y apellido:.....		
Edad:.....		
Escuela:.....		
Grado:.....	Sección:.....	Turno:.....

Nombre y apellido del docente:.....

Firma del docente.....

Fecha de evaluación:.....

DESEMPEÑO EN EL TEST

Puntaje directo:

Rango percentil
(baremo según grado escolar):

Rango percentil
(baremo según grado escolar y nivel de oportunidades educativas):

Ubicación en el ranking del grado

Último ítem respondido Nº errores Nº omisiones

Observaciones:.....
.....
.....

ÍTEMS DE PRÁCTICA

Marcá con una la palabra que completa la oración.

1. Tu pelota es de color...
<input type="checkbox"/> vedre <input type="checkbox"/> veque <input type="checkbox"/> verte <input type="checkbox"/> verde
2. El caballo tenía la pata...
<input type="checkbox"/> ropa <input type="checkbox"/> rota <input type="checkbox"/> rofa <input type="checkbox"/> rona

ÍTEMS DE PRUEBA

Marcá con una la palabra que completa la oración.

1. En el tren viajan muchas... <input type="checkbox"/> personas <input type="checkbox"/> perchonas <input type="checkbox"/> permonas <input type="checkbox"/> perdonas
2. Le gustaba hablar con sus... <input type="checkbox"/> amicos <input type="checkbox"/> amigos <input type="checkbox"/> amibos <input type="checkbox"/> arribos
3. Me visita cada dos... <input type="checkbox"/> deas <input type="checkbox"/> dúas <input type="checkbox"/> digas <input type="checkbox"/> días
4. Hemos viajado por todo el... <input type="checkbox"/> mumbo <input type="checkbox"/> mundo <input type="checkbox"/> mudo <input type="checkbox"/> munto
5. El perro lo seguía con su... <input type="checkbox"/> migada <input type="checkbox"/> ministro <input type="checkbox"/> mirada <input type="checkbox"/> micada
6. Laura quiere un... <input type="checkbox"/> tripartito <input type="checkbox"/> trificlo <input type="checkbox"/> triciclo <input type="checkbox"/> triviclo
7. Me dolía la mano y no podía... <input type="checkbox"/> escribir <input type="checkbox"/> estribir <input type="checkbox"/> escriilir <input type="checkbox"/> extinguir
8. El coche está en... <input type="checkbox"/> movibiento <input type="checkbox"/> movimienlo <input type="checkbox"/> molinete <input type="checkbox"/> movimiento
9. Juan sale de su... <input type="checkbox"/> halitación <input type="checkbox"/> habitación <input type="checkbox"/> habitafión <input type="checkbox"/> habitaron
10. Aceleró hasta el límite de... <input type="checkbox"/> vellosidad <input type="checkbox"/> verocidad <input type="checkbox"/> velocidad <input type="checkbox"/> veloridad
11. Quiso subir al... <input type="checkbox"/> camello <input type="checkbox"/> camillo <input type="checkbox"/> canello <input type="checkbox"/> canilla
12. Mi amigo viene en... <input type="checkbox"/> bidi <input type="checkbox"/> bici <input type="checkbox"/> bicho <input type="checkbox"/> dici

13. Anoche comí un...	<input type="checkbox"/> jarrón	<input type="checkbox"/> tumón	<input type="checkbox"/> turón	<input type="checkbox"/> turrón
14. María le ocultaba la...	<input type="checkbox"/> verlad	<input type="checkbox"/> vertad	<input type="checkbox"/> vendar	<input type="checkbox"/> verdad
15. El médico le aconsejó tener cuidado con su...	<input type="checkbox"/> coraje	<input type="checkbox"/> corazón	<input type="checkbox"/> conazón	<input type="checkbox"/> covazón
16. Se necesita trabajar mucho para tener más...	<input type="checkbox"/> expediencia	<input type="checkbox"/> expresado	<input type="checkbox"/> experiencia	<input type="checkbox"/> expertiencia
17. Tu auto es tan viejo que se le caen las...	<input type="checkbox"/> puertas	<input type="checkbox"/> pueltas	<input type="checkbox"/> puestas	<input type="checkbox"/> tuerpas
18. Para hacer el viaje tuvo que pedir...	<input type="checkbox"/> injormación	<input type="checkbox"/> inzormación	<input type="checkbox"/> información	<input type="checkbox"/> imposible
19. El domingo no saldremos a pescar con mi...	<input type="checkbox"/> farrilia	<input type="checkbox"/> familia	<input type="checkbox"/> famidia	<input type="checkbox"/> famosa
20. En el primer cajón, tu abuelo guarda su juego de...	<input type="checkbox"/> dados	<input type="checkbox"/> tados	<input type="checkbox"/> dodos	<input type="checkbox"/> dedos
21. Llegó a su casa y se puso a tocar la...	<input type="checkbox"/> guibarra	<input type="checkbox"/> guitarra	<input type="checkbox"/> guifarra	<input type="checkbox"/> guirnalda
22. La hija de Analía no dejó de llorar en toda la...	<input type="checkbox"/> noce	<input type="checkbox"/> nolle	<input type="checkbox"/> nota	<input type="checkbox"/> noche
23. El chico necesita ayuda para resolver el...	<input type="checkbox"/> protesta	<input type="checkbox"/> probrema	<input type="checkbox"/> problema	<input type="checkbox"/> proglema
24. El maestro se sienta en su...	<input type="checkbox"/> escritorio	<input type="checkbox"/> escritolio	<input type="checkbox"/> escrifolio	<input type="checkbox"/> escribiente
25. Su padre trabaja en el...	<input type="checkbox"/> puebro	<input type="checkbox"/> pueglo	<input type="checkbox"/> pueblo	<input type="checkbox"/> puedo

