

**MONITOREO CON BASE EN EL CURRÍCULO EN LA ESCUELA INCLUSIVA:  
ADECUACIÓN TÉCNICA DE LAS PRUEBAS MAZE EN LA CLASIFICACIÓN DEL  
RIESGO DE DISLEXIA**

**MONITORIZAÇÃO COM BASE NO CURRÍCULO NA ESCOLA INCLUSIVA:  
ADEQUAÇÃO TÉCNICA DAS PROVAS MAZE NA TRIAGEM DO RISCO DE  
DISLEXIA**

**CURRICULUM-BASED MEASUREMENT WITHIN INCLUSIVE SCHOOL:  
TECHNICAL ADEQUACY OF MAZE PROBES FOR SCREENING STUDENTS AT  
RISK FOR DYSLEXIA**

Paula Marisa Fortunato VAZ<sup>1</sup>  
Ana Paula Loução MARTINS<sup>2</sup>

**RESUMO:** Neste estudo descrevem-se provas de monitorização com base no currículo (MBC) para a compreensão da leitura (MBC-Maze). Partindo-se da existência de cinco procedimentos de cotação das mesmas, e no contexto da triagem universal do Modelo de Resposta à Intervenção, realizou-se um estudo que teve por finalidade testar a administração, a cotação e a adequação técnica de uma prova de MBC-Maze na identificação de alunos em risco de apresentarem Dislexia. Assim, analisou-se o impacto dos cinco procedimentos de cotação: a) na validade concorrente da prova; b) no poder discriminativo de verdadeiros positivos (sensibilidade) e de verdadeiros negativos (especificidade) da prova. Monitorizaram-se em três momentos 82 alunos do 3.º ano de escolaridade de um agrupamento de escolas do norte de Portugal. Os dados analisados com recurso à estatística descritiva, inferencial, correlacional e à análise ROC permitem concluir que: a) os resultados obtidos na prova Maze são válidos independentemente do procedimento de cotação usado; b) no início do ano letivo não é indiferente o procedimento de cotação que é utilizado, c) a sensibilidade da prova MBC-Maze, para predizer no início do ano letivo se um aluno terá risco de vir a apresentar uma leitura muito insatisfatória no final do ano letivo variou entre 93% e 96%, consoante o procedimento de cotação. O impacto social deste estudo manifesta-se na chamada de atenção para a importância da identificação de alunos em risco de insucesso na leitura, na disseminação dum sistema de triagem rápido, fiável e económico no contexto dum modelo de Resposta à Intervenção.

**PALAVRAS-CHAVE:** Monitorização-com-base-no-curriculo. Risco. Compreensão-leitora. Triagem. Provas maze.

<sup>1</sup> Centro de Investiga o em Educa o B sica, Instituto Polit cnico de Bragan a (CIEB/IPB), Bragan a - Portugal. Professora Adjunta, Departamento de Psicologia. ORCID: <https://orcid.org/0000-0001-7678-6781>. E-mail: paulavaz@ipb.pt

<sup>2</sup> Centro de Investiga o em Educa o, Universidade do Minho – (CIED/UM), Braga - Portugal. Professora Auxiliar, Departamento de Psicologia de la Educaci o y Educaci o Especial. ORCID: <https://orcid.org/0000-0002-2848-5754>. E-mail: apmartins@ie.uminho.pt

**RESUMEN:** *En este estudio se describen pruebas de monitoreo con base en el currículo (MBC) para la comprensión de lectura (MBC-Maze). Partiendo de la existencia de cinco procedimientos de cotización, y en el contexto de segmentación universal del Modelo de Respuesta a la Intervención, se realizó un estudio que tuvo por finalidad testar la administración, la cotización y la adecuación técnica de una prueba de MBC-Maze en la identificación de alumnos con riesgo de presentar Dislexia. Así, se analizó el impacto de los cinco procedimientos de cotización: a) en la validez concurrente de la prueba; b) en el poder discriminativo de verdaderos positivos (sensibilidad) y de verdaderos negativos (especificidad) de la prueba. Se monitoreó en tres momentos a 82 alumnos del 3º año de escolaridad de un agrupamiento de escuelas del norte de Portugal. Los datos analizados con recurso a la estadística descriptiva, inferencial, correlacional a el análisis ROC permiten concluir que: a) los resultados obtenidos en la prueba Maze son válidos independientemente del procedimiento de cotización utilizado; b) en el inicio del año lectivo no es indiferente el procedimiento de cotización que es utilizado; c) la sensibilidad de la prueba MBC-Maze, para predecir en el inicio del año lectivo si un alumno tendrá riesgo de presentar una lectura muy insatisfactoria en el final del año lectivo varió entre 93% y 96%, en conformidad con procedimiento de cotización. El impacto social de este estudio se manifiesta en la llamada a la atención a la importancia de identificación de los estudiantes en riesgo de fallo de lectura, en la difusión de un sistema de un examen fiable y rentable en el contexto de un modelo de respuesta a la intervención.*

**PALABRAS CLAVE:** *Monitoreo-con-base-en-el-curriculum. Riesgo. Comprensión-lectora. Clasificación. Pruebas maze.*

**ABSTRACT:** *This study describes curriculum-based measurement probes (CBM) for reading comprehension (CBM-Maze). Based on the existence of five scoring procedures, and in the context of the universal screening within Response to Intervention Educational Model, a study was conducted that aimed to test the administration, scoring and technical adequacy of a CBM-Maze probe in the identification of students at risk for Dyslexia. Thus, the impact of the five scoring procedures was analyzed: a) on the concurrent validity of the test; b) in the discriminative power of true positives (sensitivity) and true negatives (specificity) of the test. In three moments, 82 students from the 3rd year of schooling of a group of schools in northern Portugal were monitored. The data analyzed using descriptive, inferential, correlational statistics and ROC analysis allow us to conclude that: a) the results obtained in the CBM-Maze probe are valid regardless of the used scoring procedure ; b) at the beginning of the school year is not indifferent the scoring procedure that is used; c) the sensitivity of the CBM-Maze probe, to predict at the beginning of the school year whether a student will be at risk of presenting a very unsatisfactory reading at the end of the school year ranged between 93% and 96%, depending on the scoring procedure. The social impact of this study is manifested in the call for attention to the importance of identification of students at risk of reading failure, in the dissemination of a system of reliable and cost-effective screening in the context of an Intervention Response model.*

