

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

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

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

Paula Marisa Fortunato VAZ¹
Ana Paula Loução MARTINS²

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 prever 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-currículo. Risco. Compreensão-leitora. Triagem. Provas maze.

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*

¹ Centro de Investigação em Educação Básica, Instituto Politécnico de Bragança (CIEB/IPB), Bragança - Portugal. Professora Adjunta, Departamento de Psicologia. ORCID: <https://orcid.org/0000-0001-7678-6781>. E-mail: paulavaz@ipb.pt

² Centro de Investigação em Educação, Universidade do Minho – (CIED/UM), Braga - Portugal. Professora Auxiliar, Departamento de Psicologia da Educação e Educação Especial. ORCID: <https://orcid.org/0000-0002-2848-5754>. E-mail: apmartins@ie.uminho.pt

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-currículo. 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

O Modelo de Resposta à Intervenção é um modelo multinível que pressupõe a existência de um sistema de triagem universal, de monitorização regular do progresso académico dos alunos e de um apoio precoce e intenso para os alunos que não respondem positivamente ao ensino eficaz que é ministrado na sala de aula regular. Assim, a implementação deste modelo permite a identificação de alunos considerados em risco, antes de as suas dificuldades se agravarem e de experimentarem insucesso por um longo período de tempo, através de uma atuação precoce e preventiva consubstanciada em intervenções rigorosas e fidedignas de elevada qualidade (FUCHS; MOCK; MORGAN; YOUNG, 2003).

No contexto deste Modelo de Resposta à Intervenção, a monitorização com base no curriculum (MBC) assume especial importância. A MBC permite monitorizar o progresso dos alunos, em áreas como a leitura, a escrita e a matemática, de forma viável e tecnicamente robusta, fácil e rápida. É realizada com vista à obtenção do nível de realização e de estimativas das taxas de crescimento dos alunos, à identificação de alunos que não progridem de forma adequada (alunos em risco), à comparação da eficácia de diferentes formas de ensino e intervenções e à elaboração de programas mais eficientes e individualizados para os alunos em risco (FUCHS; FUCHS, 2007b). Na área da leitura, a MBC pode ser efetuada recorrendo a diferentes tipos de prova em função do ano de escolaridade ou do ano do nível do aluno. Assim, pode recorrer-se a provas de identificação do som das letras (MBC-letras), a provas de identificação de palavras (MBC-palavras), a provas de leitura oral (MBC-Oral) e a provas de compreensão da leitura (MBC-Maze) (FUCHS; FUCHS, 2007a).

Neste artigo analisaremos o uso da prova de MBC-Maze num sistema de triagem universal para a leitura. Sublinha-se que um sistema escolar recorre a provas de MBC-Maze para monitorizar o nível de compreensão da leitura (BROWN-CHIDSEY; DAVIS; MAYA, 2003; BUSCH; LEMBKE, 2005; SHINN; SHINN, 2002), mas da sua administração e subsequente cotação, realizadas de modo standardizado, resultam, também, indicadores gerais da realização dos alunos na área da leitura oral (BUSCH; LEMBKE, 2005). A MBC-Maze é, portanto, um indicador global da performance dos alunos na leitura e não mede somente a compreensão, o que tem sido mostrado pelas elevadas correlações destas provas, quer com medidas de fluência, quer com outras medidas de compreensão da leitura (BUSCH; LEMBKE, 2005).