26. Ana puso los platos sobre la...	<input type="checkbox"/> meva	<input type="checkbox"/> mesa	<input type="checkbox"/> mefa	<input type="checkbox"/> meca
27. Juana nos cuenta cuentos con mucha...	<input type="checkbox"/> imaginafi3n	<input type="checkbox"/> imapinaci3n	<input type="checkbox"/> imantado	<input type="checkbox"/> imaginaci3n
28. El policia encarcel3 a los...	<input type="checkbox"/> delinpuentes	<input type="checkbox"/> debicuentes	<input type="checkbox"/> delincuentes	<input type="checkbox"/> declinantes
29. Entre las flores hay una...	<input type="checkbox"/> margarita	<input type="checkbox"/> magrarina	<input type="checkbox"/> marcarita	<input type="checkbox"/> marcador
30. No aprendi3 ingl3s en toda su...	<input type="checkbox"/> viba	<input type="checkbox"/> vino	<input type="checkbox"/> vica	<input type="checkbox"/> vida
31. Su hermano estudia en la...	<input type="checkbox"/> universidad	<input type="checkbox"/> unidersidad	<input type="checkbox"/> unipersonal	<input type="checkbox"/> universilad
32. En sus 3ltimas vacaciones recorrieron bastantes...	<input type="checkbox"/> kilodramos	<input type="checkbox"/> quir3fano	<input type="checkbox"/> quir3metro	<input type="checkbox"/> kil3metros
33. Era una soluci3n bastante...	<input type="checkbox"/> raformable	<input type="checkbox"/> rastrillada	<input type="checkbox"/> razonable	<input type="checkbox"/> razomable
34. Las cajas eran tan pesadas que no pudimos...	<input type="checkbox"/> traslatarlas	<input type="checkbox"/> trasladarlas	<input type="checkbox"/> trasfadarlas	<input type="checkbox"/> trasparentes
35. La secretaria de Carlos tiene la cara llena de...	<input type="checkbox"/> lunares	<input type="checkbox"/> lunarres	<input type="checkbox"/> lunetas	<input type="checkbox"/> luranes
36. Las fuertes lluvias provocaron una...	<input type="checkbox"/> inmundaci3n	<input type="checkbox"/> untaci3n	<input type="checkbox"/> inmutaci3n	<input type="checkbox"/> inundaci3n
37. Se sent3 frente a la ventana para hablar por...	<input type="checkbox"/> telescopio	<input type="checkbox"/> tel3zono	<input type="checkbox"/> tel3fono	<input type="checkbox"/> tel3jono
38. El nuevo ministro fue entrevistado por el...	<input type="checkbox"/> perriodiza	<input type="checkbox"/> periodista	<input type="checkbox"/> per3metro	<input type="checkbox"/> petriodista

39. Todas las mañanas caminábamos por la playa buscando...	<input type="checkbox"/> calacores	<input type="checkbox"/> taracoles	<input type="checkbox"/> cartulinas	<input type="checkbox"/> caracoles
40. La cuñada de Sara no pudo llegar a...	<input type="checkbox"/> tiembo	<input type="checkbox"/> tiento	<input type="checkbox"/> tiespo	<input type="checkbox"/> tiempo
41. Para coserle la camisa buscó el hilo y la...	<input type="checkbox"/> aguja	<input type="checkbox"/> abuja	<input type="checkbox"/> abeja	<input type="checkbox"/> baguja
42. Si no tiene cuidado va a sufrir las...	<input type="checkbox"/> colsecuencias	<input type="checkbox"/> consejos	<input type="checkbox"/> consecuencias	<input type="checkbox"/> consetuencias
43. No consiguió el libro y tiene que hacer...	<input type="checkbox"/> fotógrafos	<input type="checkbox"/> fotonopias	<input type="checkbox"/> fotocobias	<input type="checkbox"/> fotocopias
44. Tus sobrinas se van a resfriar por andar...	<input type="checkbox"/> descalzas	<input type="checkbox"/> despaldas	<input type="checkbox"/> dercalzas	<input type="checkbox"/> descartas
45. En el hospital le sacaron una...	<input type="checkbox"/> rabiografía	<input type="checkbox"/> radiografía	<input type="checkbox"/> raquiografía	<input type="checkbox"/> radiodifusor
46. La familia de Cecilia va a la sierra porque le gusta la...	<input type="checkbox"/> naturaleza	<input type="checkbox"/> natunaleza	<input type="checkbox"/> nacionalidad	<input type="checkbox"/> naturalefa
47. Luis va al zoológico y quiere dar de comer a todos los...	<input type="checkbox"/> animarles	<input type="checkbox"/> animates	<input type="checkbox"/> animales	<input type="checkbox"/> alimanes
48. En la estación del tren, Francisca notó que estaba...	<input type="checkbox"/> desadrigaba	<input type="checkbox"/> desabrigada	<input type="checkbox"/> desabritada	<input type="checkbox"/> desabrochar
49. En medio de la tormenta, el capitán daba órdenes a los...	<input type="checkbox"/> tripulantes	<input type="checkbox"/> trifulantes	<input type="checkbox"/> tridentes	<input type="checkbox"/> fritulantes
50. Para ubicar los puntos cardinales, se necesita una...	<input type="checkbox"/> burbuja	<input type="checkbox"/> brújula	<input type="checkbox"/> blujula	<input type="checkbox"/> brúgula
51. El deportista no se curó bien la herida y le quedó una...	<input type="checkbox"/> quisatriz	<input type="checkbox"/> emperatriz	<input type="checkbox"/> citatriz	<input type="checkbox"/> cicatriz