**KEYWORDS:** *Curriculum-based measurement. Risk. Reading comprehension. Screening. Maze probes.*

## Introdução

El Modelo de Respuesta a la Intervención es un modelo multinivel que presupone la existencia de un sistema de evaluación universal, de monitoreo regular del progreso académico de los alumnos y de un apoyo precoz e intenso para los alumnos que no responden positivamente a la enseñanza eficaz que se imparte en el aula regular. Así, la implementación de este modelo permite la identificación de alumnos considerados en riesgo, antes de que sus dificultades se agravaran y se experimentaran fracaso por un largo período de tiempo, desde una actuación precoz y preventiva consubstanciada en intervenciones rigurosas y fidedignas de elevada calidad (FUCHS; MOCK; MORGAN; YOUNG, 2003).

En el contexto de este Modelo de Respuesta a la Intervención, al monitoreo con base en el currículum (MBD) asume especial importancia. La MBC permite monitorear el progreso de los alumnos, en áreas como la lectura y la escritura y las matemáticas, de forma viable y técnicamente robusta, fácil y rápida. Se realiza con miras a la obtención del nivel de realización y de estimativas de las tasas de crecimiento de los alumnos, a la identificación de alumnos que no progresan de forma adecuada (alumnos en riesgo), a la comparación de la eficacia de diferentes formas de enseñanza e intervenciones y a la elaboración de programas más eficaces e individualizados para los alumnos en riesgo (FUCHS; FUCHS, 2007b). En el área de la lectura, la MBC se puede efectuar recurriendo a diferentes tipos de prueba en función del año de escolaridad o del año del nivel del alumno. Así, puede recorrirse a pruebas de identificación del sonido de las letras (MBC-letras), a pruebas de identificación de palabras (MBC-palabras), a pruebas de lectura oral (MBC-Oral) y a pruebas de comprensión de la lectura (MBC-Maze) (FUCHS; FUCHS, 2007a).

En este artículo analizaremos el uso de la prueba de MBC-Maze en un sistema de evaluación universal para la lectura. Se subraya que un sistema recurre a pruebas de MBC-Maze para monitorear el nivel de comprensión de lectura (BROWN-CHIDSEY; DAVIS; MAYA, 2003; BUSCH; LEMBKE, 2005; SHINN; SHINN, 2002), pero de su administración y subsecuente cotación, realizadas de modo estandarizado, resultan, también, indicadores generales de la realización de los alumnos en el área de la lectura oral (BUSCH; LEMBKE, 2005). La MBC-Maze es, por lo tanto, un indicador global de la performance de los alumnos en la lectura y no mide solo la comprensión, lo que ha sido mostrado por las elevadas correlaciones de estas pruebas, ya sea con medidas de fluencia, o con otras medidas de comprensión de lectura (BUSCH; LEMBKE, 2005).

En la prueba MBC-Maze, se fornece al alumno un trecho de un texto en el que la primera frase está intacta y, a partir de la segunda frase, cada séptima palabra fue remplazada por tres palabras dentro de paréntesis, siendo una de ellas la palabra original (la que hace sentido en el texto) y las demás, palabras que no hacen sentido en el texto (funcionan como distractores) (FUCHS; FUCHS, 2008; SHINN; SHINN, 2002). El alumno tiene que leer el texto, a la medida que marca las palabras que hacen sentido, durante 1 minuto (WILEY; DENO, 2005), dos minutos (BROWN-CHIDSEY; JOHNSON; FERNSTROM, 2005; PIERCE; MCMASTER; DENO, 2010), dos minutos y 30 segundos (BUSCH; LEMBKE, 2005; FUCHS; FUCHS, 2007a; STECKER; LEMBKE; FOEGEN, 2008), tres minutos (BUSCH; LEMBKE, 2005; GRANEY; MARTÍNEZ; MISSALL; ARICAK, 2010; RICHARDSON; HAWKEN; KIRCHER, 2012; SHIN; DENO; ESPIN, 2000; SHINN; SHINN, 2002) o cuatro minutos (ESPIN; WALLACE; LEMBKE; CAMPBELL *et al.*, 2010). La elección de la opción correcta señala que el alumno comprendió el significado del texto (BROWN-CHIDSEY; DAVIS; MAYA, 2003).

En cada momento de monitoreo se puede administrar una prueba con un texto (SHIN; DENO; ESPIN, 2000), una prueba con dos textos (ESPIN; WALLACE; LEMBKE; CAMPBELL *et al.*, 2010), o una prueba con tres textos (BUSCH; LEMBKE, 2005; WILEY; DENO, 2005). La administración de tres textos (la más utilizada) presupone el cálculo de la mediana de los resultados obtenidos (WILEY; DENO, 2005) y la administración de dos textos presupone el cálculo de la media (ESPIN; WALLACE; LEMBKE; CAMPBELL *et al.*, 2010). Mediana o media, respectivamente, serán utilizadas en los análisis subsecuentes. Son pruebas de lectura silenciosa (BROWN-CHIDSEY; DAVIS; MAYA, 2003; BUSCH; RESCHLY, 2007; RICHARDSON; HAWKEN; KIRCHER, 2012; SHINN; SHINN, 2002; WILEY; DENO, 2005), sin embargo se pueden realizarlas en voz alta si aplicadas individualmente (HALE; HAWKINS; SHEELEY; REYNOLDS *et al.*, 2011). Por lo general, son efectuadas en pequeño o gran grupo (grupos completos, en el aula regular), pero también pueden ser individualmente (BUSCH; LEMBKE, 2005; SHINN; SHINN, 2002).