Na prova MBC-Maze, é fornecida ao aluno uma passagem de um texto em que a primeira frase está intacta e, a partir da segunda, cada sétima palavra foi substituída por três palavras

dentro de parênteses, sendo uma delas a palavra original (a que faz sentido no texto) e as restantes, palavras que não fazem sentido no texto (funcionam como distratores) (FUCHS; FUCHS, 2008; SHINN; SHINN, 2002). O aluno tem de ler o texto, à medida que vai marcando as palavras que fazem sentido no mesmo, durante 1 minuto (WILEY; DENO, 2005), dois minutos (BROWN-CHIDSEY; JOHNSON; FERNSTROM, 2005; PIERCE; MCMASTER; DENO, 2010), dois minutos e 30 segundos (BUSCH; LEMBKE, 2005; FUCHS; FUCHS, 2007a; STECKER; LEMBKE; FOEGEN, 2008), três minutos (BUSCH; LEMBKE, 2005; GRANEY; MARTÍNEZ; MISSALL; ARICAK, 2010; RICHARDSON; HAWKEN; KIRCHER, 2012; SHIN; DENO; ESPIN, 2000; SHINN; SHINN, 2002) ou quatro minutos (ESPIN; WALLACE; LEMBKE; CAMPBELL *et al.*, 2010). A escolha da opção correta indica que o aluno compreendeu o significado do texto (BROWN-CHIDSEY; DAVIS; MAYA, 2003).

Em cada momento de monitorização pode ser administrada uma prova com um texto (SHIN; DENO; ESPIN, 2000), uma prova com dois textos (ESPIN; WALLACE; LEMBKE; CAMPBELL *et al.*, 2010), ou uma prova com três textos (BUSCH; LEMBKE, 2005; WILEY; DENO, 2005). A administração de três textos (a mais utilizada) pressupõe o cálculo da mediana dos resultados obtidos (WILEY; DENO, 2005) e a administração de dois textos pressupõe o cálculo da média (ESPIN; WALLACE; LEMBKE; CAMPBELL *et al.*, 2010). Mediana ou média, respetivamente, serão usadas nas análises subsequentes. São provas de leitura silenciosa (BROWN-CHIDSEY; DAVIS; MAYA, 2003; BUSCH; RESCHLY, 2007; RICHARDSON; HAWKEN; KIRCHER, 2012; SHINN; SHINN, 2002; WILEY; DENO, 2005), não obstante poderem realizar-se em voz alta quando aplicadas individualmente (HALE; HAWKINS; SHEELEY; REYNOLDS *et al.*, 2011). Geralmente são efetuadas em pequeno ou grande grupo (turmas inteiras, na sala de aula regular), mas também podem ser individualmente (BUSCH; LEMBKE, 2005; SHINN; SHINN, 2002).

No contexto da utilização destas provas pode recorrer-se a um de cinco procedimentos para as cotar (PIERCE; MCMASTER; DENO, 2010). Um dos procedimentos de cotação consiste na contagem das respostas corretas (RC) e é usado por exemplo nos estudos de Stecker, Fuchs e Fuchs (2005), Shin, Deno e Espin (2000) e Wiley e Deno (2005). Um segundo procedimento é descrito por Pierce *et al.*, (2010) e consiste em subtrair ao número de respostas corretas, o número de respostas incorretas (RC-RI). Brown-Chidsey *et al.*, (2003) descrevem um terceiro procedimento de cotação, no qual metade do número de respostas incorretas é subtraído ao número de respostas corretas (RC-1/2RI). Um quarto procedimento consiste em interromper a cotação com o aparecimento de três erros consecutivos e contar as respostas corretas (RC3E) dadas até esse momento (BUSCH; LEMBKE, 2005). Por fim, o quinto

procedimento consiste em interromper a cotação após o aparecimento de dois erros consecutivos e contar as respectivas respostas corretas (RC2E) até essa parte do texto (PIERCE; MCMASTER; DENO, 2010). Segundo Pierce et al. (2010), os quatro procedimentos mais elaborados visam reduzir o efeito da possível resposta correta dada de forma aleatória, ou seja, adivinhada. Independentemente do procedimento de cotação são consideradas respostas corretas aquelas em que os alunos circundam a palavra que faz sentido no texto (SHINN; SHINN, 2002). Consideram-se erros as respostas incorretas, as respostas em branco (SHINN; SHINN, 2002), duas respostas circundadas ou quando, por qualquer outra razão, não se tem a certeza relativamente à resposta (BUSCH; LEMBKE, 2005).