52. Para abrir más rápido su regalo, Lucía rompió el...	<input type="checkbox"/> carta	<input type="checkbox"/> carjón	<input type="checkbox"/> cartón	<input type="checkbox"/> carpón
53. Tu padre trabajaba como carpintero antes de que su jefe lo dejara...	<input type="checkbox"/> desempleabo	<input type="checkbox"/> desempleado	<input type="checkbox"/> desplegado	<input type="checkbox"/> desempleado
54. El que se despierte temprano mañana podrá venir con...	<input type="checkbox"/> nosopros	<input type="checkbox"/> norteños	<input type="checkbox"/> nonotros	<input type="checkbox"/> nosotros
55. Observando al planeta, vio unos satélites girando a su...	<input type="checkbox"/> alrebedor	<input type="checkbox"/> alternativa	<input type="checkbox"/> alrededor	<input type="checkbox"/> alsededor
56. Jorge no logra dormirse porque desde su habitación oye al...	<input type="checkbox"/> trompetista	<input type="checkbox"/> tropezaba	<input type="checkbox"/> trompefista	<input type="checkbox"/> trompepista
57. En las redes del barco pesquero, quedó atrapado un...	<input type="checkbox"/> deflín	<input type="checkbox"/> desliz	<input type="checkbox"/> deltín	<input type="checkbox"/> delfín
58. No tuvo problemas para encontrar sus zapatillas porque todas eran...	<input type="checkbox"/> difíciles	<input type="checkbox"/> diferenles	<input type="checkbox"/> dicerentes	<input type="checkbox"/> diferentes
59. En las tareas que no le gustan, la nueva empleada no se esfuerza...	<input type="checkbox"/> decorado	<input type="checkbox"/> demasiado	<input type="checkbox"/> debasiado	<input type="checkbox"/> demariado
60. Antes de irse le prometió que siempre trataría de...	<input type="checkbox"/> portegerla	<input type="checkbox"/> proteperla	<input type="checkbox"/> proteínas	<input type="checkbox"/> protegerla
61. Decile a Josefina que no moje su reloj porque no es...	<input type="checkbox"/> sumergible	<input type="checkbox"/> sumerdible	<input type="checkbox"/> somergible	<input type="checkbox"/> sustituirle
62. El ascensor estaba descompuesto y la gente tuvo que bajar por las...	<input type="checkbox"/> escalenos	<input type="checkbox"/> escorelas	<input type="checkbox"/> escaleras	<input type="checkbox"/> escalinas
63. Interrumpieron la transmisión de la final para dar una...	<input type="checkbox"/> princesa	<input type="checkbox"/> prinicia	<input type="checkbox"/> primicia	<input type="checkbox"/> primitia
64. El periodista dijo que el resultado del partido era...	<input type="checkbox"/> irrompible	<input type="checkbox"/> previsible	<input type="checkbox"/> predidible	<input type="checkbox"/> presivable

C.2. TECLE Buenos Aires: normas de aplicación y calificación

Las normas de aplicación y calificación del test fueron elaboradas con el objeto de asegurar la uniformidad de los procedimientos necesaria para minimizar los errores de medición (Anastasi & Urbina, 1998). Por lo tanto, es fundamental que sean respetadas con la mayor rigurosidad posible.

C.2.1. Normas de aplicación del test

Las normas de aplicación consisten en las instrucciones para la administración colectiva e individual del test. Se proponen reducir al mínimo posible las diferencias en las condiciones de evaluación con el objeto de que las respuestas a los ítems reflejen con la mayor exactitud posible el nivel de eficacia lectora de cada niño.

Instrucciones para la administración colectiva del test en el aula

En primer lugar, es fundamental *respetar las instrucciones para la aplicación del test* antes de dar comienzo a la evaluación, así como contar con el *material necesario*, a saber, cantidad suficiente de protocolos de respuesta, biromes, y cronómetro.

Es igualmente importante asegurarse de que la evaluación se desarrollará en un *ambiente cómodo y libre de distracciones* (ruido e interrupciones). Particularmente, se debe solicitar el apagado de celulares e impedir el ingreso al aula de alumnos rezagados y de toda persona ajena a la evaluación.

Se recomienda contar con asistencia para la distribución y recolección de los protocolos de respuesta a fin de evitar que los alumnos continúen realizando la tarea propuesta una vez finalizado el tiempo límite permitido.