En el contexto de la utilización de estas pruebas se puede recurrir a uno de cinco procedimientos para las cuotas (PIERCE; MCMASTER; DENO, 2010). Uno de los procedimientos de cotización consiste en contar las respuestas correctas (RC) y es utilizado por ejemplo en los estudios de Stecker, Fuchs y Fuchs (2005), Shin, Deno y Espin (2000) y Wiley y Deno (2005). Un segundo procedimiento es descrito por Pierce *et al.*, (2010) y consiste en sustraer al número de respuestas correctas, el número de respuestas incorrectas (RC-RI). Brown-Chidsey *et al.*, (2003) describen un tercer procedimiento de cotización, en el

cual mitad del número de respuestas incorrectas es sustraído al número de respuestas correctas (RC-½RI). Un cuarto procedimiento consiste en interrumpir la cotización con el apareamiento de tres equívocos consecutivos y contar las respuestas correctas (RC3E) dadas hasta este momento (BUSCH; LEMBKE, 2005). Finalmente, el quinto procedimiento consiste en interrumpir la cotización después del apareamiento de los errores consecutivos y contar las respectivas respuestas correctas (RC2E) hasta esta parte del texto (PIERCE; MCMASTER; DENO, 2010). Según Pierce et al. (2010), los cuatro procedimientos más elaborados tienen en cuenta reducir el efecto de la posible respuesta correcta dada de forma aleatoria, o sea, adivinada. Independientemente del procedimiento de cotización se consideran respuestas correctas las que los alumnos circundan la palabra que hace sentido en el texto (SHINN; SHINN, 2002). Se consideran equívocos las respuestas incorrectas, las respuestas en blanco (SHINN; SHINN, 2002), dos respuestas circundadas o cuando, por cualquier otra razón, no se tiene por cierto relativamente a la respuesta (BUSCH; LEMBKE, 2005).

En el contexto del Modelo de Respuesta a la Intervención Deno *et al.*, (2009) sugieren que en la evaluación universal (realizada en el inicio, medio y final del año lectivo) se utilice el valor del percentil 20 de los resultados obtenidos con la MBC para identificarse los alumnos en riesgo, justificando que los 20% de alumnos con resultados más bajos, además de incluir a los alumnos con más dificultades, se traducen en un número de alumnos que se considera razonable monitorear individualmente. Se asume, por lo tanto, que la baja realización relativamente a un punto de referencia, obtenido en un determinado período de año escolar, es una evidencia de que un alumno está teniendo fracaso y necesita de intervención preventiva. En este contexto, para que un sistema de evaluación universal sea eficaz, los procedimientos que lo constituyen deben identificar una gran porcentaje de verdaderos positivos y limitar la existencia de falsos positivos (FUCHS; FUCHS, 2007b). Johnson, Jenkins, Petscher e Catts (2009) señalan que en la literatura se identifican tres características que se deben considerar en un instrumento de evaluación que aumente la posibilidad de identificar verdaderos positivos y disminuya la posibilidad de identificarse falsos positivos. Así el instrumento:

1) *Es fiable*: se relaciona con la sensibilidad y la especificidad. La sensibilidad de un instrumento representa el grado en que identifica efectivamente como estudiantes en riesgo a aquellos que obtendrán logros insatisfactorios en un instrumento referido al criterio (JENKINS; HUDSON; JOHNSON, 2007). Estos estudiantes son considerados como verdaderos positivos, es decir, aquellos que corren el riesgo de fracasar académicamente en el área evaluada. Un instrumento con buena sensibilidad también reduce el número de falsos

negativos, lo que es fundamental en un modelo de respuesta a la intervención, en el que los estudiantes en situación de riesgo disfrutan de una intervención adicional a la enseñanza practicada en el aula. Cabe destacar que la sensibilidad depende del punto de corte que se utilice (JOHNSON; JENKINS; PETSCHER; CATTS, 2009). La *especificidad* de un instrumento indica el grado en que identifica efectivamente como estudiantes que no corren riesgo a aquellos que lograrán resultados satisfactorios en un instrumento referido al criterio. Estos estudiantes son considerados como verdaderos negativos. Una prueba con buena especificidad también reduce el número de falsos positivos, lo que es fundamental en un modelo de respuesta a la intervención, ya que a estos estudiantes se les proporciona una intervención adicional que no es necesaria y que consume tiempo y dinero (JOHNSON; JENKINS; PETSCHER; CATTS, 2009).

2) *Es práctico*: se refiere al instrumento corto, sencillo, barato y fácil de aplicar y citar para identificar a los estudiantes que están teniendo logros por debajo de sus compañeros. Esto permite una aplicación a gran escala por parte de diversos profesionales y puede llevarse a cabo en el aula. Además, esta característica del instrumento tiende a reducir la ansiedad de los estudiantes y facilita el cálculo e interpretación de los resultados (JOHNSON; JENKINS; PETSCHER; CATTS, 2009). Busch y Lembke (2005) señalan que una de las razones para que la MBC-Maze sea útil para un sistema educacional es el hecho de ser fácil y rápido el proceso de realización, cotización y visualización de los resultados.

3) *Tiene validez consecuyente (Consequential Validity)*: relaciona el instrumento de detección con el efecto positivo o negativo que los resultados tienen en el sistema de apoyo multinivel. Esto significa que mediante la aplicación de reglas de análisis de los resultados existe un vínculo entre los resultados de la vigilancia y una posterior intervención de apoyo que es eficaz y sólo para aquellos estudiantes que lo necesitan (JOHNSON; JENKINS; PETSCHER; CATTS, 2009).