No contexto do Modelo de Resposta à Intervenção, Deno *et al.*, (2009) sugerem que na triagem universal (realizada no início, meio e final do ano letivo) se utilize o valor do percentil 20 dos resultados obtidos com a MBC para se identificarem alunos em risco, justificando que os 20% de alunos com resultados mais baixos, além de incluírem os alunos com mais dificuldades, se traduzem num número de alunos que se considera ser razoável monitorizar individualmente. Assume-se, portanto, que a baixa realização relativamente a um ponto de referência, obtido num determinado período do ano escolar, é uma evidência de que um aluno está a ter insucesso e necessita de intervenção preventiva. Neste contexto, para que um sistema de triagem universal seja eficaz, os procedimentos que o constituem devem identificar uma grande percentagem de verdadeiros positivos e limitar a existência de falsos positivos (FUCHS; FUCHS, 2007b). Johnson, Jenkins, Petscher e Catts (2009) indicam que na literatura se identificam três características que devem ser consideradas num instrumento de triagem que aumente a possibilidade de se identifiquem verdadeiros positivos e diminua a possibilidade de se identificarem falsos positivos. Assim, o instrumento:

1) *É fiável*: relaciona-se com a sensibilidade e especificidade. A sensibilidade de um instrumento representa o grau com que, de forma eficaz, identifica como alunos em risco aqueles que obterão realizações insatisfatórias num instrumento referenciado ao critério (JENKINS; HUDSON; JOHNSON, 2007). Estes alunos são considerados como verdadeiros positivos, ou seja, aqueles que estão em risco de insucesso académico na área avaliada. Um instrumento com boa sensibilidade também diminui o número de falsos negativos, o que é crítico num modelo de resposta à intervenção, no qual os alunos em risco usufruem de uma intervenção adicional ao ensino praticado na sala de aula. De sublinhar que a sensibilidade depende do ponto de corte que é utilizado (JOHNSON; JENKINS; PETSCHER; CATTS, 2009). A *especificidade* de um instrumento indica o grau com que, de forma eficaz, identifica como alunos que não estão em risco aqueles que obterão realizações satisfatórias num

instrumento referenciado ao critério. Estes alunos são considerados como verdadeiros negativos. Uma prova com boa especificidade também diminui o número de falsos positivos, o que é crítico num modelo de resposta à intervenção, uma vez que a estes alunos é proporcionada uma intervenção adicional que não é necessitada e que consome tempo e dinheiro desnecessários (JOHNSON; JENKINS; PETSCHER; CATTS, 2009).

2) *É prático*: relaciona-se com a forma breve, simples, económica e fácil de aplicar e cotar que o instrumento tem para identificar aqueles alunos que estão a ter realizações abaixo dos seus colegas. Tal permite uma implementação em larga escala por profissionais diversos e pode ser realizado na sala de aula. Adicionalmente, esta característica do instrumento tende a reduzir a ansiedade dos alunos e a facilitar o cálculo e a interpretação dos resultados (JOHNSON; JENKINS; PETSCHER; CATTS, 2009). Busch e Lembke (2005) salientam que uma das razões para que a MBC-Maze seja útil para um sistema educativo é o facto de ser fácil e rápido o processo de realização, cotação e visualização dos resultados.

3) *Tem validade consequencial (Consequential Validity)*: relaciona o instrumento de triagem com o efeito positivo ou negativo que os resultados têm no sistema de apoio multinível. Tal significa que mediante a aplicação de regras de análise dos resultados existe uma interligação entre os resultados da monitorização e uma intervenção de apoio posterior que é eficaz e apenas para aqueles alunos que dela necessitam (JOHNSON; JENKINS; PETSCHER; CATTS, 2009).