Una vez dadas las condiciones para la administración del test, se distribuyen los protocolos de respuesta y se informa a los alumnos acerca de la actividad con la siguiente consigna:

“Ahora vamos a hacer un ejercicio de lectura que sirve para conocer cómo leen y comprenden lo que leen. El resultado del ejercicio no lo vamos a usar como nota de Lengua en el boletín, pero es necesario que hagan el ejercicio seriamente, prestando atención.

Les voy a repartir un conjunto de hojas. En la primera, van a ver el recuadro en el que está escrito “Datos del alumno”. Completen solamente los datos de ese recuadro y cuando terminen, esperen mis instrucciones. Es importante que no den vuelta las hojas hasta que yo se los indique.”

Una vez que todos los alumnos hayan completado sus datos personales, se continúa con la explicación de la tarea:

“En las hojas hay una serie de oraciones incompletas a las que les falta la última palabra. Debajo de cada oración hay cuatro escrituras y cada una está acompañada de un casillero vacío. El ejercicio consiste en leer en silencio la oración incompleta y elegir solo una de las cuatro escrituras que complete a esa oración, dibujando una cruz o una equis en el casillero correspondiente. Es decir que **no** tienen que escribir la escritura elegida al final de cada oración, sino señalarla con una cruz o una equis. Vamos a ver unos ejemplos que están al final de la primera hoja.”

A continuación, se escribe en el pizarrón el primer ejemplo, tal como se presenta en la primer hoja del protocolo (la oración incompleta seguida de las cuatro opciones, cada una con el casillero vacío correspondiente), se lee en voz alta la oración incompleta y las opciones, se muestra cómo quedaría la oración si se la completara con cada una de las opciones, se pregunta a los alumnos cuál es la opción correcta, y se la señala con una equis en el casillero correspondiente. Es importante advertir a los alumnos acerca de la similitud entre las opciones. Luego se presenta el segundo ejemplo, pidiendo a los alumnos que marquen la opción correcta en su hoja de respuesta y se continúa con la consigna de la siguiente manera:

“La idea es que cada uno complete **solo y sin ayuda**, la mayor cantidad posible de oraciones en 5 minutos, empezando con la oración número 1. Cuando terminen de completar las oraciones de esa hoja continúen con la siguiente y así hasta que yo les avise que terminó el tiempo. Yo voy a decir “¡ya!” cuando puedan empezar la tarea y voy a decir “¡basta!” cuando se hayan cumplido los 5 minutos. En ese momento, todo el mundo tiene que dejar la birome, dar vuelta el montón de hojas contra la mesa y levantar ambas manos.

Si una oración les resulta difícil porque no conocen alguna de las palabras o no están seguros de elegir la respuesta correcta, no se detengan mucho tiempo, salteen esa oración y continúen con la siguiente.

Si les sucede que en una oración se arrepienten de haber elegido una respuesta, táchenla y marquen la nueva elección con una equis en el casillero correspondiente.”

En este momento, se debe verificar si todos los alumnos comprendieron la consigna, aclarar las dudas eventuales y dar comienzo a la evaluación con la siguiente instrucción:

“Entonces, cuando yo diga “¡ya!”, vayan a la hoja que empieza con la oración número 1, tomen la birome y empiecen a completar las oraciones. Preparados... listos... ¡ya!”

Es importante estar atento y preparado ante cualquier eventualidad durante los 5 minutos de evaluación, que puede dar como resultado la invalidación de la administración de la prueba por no cumplir con las condiciones requeridas. Algunos ejemplos de lo que puede ocurrir son los siguientes: que alguna birome se quede sin tinta, que algún alumno consulte la respuesta correcta para alguna oración, o inclusive, que algún alumno dé por finalizada la tarea al completar las oraciones de la primer hoja sin advertir que hay más. En estos casos, es necesario resolver la situación de la manera más rápida posible, sin provocar la distracción del resto de los alumnos, así como consignarla en el espacio destinado a las observaciones.

Asimismo, es necesario estar atento al cronómetro, ya que 5 minutos pasan rápido. Al cumplirse el tiempo permitido se debe dar la siguiente instrucción:

“¡Basta! Den vuelta las hojas, dejen sus biromes y levanten las manos.”

En este momento, un asistente debe comenzar la recolección de los protocolos de respuesta, verificando que estén completos los datos personales del alumno, mientras el evaluador debe estar atento para que ningún alumno continúe completando oraciones.

Instrucciones para la administración individual del test

Para la evaluación individual, al igual que para la colectiva, es fundamental respetar las instrucciones para la aplicación del test, contar con el material necesario (protocolo de respuesta, birome y cronómetro) y asegurarse de que el procedimiento se desarrollará en un ambiente cómodo y libre de distracciones, antes de comenzar con la administración de la prueba.

Las instrucciones son esencialmente las mismas que para la evaluación colectiva en lo referido a la explicación brindada al sujeto evaluado acerca de la tarea, incluidos los ejemplos, y el tiempo límite permitido.

C.2.2. Normas de calificación

Obtención del puntaje directo

El puntaje directo obtenido en el test resulta de la sumatoria de los ítems respondidos correctamente.