Acorde con la revisión realizada por Jenkins; Hudson y Johnson (2007), el contexto de un sistema de detección, el fracaso en la lectura ha sido clasificado como *lectura insatisfactoria* y *lectura muy satisfactoria*. Se consideran con *lectura insatisfactoria* los alumnos que obtienen un resultado que se sitúa: a) un año abajo del año escolar, b) abajo del percentil 25, o c) debajo de un determinado punto de corte. Dependiendo de cómo se define, esta clasificación de *lectura insatisfactoria* tiene la capacidad de identificar un largo porcentaje de alumnos en riesgo (entre 50% y 75%). Adicionalmente, cuando el objetivo de un sistema de detección es encontrar los alumnos con una *lectura muy insatisfactoria*, es decir, aquellos alumnos con una realización que sitúa el alumno abajo el percentil 10. Este

criterio permite identificar los alumnos con las peores realizaciones del grupo, es decir, los que se sospecha poder tener una dificultad de aprendizaje específica en la lectura (Dislexia) (JENKINS; HUDSON; JOHNSON, 2007). Estos casos pueden ser alumnos que en el futuro puedan tener un diagnóstico de dislexia y que, en un modelo de respuesta a la intervención, deben beneficiar de intervenciones individualizadas, como las centradas en la tutoría instruccional (véase el estudio de MACHADO; CAPELLINI, 2014).

Así, en el contexto de la evaluación universal del Modelo de Respuesta a la Intervención se desarrolló en Portugal un estudio que tuvo por finalidad testar la administración, la cotización y la adecuación técnica de una prueba de MBC-Maze en la identificación de alumnos en riesgo de presentar Dislexia. Emergen de esta finalidad dos objetivos que se relacionan con el conocimiento del impacto de cinco procedimientos de cotización de la prueba MBC-Maze: a) en la validez concurrente de la prueba; b) en el poder discriminativo de verdaderos positivos (sensibilidad) y de verdaderos negativos (especificidad) de la prueba.

## **Método**

### *Población*

82 alumnos del 3º grado del 1º Ciclo de Enseñanza Básica de un Agrupamiento de Escuelas del Norte de Portugal con 8 años de edad, sin necesidades educativas especiales identificadas, sin apoyo de los servicios de educación especial y distribuidos por seis grupos participaron de esta investigación. De estos, 47 (57,3%) son del género femenino y 35 (42,7%) del género masculino, que constituyen la población del dicho año y ciclo de enseñanza del agrupamiento.

### *Instrumento de recopilación de datos*

En esta investigación se utilizó dos instrumentos de recopilación de datos: (1) una prueba de MBC para la comprensión de la lectura – prueba MBC-Maze (BROWN-CHIDSEY; DAVIS; MAYA, 2003; BUSCH; LEMBKE, 2005; SHINN; SHINN, 2002; STEVENSON, 2017); y (2) el Test de Edad de Lectura (TIL), que evalúa la comprensión de la lectura (SUCENA; CASTRO, 2010), que se describen a continuación.

*1) Prueba de MBC para la comprensión de la lectura – prueba MBC Maze: Su*

preparación se inició con un trabajo previo de investigación de textos de manuales escolares del 3º año de escolaridad, diferentes del adoptado en el agrupamiento de escuelas, con miras a seleccionar lo que, por sus características, mejor se adecuara a las reglas de elaboración de la misma. En este sentido, tras seleccionado el texto y habiendo la indicación de que éste no había sido trabajado en ninguno de los grupos, se decidió por la elaboración de la prueba con base en el mismo y se acordó con los docentes que estos no lo trabajarían en sus clases. También se tuvo el cuidado de elegir un texto que no fuera alusivo a temas específicos, como por ejemplo las estaciones del año o cuadas festivas.

Elegido el texto, empezó la construcción de la prueba, respetando las reglas que se siguen: (1) Dejar intacta la primera frase; (2) Borrar, a partir de la primera frase, cada séptima palabra y en su lugar poner tres palabras: la palabra original que pertenece al texto (palabra correcta) y dos otras que no hacen sentido en el texto (distractores); (3) Estar seguro que ninguno de los distractores hace sentido en el texto. Solo la palabra correcta puede hacer sentido en el texto. Relativamente a los distractores, se tuvieron en cuenta las siguientes reglas: (a) No elegir los distractores al azar. Uno de ellos debe ser del mismo tipo de la palabra correcta (nombre, verbo, adverbio) y el otro una palabra sacada del texto al azar; (b) Elegir distractores de la misma amplitud que la palabra correcta; (c) Seleccionar distractores que no rimen con la palabra correcta; (d) Elegir distractores que no empiecen por la misma letra de la palabra correcta (BUSCH; LEMBKE, 2005). Se consideran, también, algunas sugerencias adicionales según Busch y Lembke (2005): (1) Subrayar y poner en negrita las tres palabras (la correcta y los distractores); (2) No dejar separadas las tres palabras en el final de las líneas del texto; (3) Pasar para la palabra anterior o para la palabra siguiente cuando la séptima palabra sea un nombre; (4) Variar la localización de la palabra correcta relativamente a los distractores; (5) Escribir las tres palabras con letra mayúscula, cuando la séptima palabra sea la primera palabra de una frase.

La prueba es constituida por una hoja de portada que contiene un ejemplo constituido por tres frases de entrenamiento. Se sigue el texto con 33 situaciones (resultado máximo) en que el alumno debe seleccionar la palabra que hace sentido. Se definió el límite de tiempo de tres minutos y se optó por la realización colectiva en cada grupo. Las pruebas fueron cotizadas recurriendo a los cinco procedimientos de cotización anteriormente descriptos.

2) *Test de Edad de Lectura (TIL)*: Este test consiste en una adaptación del sub-test Lobrot L3, que es una prueba estandarizada para el francés y de utilización generalizada por investigadores francófonos, más bien en el área de la dislexia y de la psicología escolar. Se trata de un test que evalúa la lectura y que puede ser utilizado para la realización de un primer



diagnóstico de dificultades de lectura/dislexia. El TIL fue elaborado para poder ser administrado de forma colectiva. Durante su realización, el alumno tiene que leer silenciosamente frases aisladas, que no están completas, y elegir una de las cinco palabras que se encuentran dentro de paréntesis para completar cada frase. Para ello deben subrayar la palabra elegida. De la prueba hacen parte cuatro frases para los alumnos entrenar (presentes en una primera página) y 36 frases experimentales (escritas en la segunda página y distribuidas por dos columnas, de forma equitativa). El tiempo destinado a la realización del TIL es de 5 minutos y la cotización consiste en la suma de las frases completadas correctamente, en la multiplicación por 100 del número obtenido y en la división del producto resultante por 36 (número total de frases (SUCENA; CASTRO, 2010)).