De acordo com a revisão feita por Jenkins; Hudson e Johnson (2007), no contexto de um sistema de triagem, o insucesso na leitura tem sido classificado como *leitura insatisfatória* e *leitura muito insatisfatória*. São considerados com *leitura insatisfatória* aqueles alunos que obtêm um resultado que se situa: a) um ano abaixo do ano escolar, b) abaixo do percentil 25, ou c) abaixo de um determinado ponto de corte. Dependendo de como é definida, esta classificação de *leitura insatisfatória* tem a capacidade de identificar uma larga percentagem de alunos em risco (entre 50% e 75%). Adicionalmente, quando o objetivo de um sistema de triagem é encontrar os alunos com uma *leitura muito insatisfatória*, ou seja, aqueles alunos com uma realização muito afastada daquela dos restantes colegas, é, muitas vezes, usado como critério o resultado que situa o aluno abaixo do percentil 10. Este critério permite identificar os alunos com as piores realizações da turma, ou seja aqueles que se suspeita poderem ter uma dificuldade de aprendizagem específica na leitura (Dislexia) (JENKINS; HUDSON; JOHNSON, 2007). Estes casos podem ser alunos que no futuro venham a ter um diagnóstico de dislexia e que, num modelo de resposta à intervenção, devem beneficiar de intervenções individualizadas, como as centradas na tutoria instrucional (ver o estudo de MACHADO;

CAPELLINI, 2014).

Assim, no contexto da triagem universal do Modelo de Resposta à Intervenção foi desenvolvido em Portugal 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. Emergem desta finalidade dois objetivos que se relacionam com o conhecimento do impacto de cinco procedimentos de cotação da prova MBC-Maze: a) na validade concorrente da prova; b) no poder discriminativo de verdadeiros positivos (sensibilidade) e de verdadeiros negativos (especificidade) da prova.

Método

População

Participaram na presente investigação 82 alunos do 3º ano do 1º Ciclo do Ensino Básico de um Agrupamento de Escolas do Norte de Portugal, com 8 anos de idade, sem necessidades educativas especiais identificadas, sem apoio dos serviços de educação especial e distribuídos por seis turmas. Destes, 47 (57,3%) são do género feminino e 35 (42,7%) do género masculino, que constituem a população do referido ano e ciclo de ensino do agrupamento.

Instrumentos de recolha de dados

Na presente investigação utilizaram-se dois instrumentos de recolha de dados: (1) uma prova de MBC para a compreensão da leitura – prova MBC-Maze (BROWN-CHIDSEY; DAVIS; MAYA, 2003; BUSCH; LEMBKE, 2005; SHINN; SHINN, 2002; STEVENSON, 2017); e (2) o Teste de Idade de Leitura (TIL), que avalia a compreensão da leitura (SUCENA; CASTRO, 2010), que se descrevem de seguida.

1) Prova de MBC para a compreensão da leitura – prova MBC Maze: A preparação desta prova teve início com um trabalho prévio de pesquisa de textos de manuais escolares do 3º ano de escolaridade, diferentes do adotado no agrupamento de escolas, com vista a seleccionar o que, pelas suas características, melhor se adequasse às regras de elaboração da mesma. Neste sentido, após selecionado o texto e havendo a indicação de que este não havia sido trabalhado em nenhuma das turmas, decidiu-se pela elaboração da prova com base no mesmo e acordou-se com os docentes que estes não o trabalhariam nas suas aulas. Teve-se ainda o cuidado de escolher um texto que não fosse alusivo a temas específicos, como por exemplo as estações do

ano ou quadras festivas.