Se considera que un ítem fue respondido correctamente si el alumno señaló la opción que completa coherentemente la oración, en los siguientes casos:

- Utilizando la marca solicitada en la consigna.

- Utilizando una marca diferente a la solicitada en la consigna (por ejemplo, subrayando o encerrando con un círculo dicha opción, en vez de marcar el casillero correspondiente).
- Luego de una o más autocorrecciones, siempre que las opciones previamente señaladas y descartadas hayan sido tachadas.

En la primera hoja del protocolo de respuestas, existe un espacio destinado a consignar el puntaje directo obtenido.

Eventualmente, puede ocurrir que un protocolo de respuestas deba ser invalidado ante la presencia de indicadores de que las respuestas no reflejan el nivel de “eficacia lectora” del sujeto evaluado, sino que cabe la sospecha de falta de comprensión de la consigna o de poco cuidado en la ejecución de la tarea. Por ejemplo, si el sujeto se saltó hojas del protocolo de respuestas, completó las oraciones de atrás hacia delante, exhibió un patrón estereotipado de respuesta marcando las opciones que ocupan una misma posición. En estos casos, es necesario dejar asentada tal eventualidad en la primera hoja del protocolo de respuestas, en el espacio destinado a las observaciones, y considerar otra instancia de evaluación, separada por un intervalo de tiempo razonable, digamos, de dos a cuatro semanas.

Puede resultar útil consignar además el número del último ítem respondido, la cantidad de errores y de omisiones.

Transformación del puntaje directo a rango percentil

Una vez obtenido el puntaje directo del niño, es necesaria su interpretación referida al desempeño de su grupo de referencia. Para ello, el puntaje directo debe ser transformado al rango percentil correspondiente exhibido en dos tablas: la que presenta el baremo para el grado escolar del niño y la que presenta el baremo para el grado escolar y el nivel de oportunidades educativas.

El rango percentil según el grado escolar permite ubicar el desempeño del niño en relación con el de la muestra general de niños del mismo grado

escolar, mientras que el rango percentil según el grado escolar y el nivel de oportunidades educativas permite ubicarlo en relación con el de un grupo más homogéneo de niños del mismo grado escolar que comparte además el nivel de oportunidades educativas. Adicionalmente, es útil confeccionar un ranking con el desempeño de los alumnos del curso y ubicar en ese ranking el desempeño del niño.

Así, por ejemplo, si el niño obtiene rangos percentiles bajos en ambos tipos de baremos, y además su desempeño es bajo en comparación con el de sus compañeros de curso, esto puede interpretarse como la indicación de un riesgo de desarrollar o presentar dificultades en el aprendizaje de la lectura. En cambio, si el niño obtiene un rango percentil en torno a 50 o superior en el baremo para el grado escolar y el nivel de oportunidades educativas correspondiente y su desempeño se ubica en torno a la mitad en el ranking del curso, pero obtiene un rango percentil bajo (inferior a 25) en el baremo general para el grado escolar, esto podría reflejar una estrategia para la enseñanza de la lectura inadecuada o insuficiente, así como pobres hábitos de lectura en el entorno de crianza del niño.

En el primer caso, se recomienda incorporar al niño en un programa de enseñanza de la lectura con estrategias probadamente eficaces para el entrenamiento de las habilidades de decodificación, por un lapso razonable de tiempo, seguido de una nueva evaluación destinada a medir el progreso en sus habilidades de lectura. En el segundo caso, se recomienda revisar la estrategia utilizada para la enseñanza de la lectura.

A continuación, se muestra una planilla que reúne los datos del rendimiento en el test de los alumnos del curso.

C.3. Planilla de desempeño en el TECLE Buenos Aires de los alumnos del curso

Escuela:.....

Grado / sección / turno:.....

Nivel de oportunidades educativas de la escuela
(consignar si es bajo, medio o alto):.....

Fecha de evaluación:.....

Medidas de posición del grupo de referencia

Baremos	P16	P25	P50	P75
Grado escolar				
Grado escolar según nivel oportunidades educativas				

Nº	Alumno (ordenados de mayor a menor en función del puntaje directo en el test)	Puntaje directo	Rango percentil GE	Rango percentil GE-NOE (*)
1				
2				
3				
4				
5				
6				
7				
8				
9				
10				
11				
12				
13				
14				
15				
16				
17				
18				