### *Procedimientos de recopilación de datos*

Tras la aprobación del proyecto de investigación por el Consejo Científico de una Universidad Portuguesa y de obtenidas las autorizaciones de las instituciones educativas se procedió a la entrega a los encargados de educación de una declaración de autorización de participación en la investigación por parte de sus estudiantes. Todos los encargados de educación autorizan sus estudiantes a participar en la investigación. Adicionalmente, se explicó a los alumnos qué se pretendía de ellos y dada libertad, en los diferentes monitoreo, para realizar o no la prueba, habiendo todos participado.

Con la prueba MBC-Maze, ocurrió un primer monitoreo en el otoño (inicio del año lectivo), una segunda monitorización, aproximadamente dos semanas después (un retest para análisis de fiabilidad), una tercera monitorización en el invierno y una cuarta monitorización en la primavera (final del año lectivo). Adicionalmente, en la primavera tuvo lugar una evaluación con el TIL.

Para garantizar la fiabilidad de la administración de ambas las pruebas se llenó por los profesores el documento de *Validación de la implementación de la prueba Maze* de (PATRÃO, 2010) en todas las monitorizaciones con la prueba Maze y el documento de *Validación de la implementación del TIL*<sup>3</sup> en las evaluaciones con el TIL. Las pruebas y el test fueron administrados y cotizados por el investigador principal de este estudio.

### *Procedimientos de análisis y de datos*

<sup>3</sup> Preparado en el contexto de este estudio sobre la base del documento de PATRÃO (2010).

Considerando los resultados obtenidos en las diferentes monitorizaciones se evaluó los supuestos de aplicabilidad, más bien, homogeneidad e independencia de los residuos obtenidos, con recurso al test de Kolmogorov-Smirnov para la normalidad, test de Levene para la homogeneidad de las variancias y para la independencia del gráfico de los residuos. La fiabilidad y la validad de la prueba fueron evaluadas recurriendo al coeficiente de correlación Pearson ( $r$  de Pearson). Se hizo un análisis ROC (*Receiver Operating Characteristic*), con metodología DeLong et al. (1988), para obtención de estadísticas de eficiencia de diagnóstico de la prueba de MBC-Maze (sensibilidad, especificidad, valor predictivo positivo, valor predictivo negativo y porcentaje de clasificaciones correctas). Para el cálculo de la sensibilidad y de la especificidad, como en el caso *s no existe un “e ar oro”* que tenga condiciones de realmente afirmar la condición existente o no en un determinado alumno, se utilizó un criterio resultante de la aplicación del TIL (SUCENA; CASTRO, 2010), que se consideró el test de criterio. Teniendo por objetivo encontrar los alumnos estándar oro, o sea, como variable que indicó quién tenía y no tenía una *comprensión en la lectura muy insatisfactoria*. Este criterio permitió identificar los alumnos con realizaciones peores del grupo, los que se sospecha poder tener dislexia.

## Resultados

Se sigue la presentación de los resultados obtenidos, cuando utilizados los cinco procedimientos de cotización<sup>4</sup>, teniendo por base los objetivos del estudio. Se subraya que están verificados los supuestos de aplicabilidad de testes paramétricos (PESTANA; GAGEIRO, 2008).

a) Impacto de cinco procedimientos de cotización en la validad concurrente de la prueba.

Con miras al estudio de la validad concurrente de cada uno de los procedimientos de cotización, se recurrió al  $r$  de Pearson para calcular los coeficientes de correlación entre los resultados obtenidos en la prueba CBM Maze (con la utilización de cada uno de los cinco procedimientos de cotización) realizada en la primavera y los resultados obtenidos en el TIL

---

<sup>4</sup> Procedimiento de cotización nº 1 (Número de respuestas correctas menos las incorrectas, RC-RI), procedimiento de cotización nº2 (Número de respuestas correctas menos mitas de las incorrectas, RC-½RI), procedimiento de cotización nº3 (Número de respuestas correctas con interrupción de la cotización después de tres años consecutivos, RC3E), procedimiento de cotización nº 5 (Número total de respuestas correctas, RC).

administrado también en la primavera. Como se puede observar en Cuadro 1, se verifica que los coeficientes de correlación obtenidos varían entre .776 y .806, valores indicativos de una asociación linear alta (PESTANA; GAGEIRO, 2008).

**Cuadro 1** - Correlaciones entre Maze e TIL por procedimiento de cotización

Procedimiento	<i>r</i> de Pearson (* significativo a 1%)
RC-RI	.776*
RC-½RI	.794*
RC2E	.793*
RC3E	.780*
RC	.806*

Fuente: Elaborado por las autoras.

El coeficiente de correlación más elevado es obtenido para el procedimiento de cotización RC, y el más bajo para el procedimiento RC-RI. Significa eso que, teóricamente se podrá tener más confianza en los resultados obtenidos con la utilización de los procedimientos de cotización RC. Sin embargo, los resultados obtenidos con los diferentes procedimientos de cotización son mayoritariamente indicativos de asociaciones lineales altas, habiendo solo un resultado indicativo de asociación lineal moderada (PESTANA; GAGEIRO, 2008), y todos son significativos a 1%. Y eso quiere decir, en la práctica, los resultados obtenidos con todos los procedimientos de cotización son fiables.

b) Impacto de cinco procedimientos de cotización en el poder discriminativo de verdaderos positivos (sensibilidad) y verdaderos negativos (especificidad) de la prueba.

Los resultados del *Test ANOVA a 1 factor* muestran que hay diferencias estadísticamente significativas entre procedimientos de cotización solo en la monitorización del otoño (ET= 2.655, para  $p=.033$ ). En las subsecuentes monitorizaciones, realizadas a lo largo del año lectivo, no hay diferencias estadísticamente significativas entre los procedimientos de cotización. El test *pos-hoc Tukey* evidenció que los procedimientos de cotización que difieren estadísticamente entre sí en el otoño, es decir, en el comienzo del año lectivo, son el procedimiento RC-RI y el procedimiento RC.