Escolhido o texto iniciou-se a construção da prova, respeitando-se as regras que se seguem: (1) Deixar intacta a primeira frase; (2) Apagar, a partir da primeira frase, cada sétima palavra e no lugar dela colocar três palavras: a palavra original que pertence ao texto (palavra correta) e duas outras que não fazem sentido no texto (distratores); (3) Ter a certeza de que nenhum dos distratores faz sentido no texto. Apenas a palavra correta pode fazer sentido no texto. Relativamente aos distratores, tiveram-se em conta as seguintes regras: (a) Não escolher os distratores ao acaso. Um deles deve ser do mesmo tipo da palavra correta (nome, verbo, advérbio) e o outro uma palavra retirada aleatoriamente do texto; (b) Escolher distratores do mesmo comprimento que a palavra correta; (c) Selecionar distratores que não rimem com a palavra correta; (d) Escolher distratores que não comecem pela mesma letra da palavra correta (BUSCH; LEMBKE, 2005). Consideraram-se, ainda, algumas sugestões adicionais segundo Busch e Lembke (2005): (1) Sublinhar e colocar em negrito as três palavras (a correta e os distratores); (2) Não deixar separadas as três palavras no final das linhas do texto; (3) Passar para a palavra anterior ou para a palavra seguinte quando a sétima palavra for um nome próprio; (4) Variar a localização da palavra correta relativamente aos distratores; (5) Escrever as três palavras com letra maiúscula, quando a sétima palavra for a primeira palavra de uma frase.

A prova é constituída por uma folha de rosto que contém um exemplo constituído por três frases de treino. Segue-se o texto com 33 situações (resultado máximo) em que o aluno tem de seleccionar a palavra que faz sentido. Definiu-se o limite de tempo de três minutos e optou-se pela realização coletiva em cada turma. As provas foram cotadas recorrendo-se aos cinco procedimentos de cotação anteriormente descritos.

2) *Teste de Idade de Leitura (TIL)*: Este teste consiste numa adaptação do sub-teste Lobrot L3, que é uma prova estandardizada para o francês e de utilização generalizada por investigadores francófonos, nomeadamente na área da dislexia e da psicologia escolar. Trata-se de um teste que avalia a leitura e que pode ser usado para a realização de um primeiro diagnóstico de dificuldades de leitura/dislexia. O TIL foi elaborado para poder ser administrado de forma coletiva. Durante a sua realização, o aluno tem de ler silenciosamente frases isoladas, que não estão completas, e escolher uma das cinco palavras que se encontram dentro de parênteses para completar cada frase. Para tal devem sublinhar a palavra escolhida. Da prova fazem parte quatro frases para os alunos treinarem (presentes numa primeira página) e 36 frases experimentais (escritas na segunda página e distribuídas por duas colunas, de forma equitativa). O tempo destinado à realização do TIL é de 5 minutos e a cotação consiste na soma das frases

completadas corretamente, na multiplicação por 100 do número obtido e na divisão do produto resultante por 36 (número total de frases) (SUCENA; CASTRO, 2010).

Procedimentos de recolha de dados

Após a aprovação do projeto de investigação pelo Conselho Científico de uma Universidade Portuguesa e de obtidas as autorizações das instituições educativas procedeu-se à entrega aos encarregados de educação de uma declaração de autorização de participação na investigação por parte dos seus educandos. Todos os encarregados de educação autorizaram os seus educandos a participar na investigação. Adicionalmente, foi explicado aos alunos o que se pretendia deles e dada liberdade, nas diferentes monitorizações, para realizarem ou não a prova, tendo todos participado.

Com a prova MBC-Maze, ocorreu uma primeira monitorização no outono (início do ano letivo), uma segunda monitorização, aproximadamente duas semanas depois (um reteste para análise de fiabilidade), uma terceira monitorização no inverno e uma quarta monitorização na primavera (final do ano letivo). Adicionalmente, na primavera teve lugar uma avaliação com o TIL.

Para garantir a fiabilidade da administração de ambas as provas foi preenchido pelos professores o documento *Validação da implementação da prova Maze* de (PATRÃO, 2010) em todas as monitorizações com a prova Maze e o documento *Validação da implementação do TIL*³ nas avaliações com o TIL. As provas e o teste foram administrados e cotados pelo investigador principal deste estudo.