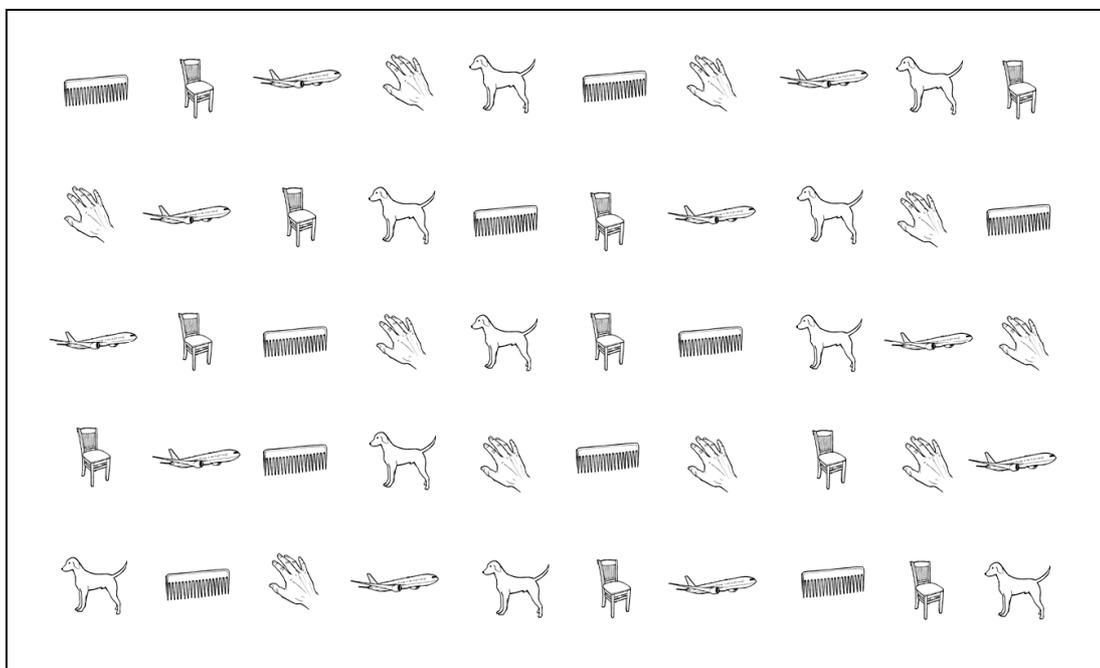
(*) Se sugiere interpretar los resultados del desempeño de cada niño según el siguiente detalle, considerando además su ubicación en el ranking del curso: si $RP \leq 24$ necesita mejorar; si $RP25-49$ desempeño bueno; si $RP50-74$ desempeño muy bueno; $RP \geq 75$ desempeño superior.

Anexo D

Instrumentos administrados en el estudio de validez de constructo

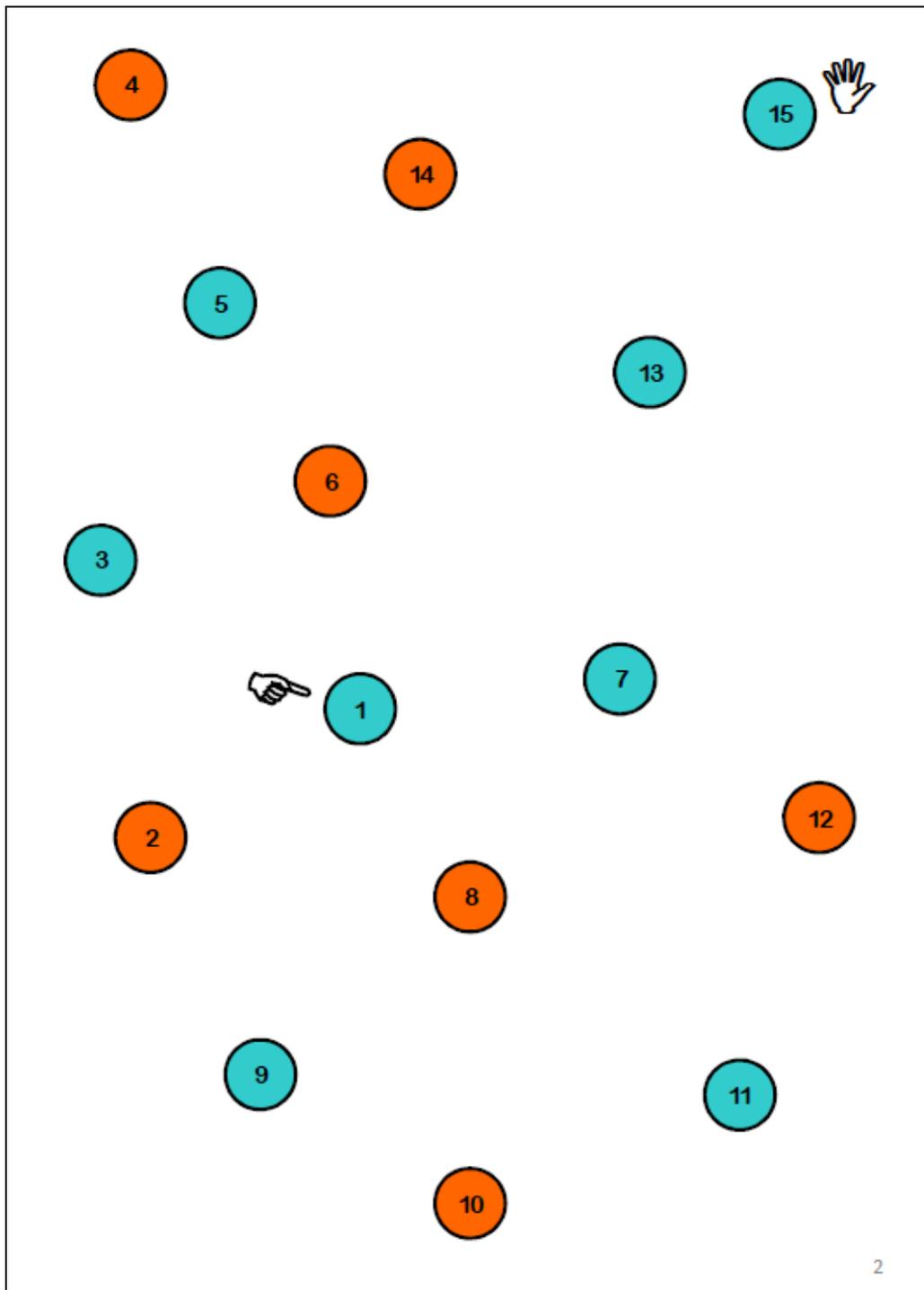
En el presente anexo, se presentan dos instrumentos no publicados utilizados en el estudio de validez de constructo: la lámina de dibujos para la denominación automatizada rápida del test DAR (Morello García, China, & Ferreres, 2015) y las láminas de prueba de las partes 1 y 2 de la prueba de trazado de senderos, diseñada para la presente tesis. Además, se presenta la prueba de *screening* sobre el texto expositivo “Un almuerzo prehistórico” del Test Leer para Comprender (Abusamra, Ferreres, Raiter, De Beni, & Cornoldi, 2010), disponible en http://www.paidosdep.com.ar/html/adjuntos/Un_almuerzo_Prehistorico.pdf.