Las pruebas de MBC-Maze tuvieron muy buena aceptación por parte de los profesores, de las familias y de los alumnos involucrados, habiendo todos autorizado y/o participado. La prueba se mostró sencilla y práctica en la aplicación en grupo. El trabajo de administración de las pruebas llevó cerca de 15 minutos en cada grupo, tiempo que comprendió las informaciones a los alumnos sobre lo que se realizaría a continuación, la

obtención de su consentimiento verbal, su organización en el aula, la lectura de las respectivas instrucciones, la realización de las pruebas y su recogida. Este tiempo fue disminuido a lo largo del año lectivo y a la medida que los alumnos se familiarizaban con los procedimientos. Independientemente del procedimiento de cotización, cada prueba pudo ser corregida en menos de cinco minutos.

En el Cuadro 2 se presentan las cinco estadísticas de eficiencia de diagnóstico, más bien sensibilidad (S), especificidad (E), valor predictivo positivo (VPP), valor predictivo negativo (VPN) y porcentaje de clasificaciones correctas (PCC) para el punto de corte (PC) relativo al percentil 20 del inicio y del final de año lectivo, que fue lo considerado para la identificación de alumnos en riesgo, tal como sugerido por (DENO, 1985).

**Cuadro 2** - Eficacia de diagnóstico de la prueba MBC-Maze

Procedimiento	PC	S (%)	E (%)	VPP (%)	VPN (%)	PCC (%)
RC-RI	< 4	94	33	89	50	85
	< 9	93	25	88	38	83
RC-½RI	< 5	96	38	89	62	87
	< 10	94	31	88	50	84
RC2E	< 5	93	27	89	38	84
	< 10	96	38	89	63	87
RC3E	< 6	94	33	89	50	85
	< 11	95	36	88	63	85
RC	< 6	94	36	91	50	87
	< 11	96	38	89	63	87

Fuente: Elaborado por las autoras.

El procedimiento de cotización que presentó una mayor sensibilidad en el inicio del año fue el RC-½RI. La sensibilidad de la prueba MBC-Maze, cotizada con el procedimiento RC-½RI, para predecir, en el inicio del año lectivo, si un alumno estará en riesgo de venir a presentar una lectura muy insatisfactoria en el final del año lectivo, es de 96%. La especificidad de la prueba, cuando cotizada con este procedimiento de cotización, o sea, la posibilidad de excluir del grupo de riesgo los alumnos que tendrán una lectura satisfactoria y, consecuentemente, una intervención innecesaria, es de 38%. En el inicio del año lectivo, este es uno de los procedimientos más eficaces (87% de clasificaciones correctas) para ser utilizados en la cotización del instrumento de clasificación.

## Consideraciones finales

La utilización de un sistema universal de clasificación para la detección precoz del riesgo de dislexia es uno de los más grandes y más promisoros desafíos de una reforma de la educación. La clasificación determina el nivel de riesgo para problemas de lectura en general y el riesgo potencial de tener o desarrollar dislexia. Los procedimientos de clasificación para el riesgo de dislexia deben ser eficientes y baratos y deben ser utilizados para todos los alumnos en aula de clase.

Este estudio testó la administración, la cotización y la adecuación técnica de una prueba de MBC-Maze en la identificación de alumnos en riesgo de presentar dislexia y analizó el efecto de diferentes procedimientos de cotización de esta prueba. A partir de los resultados obtenidos fue posible sacar las conclusiones que se presentan a continuación:

### **1. Los resultados obtenidos en la prueba Maze son válidos independientemente del procedimiento de cotización utilizado**

En lo que respecta a la validez, se estudió la validez concurrente. Los coeficientes de correlación de la prueba Maze con el TIL obtenidos en la primavera, cuando utilizados los cinco procedimientos de cotización, remiten al hecho de los resultados en la prueba Maze ser válidos, dado que todos los dichos coeficientes (entre .776 y .806) son indicativos de asociaciones lineales altas de acuerdos con Pestana y Gageiro (2008). Se trata de coeficientes de correlación compatibles a los obtenidos por Tichá, Espin y Wayman (2009) (.82 entre la prueba Maze y el MBST y .88 entre la prueba Maze y el WJ-III) y por Hale *et al.*, (2011) (.859, entre los resultados en pruebas Maze realizadas en tres minutos y los resultados en el WJ-III ACH Broad Reading Cluster Score e .860 entre los resultados en las mismas pruebas Maze y los resultados en el WCPM), dado que corresponden también a asociaciones lineales (PESTANA; GAGEIRO, 2008). Son semejantes a los encontrados por Pierce, McMaster e Deno (2010) (.73 e .82) cuando al comparar cada uno de los cinco procedimientos de cotización, correlacionaron los respectivos resultados con tres medidas de criterio (o KTEA-II Letter Word Overall, o KTEA-II Reading Comprehension Overall e Provas de Leitura Oral), globalmente superiores a los encontrados en otros estudios, a saber por Jenkins y Jewell (1993), que correlacionaron los resultados obtenidos en pruebas Maze (realizadas por alumnos del 2º al 6º grados, en un minuto, y cotizadas recurriendo al procedimiento de cotización RC3E con los resultados en tesis como el Gates-MacGinitie Reading Tests (GM-TR) y el Metropolitan Achievement Tests (MAT), para estudiar la validez concurrente de los

resultados obtenidos en las pruebas Maze), y obtuvieron correlaciones entre .65 y .76 con el primero entre .66 y .76 con el segundo.