D.1. Lámina de dibujos para la denominación automatizada rápida del test DAR (Morello García, China, & Ferreres, 2015)

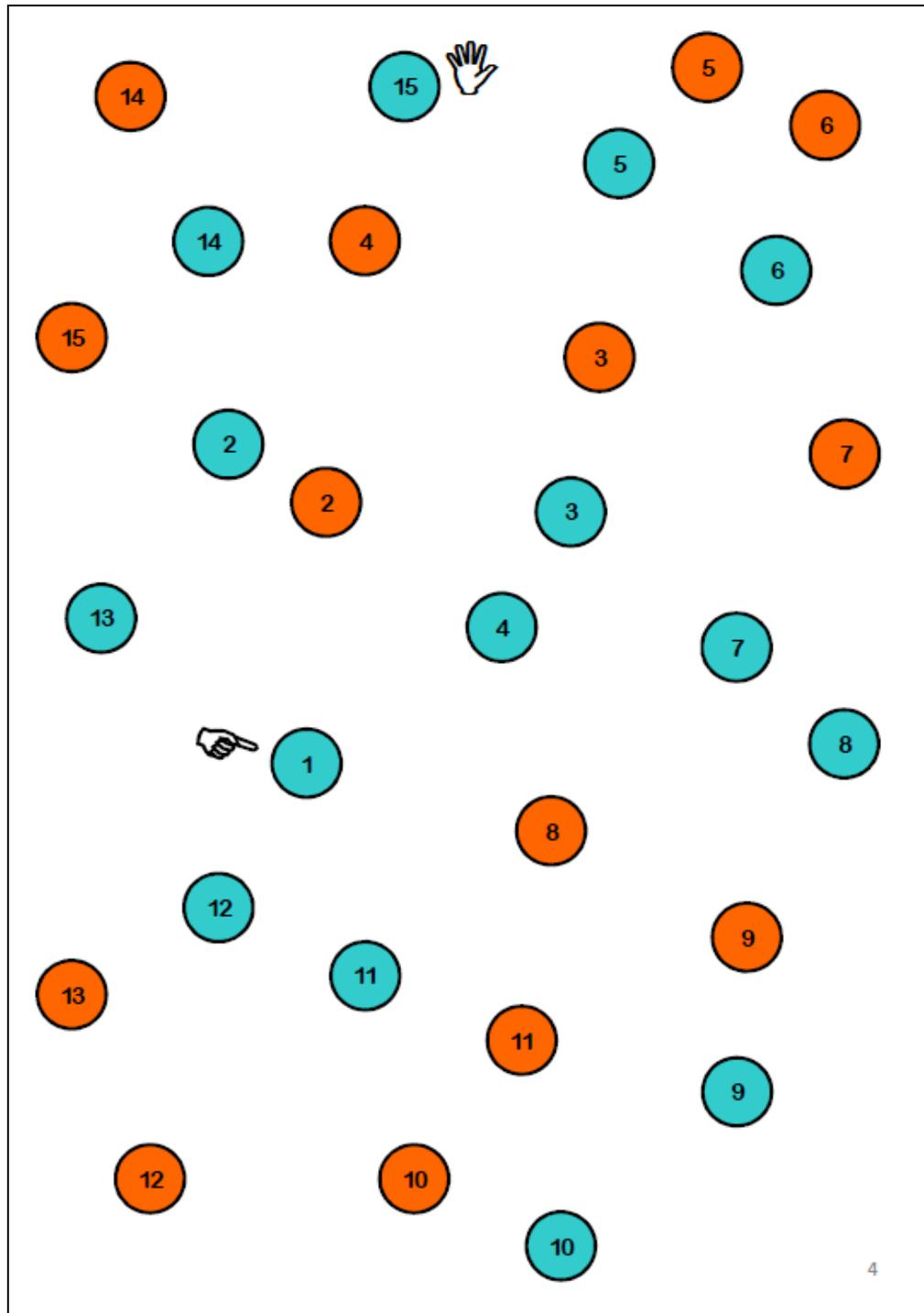


D2. Láminas de prueba de las partes 1 y 2 de la prueba de trazado de senderos diseñada para esta tesis (China, no publicado)

D.2.1. Lámina de prueba de la parte 1



D.2.2. Lámina de prueba de la parte 2



D.3. Prueba de screening sobre el texto expositivo “Un almuerzo prehistórico” del Test Leer para Comprender (Abusamra, Ferreres, Raiter, De Beni, & Cornoldi, 2010)

Nombre:

Edad:

Leer atentamente el siguiente texto.