## **2. En el inicio del año lectivo no es indiferente el procedimiento de cotización que se utiliza.**

En el ámbito de la comparación de los diferentes procedimientos de cotización, importa también destacar las diferencias estadísticamente significativas entre las medias obtenidas cuando utilizados los procedimientos de cotización RC-RI y RC, en el monitoreo del otoño, inicio de un año lectivo y en que los alumnos se encuentran todavía en un período “inicial” del aprendizaje de lectura (aunque se trate de un 3º nivel de escolaridad). Significa entonces que en este momento del año lectivo no es indiferente la opción por uno de estos dos procedimientos de cotización, debiendo el profesor ponderar su elección, también en función de otras variables que interese considerar. Ya en las monitorizaciones que tuvieron lugar en el invierno y en la primavera se puede elegir cualquier uno de los procedimientos de cotización, dado que las diferencias en los resultados obtenidos no son estadísticamente significativas.

## **3. La sensibilidad de la prueba MBC-Maze, para predecir en el inicio del año lectivo si un alumno tendrá posibilidad de presentar una lectura muy insatisfactoria en el final del año lectivo, cambio entre 93% y 96%, con relación al procedimiento de cotización.**

Aunque un sistema universal de clasificación para detección del riesgo de dislexia pueda potencialmente ayudar muchos estudiantes que se deparan con problemas en el aprendizaje de la lectura, ella también presenta riesgos y retos para los sistemas escolares y para alumnos dentro de ellos.

La implementación de un sistema universal de clasificación eficaz para detectar precozmente el riesgo de dislexia no es una cuestión sencilla de elegir y administrar un test único para seleccionar niños. Por lo contrario, ella deberá asentar en la utilización de pruebas de fácil implementación, breve realización y cotización, válidas y fiables, que puedan ser implementadas por los profesores por lo menos tres veces en el año lectivo, a todo los alumnos de todos los años de escolaridad en el 1º Ciclo de Enseñanza Básica. Un sistema eficaz de clasificación para la dislexia conlleva a riesgos asociados a dos tipos de resultados de clasificación incorrectos: aquél en el que el examinar identifica el alumno como en riesgo, verificando posteriormente que él no presenta resultados debajo de un determinado valor de referencia (error de identificación de falso positivo); aquél en el que el examinador dice que el alumno no está en riesgo, pero presenta posteriormente resultados debajo de un valor de

referencia especificado (error de identificación de falso negativo). El problema de la identificación de falsos positivos debe ser encuadrado en el contexto de los recursos existentes en la escuela. Si en el proceso de clasificación universal el examinador identifica muchos falsos positivos, puede que el contribuya para que se desperdicien recursos al estar a proporcionar intervención y/o evaluación adicional a alumnos que, en realidad, de ellos no necesitan. Así, las escuelas estarán a gestionar preciosos recursos de intervención con alumnos que efectivamente no necesitan de intervención adicional, lo que puede hacer que los recursos sean insuficientes para los alumnos que de hecho necesitan de ellos. Por lo tanto, la clasificación debe buscar proporcionar identificar correctamente 80% o más de los alumnos que no están en riesgo. Sin embargo, las consecuencias de la identificación de falsos negativos, que puede significar falta de acceso a intervenciones precoces en la lectura a alumnos que no logran suceso sin intervención, tal vez sean más graves, puesta la necesidad de los estudiantes en riesgo de dislexia recibir intervención en la lectura lo más temprano que pueda, en el período en el que la adquisición de la lectura es enfatizada.

Así se puede considerar que las pruebas de MBC-Maze revelan buenas características técnicas, validez, sensibilidad y especificidad, que vuelven segura su utilización por parte de los profesionales en el monitoreo y especificidad, que vuelven segura su utilización por parte de los profesores en el monitoreo del progreso de los alumnos en la lectura. Adicionalmente, en el contexto de un Modelo de Respuesta a la Intervención, además de la clasificación universal con recurso a pruebas de MBC, los investigadores internacionales sugieren que las escuelas implementen un apoyo multinivel, en el sentido de, progresivamente, providenciar intervenciones más intensivas, que de modo directo y explícito trabajen las necesidades académicas, comportamentales y sociales de los alumnos (BRADLEY; DANIELSON; HALLAHAN, 2002).

**AGRADECIMENTOS:** Esta investigación se desarrolló en el contexto de un proyecto financiado con una subvención de la Fundación para la Ciencia y la Tecnología (Portugal).

## REFERENCIAS

BRADLEY, R.; DANIELSON, L.; HALLAHAN, D. P. Specific learning disabilities: Building consensus for identification and classification. *In*: BRADLEY, R.; DANIELSON, L. *et al.*, (Ed.). **Identification of learning disabilities: Research to practice** Mahwah, NJ: Lawrence Erlbaum Associates, 2002.

BROWN-CHIDSEY, R.; DAVIS, L.; MAYA, C. Sources of variance in curriculum-based measures of silent reading. **Psychology in the Schools**, v. 40, n. 4, p. 363-377, 2003.

BROWN-CHIDSEY, R.; JOHNSON, P.; FERNSTROM, R. Comparison of Grade-Level Controlled and Literature-Based Maze CBM Reading Passages. **School Psychology Review**, v. 34, n. 3, p. 387-394, 2005.

BUSCH, T. W.; LEMBKE, E. S. **Teaching tutorial 5**: progress monitoring in reading using the CBM maze procedure. Division for Learning Disabilities of the Council for Exceptional Children, 2005.

BUSCH, T. W.; RESCHLY, A. L. Progress monitoring in reading: using curriculum-based measurement in a response-to-intervention model. **Assessment for Effective Intervention**, v. 32, n. 4, p. 223-230, 2007.

DENO, S. L. Curriculum-based measurement: the emerging alternative. **Exceptional Children**, v. 52, n. 3, p. 219-232, 1985.

DENO, S. L.; RESCHLY, A. L.; LEMBKE, E. S.; MAGNUSSON, D. *et al.* Developing a school-wide progress-monitoring system. **Psychology in the Schools**, v. 46, n. 1, p. 44-55, 2009.

ESPIN, C.; WALLACE, T.; LEMBKE, E.; CAMPBELL, H. *et al.* Creating a progress-monitoring system in reading for middle-school students: tracking progress toward meeting high-stakes standards. **Learning Disabilities Research & Practice**, v. 25, n. 2, p. 60-75, 2010.

FUCHS, D.; MOCK, D.; MORGAN, P. L.; YOUNG, C. L. Responsiveness-to-intervention: Definitions, evidence, and implications for the learning disabilities construct. **Learning Disabilities Research & Practice**, v. 18, n. 3, p. 157-171, 2003.

FUCHS, L. S.; FUCHS, D. The role of assessment in the three-tier approach to reading instruction. *In*: HAAGER, D.; KLINGNER, J. *et al.*, (Ed.). **Evidence-based reading practices for response to intervention**. Baltimore, Maryland: Paul H. Brookes Publishing Co., 2007a. p. 29-42.

FUCHS, L. S.; FUCHS, D. **Using CBM for progress monitoring in reading**. United States Office of Special Education Program; Student Progress Monitoring, 2007b.

FUCHS, L. S.; FUCHS, D. The role of assessment within the RTI framework. *In*: FUCHS, D.; FUCHS, L. S. *et al.*, (Ed.). **Response to intervention: A framework for reading educators**. Newark: International reading association, 2008. p. 27-49.

GRANEY, S. B.; MARTÍNEZ, R. S.; MISSALL, K. N.; ARICAK, O. T. Universal screening of reading in late elementary school: R-CBM versus CBM Maze. **Remedial and Special Education**, v. 31, n. 5, p. 368-377, 2010.

HALE, A. D.; HAWKINS, R. O.; SHEELEY, W.; REYNOLDS, J. R. *et al.* An investigation of silent versus aloud reading comprehension of elementary students using Maze assessment procedures. **Psychology in the Schools**, v. 48, n. 1, p. 4-13, 2011.



JENKINS, J. R.; HUDSON, R. F.; JOHNSON, E. S. Screening for at-risk readers in a response to intervention framework. **School Psychology Review**, v. 36, n. 4, p. 582-600, 2007.

JENKINS, J. R.; JEWELL, M. Examining the validity of two measures for formative teaching: reading aloud and maze. *Exceptional Children*. v. 59, n. 5, p. 421-432, 1993.

JOHNSON, E. S.; JENKINS, J. R.; PETSCHER, Y.; CATTIS, H. W. How Can We improve the accuracy of screening instruments? **Learning Disabilities Research & Practice**, v. 24, n. 4, p. 174-185, 2009.

MACHADO, A. C.; CAPELLINI, S. A. Aplicação do modelo de tutoria em tarefas de leitura e escrita para crianças com dislexia do desenvolvimento. **Revista Ibero-americana de Estudos em Educação**, v. 9, n. 1, p. 35-48, 2014.

PATRÃO, M. S. G. **Monitorização com base no currículo**: Um estudo quantitativo sobre a utilização de provas maze no contexto do nível I do modelo de atendimento à diversidade. Orientadora: Ana Paula Loução Martins. 2010. (Tese de Mestrado não publicada) -, Universidade do Minho, Braga.

PESTANA, M. H.; GAGEIRO, J. N. **Análise de dados para ciências sociais**: a complementaridade do SPSS. Lisboa: Edições Sílabo, 2008.

PIERCE, R.; MCMASTER, K.; DENO, S. L. The effects of using different procedures to score Maze measures. **Learning Disabilities Research & Practice (Blackwell Publishing Limited)**, v. 25, n. 3, p. 151-160, 2010.

RICHARDSON, R. D.; HAWKEN, L. S.; KIRCHER, J. Bias using Maze to predict high-stakes test performance among hispanic and spanish-speaking students. **Assessment for Effective Intervention**, v. 37, n. 3, p. 159-170, 2012.

SHIN, J.; DENO, S. L.; ESPIN, C. Technical adequacy of the Maze task for curriculum-based measurement of reading growth. **Journal of Special Education**, v. 34, n. 3, p. 164-172, Fall2000 2000. Article.

SHINN, M. R.; SHINN, M. M. **AIMSweb training workbook**: administration and scoring of reading curriculum-based measurement for use in general outcome measurement. 2002.

STECKER, P. M.; FUCHS, L. S.; FUCHS, D. Using curriculum-based measurement to improve student achievement: review of research. **Psychology in the Schools**, v. 42, n. 8, p. 795-819, 2005.

STECKER, P. M.; LEMBKE, E. S.; FOEGEN, A. Using progress-monitoring data to improve instructional decision making. **Preventing School Failure**, v. 52, n. 2, 2008.

STEVENSON, N. A. Comparing curriculum-based measures and extant datasets for universal screening in middle school reading. **Assessment for Effective Intervention**, p. 1-14, 2017.

SUCENA, A.; CASTRO, S. L. **Aprender a ler e avaliar a leitura. O TIL: teste de idade de leitura.** Coimbra: Almedina, 2010.

TICHÁ, R.; ESPIN, C.; WAYMAN, M. M. Reading progress monitoring for secondary-school students: reliability, validity, and sensitivity to growth of reading-aloud and maze-selection measures. **Learning Disabilities Research & Practice**, v. 24, n. 3, p. 132-142, 2009.

WILEY, H. I.; DENO, S. L. Oral reading and maze measures as predictors of success for english learners on a state standards assessment. **Remedial & Special Education**, v. 26, n. 4, p. 207-214, 2005.

### **Cómo referenciar este artículo**

VAZ, Paula Marisa Fortunato; MARTINS, Ana Paula Loução. Monitoreo con base en el currículo en la escuela inclusiva: adecuación técnica de las pruebas Maze en la clasificación del riesgo de dislexia. **Revista Ibero-Americana de Estudos em Educação**, Araraquara, v. 15, n. esp. 1, p. 1000-1017, maio 2020. e-ISSN: 1982-5587. DOI: <https://doi.org/10.21723/riaee.v15iesp.1.13513>

**Remitido el:** 11/10/2019

**Revisiones requeridas:** 20/11/2019

**Aprobado el:** 28/12/2019

**Publicado el:** 30/04/2020