Un almuerzo prehistórico

¡Qué dura debía ser la vida hace miles de años! En esos tiempos, los seres humanos vivían especialmente de la caza y de la recolección de frutos y raíces. Las tribus se desplazaban continuamente siguiendo a las presas que les servían de alimento. Los hombres eran los encargados de cazar y las mujeres de juntar frutas y raíces y, a veces, se ocupaban de la pesca. El alimento dependía de lo que podían encontrar en las diferentes regiones, pero el hambre era un fantasma permanente.

¿Cuál podía ser un almuerzo común para una familia de la prehistoria? Si tenían suerte, podían contar con osos, monos, elefantes, ciervos, jabalíes, renos o mamuts. Cuando la caza mayor escaseaba, atrapaban animales más pequeños, como ratas y murciélagos, y completaban la dieta con escarabajos, larvas, gusanos y otros insectos.

*Para poder hacer frente a animales mucho más grandes que **ellos**, los cazadores fabricaron armas cada vez más perfectas. Así surgieron el hacha y la lanza, que podían arrojar a cierta distancia, con más seguridad. **Éstas** les dieron ventajas a los hombres en la lucha con los animales.*

*Nadie sabe cuándo fue el momento exacto en el que los seres humanos descubrieron cómo utilizar el fuego, pero desde ese descubrimiento todo cambió. Se empezaron a cocinar los alimentos y esto fue un avance espectacular: el fuego acababa con los parásitos de la carne y la comida se volvió más saludable. También se pudieron aprovechar muchos vegetales como el arroz, que crudo **no resultaba nada atractivo**.*

Responder las siguientes preguntas.

Redondear la respuesta correcta.

1. ¿Por qué se dice que la vida debía ser dura?

- A. Porque debían desplazarse continuamente para sobrevivir.
- B. Porque no tenían televisores ni videojuegos.

- C. Porque los chicos tenían que salir a cazar desde muy pequeños.
- D. Porque solo podían comer frutas.

2. ¿A qué se llamaba caza mayor?

- A. A la caza de ratas, murciélagos e insectos.
- B. A la caza que realizaba la gente mayor de la tribu.
- C. A la casa más grande que había en la tribu.
- D. A la caza de osos, monos, elefantes, ciervos, jabalíes y mamuts.

3. ¿Por qué las tribus debían desplazarse continuamente?

- A. Porque necesitaban ir en busca de alimento.
- B. Porque debían huir de otras tribus salvajes.
- C. Porque de esa manera rendían culto a los dioses.
- D. Porque se aburrían de vivir en un mismo lugar.

4. Según el texto, ¿cuáles fueron las ventajas del hacha y de la lanza?

- A. Permitted mejorar las artesanías que construían.
- B. Dio ventajas al hombre para luchar contra los animales.
- C. Permitted que las tribus intercambiaran objetos.
- D. Dio la posibilidad de armar ejércitos.

5. ¿Por qué se dice que cuando apareció el fuego, la comida fue más saludable?

- A. Porque espantaba a los insectos que se acercaban a la comida.
- B. Porque como cocinaban la carne, era más sabrosa y ya no tenía parásitos.

- C. Porque el fuego les permitía abrigarse cuando hacía mucho frío.
- D. Porque como ya no estaban a oscuras, podían cocinar mejor.
6. ¿Por qué eran los hombres los encargados de cazar?
- A. Porque no les gustaba pescar.
- B. Porque no querían que las mujeres trabajaran.
- C. Porque sólo había hombres en las tribus.
- D. Porque eran más fuertes para enfrentar a los animales.
7. Cuando dice “para poder hacer frente a animales mucho más grandes que ellos...”, ¿a qué se refiere **ellos**?
- A. A los animales grandes.
- B. A los escarabajos, larvas y gusanos.
- C. A los cazadores.
- D. A los instrumentos de caza.
8. Cuando descubrieron el fuego...
- A. utilizaron el arroz para poder encender fogatas.
- B. dejaron de comer arroz.
- C. se trasladaron de un lugar a otro en busca de arroz.
- D. empezaron a comer arroz.
9. En la oración “Éstas les dieron ventajas a los hombres en la lucha contra los animales”, ¿a qué se refiere **éstas**?
- A. A las ventajas.
- B. Al hacha y a la lanza.
- C. A la lucha.
- D. A la distancia.

10. La frase “...no resultaba nada atractivo”, resaltada en las líneas 20 y 21 del texto, significa...

- A. Era atractivo.
- B. Era algo atractivo.
- C. No era atractivo.
- D. Era un poco atractivo.