Procedimentos de análise de dados

Considerando os resultados obtidos nas diferentes monitorizações avaliaram-se os pressupostos de aplicabilidade, nomeadamente, normalidade, homogeneidade e independência dos resíduos obtidos, com recurso ao teste de Kolmogorov-Smirnov para a normalidade, teste de Levene para a homogeneidade das variâncias e para a independência o gráfico dos resíduos. A fiabilidade e a validade da prova foram avaliadas recorrendo ao coeficiente de correlação Pearson (r de Pearson). Fez-se uma análise ROC (*Receiver Operating Characteristic*), com metodologia DeLong et al. (1988), para obtenção de estatísticas de eficiência de diagnóstico da

³ Elaborado no contexto do presente estudo a partir do documento de PATRÃO (2010).

prova de MBC-Maze (sensibilidade, especificidade, valor preditivo positivo, valor preditivo negativo e percentagem de classificações corretas). Para o cálculo da sensibilidade e da especificidade, como no caso português não existe um “padrão ouro” que tenha condições de realmente afirmar se a condição existe ou não num determinado aluno, usou-se um critério resultante da aplicação do TIL (SUCENA; CASTRO, 2010), que foi considerado o teste de critério. Tendo por objetivo encontrar os alunos com uma realização de *leitura muito insatisfatória* utilizamos o percentil 10 no TIL como padrão ouro, ou seja, como variável que indicou quem tinha e não tinha uma *compreensão na leitura muito insatisfatória*. Este critério permitiu identificar os alunos com realizações piores da turma, aqueles que se suspeita poderem ter dislexia.

Resultados

Segue-se a apresentação dos resultados obtidos, quando utilizados os cinco procedimentos de cotação⁴, tendo por base os objetivos do estudo. Sublinha-se que estão verificados os pressupostos de aplicabilidade de testes paramétricos (PESTANA; GAGEIRO, 2008).

a) Impacto de cinco procedimentos de cotação na validade concorrente da prova

Com vista ao estudo da validade concorrente de cada um dos procedimentos de cotação, recorreu-se ao r de Pearson para calcular os coeficientes de correlação entre os resultados obtidos na prova CBM Maze (com a utilização de cada um dos cinco procedimentos de cotação) realizada na primavera e os resultados obtidos no TIL administrado também na primavera. Como se pode observar no Quadro 1, verifica-se que os coeficientes de correlação obtidos variam entre .776 e .806, valores indicativos de uma associação linear alta (PESTANA; GAGEIRO, 2008).

Quadro 1 - Correlações entre Maze e TIL por procedimento de cotação

⁴ Procedimento de cotação nº 1 (Número de respostas corretas menos as incorretas, RC-RI), procedimento de cotação nº 2 (Número de respostas corretas menos metade das incorretas, RC-½RI), procedimento de cotação nº 3 (Número de respostas corretas com interrupção da cotação após dois erros consecutivos, RC2E), procedimento de cotação nº 4 (Número de respostas corretas com interrupção da cotação após três erros consecutivos, RC3E), procedimento de cotação nº 5 (Número total de respostas corretas, RC).

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

Fonte: Elaborado pelas autoras.

O coeficiente de correlação mais elevado é obtido para o procedimento de cotação RC, e o mais baixo para o procedimento RC-RI. Significa isto que, teoricamente, poder-se-á ter mais confiança nos resultados obtidos com a utilização do procedimento de cotação RC. Contudo, os resultados obtidos com os diferentes procedimentos de cotação são maioritariamente indicativos de associações lineares altas, havendo apenas um resultado indicativo de associação linear moderada (PESTANA; GAGEIRO, 2008), e todos são significativos a 1%. Quer isto dizer que, na prática, os resultados obtidos com todos os procedimentos de cotação são fiáveis.

b) Impacto de cinco procedimentos de cotação no poder discriminativo de verdadeiros positivos (sensibilidade) e verdadeiros negativos (especificidade) da prova

Os resultados do *Teste ANOVA a 1 fator* mostram que há diferenças estatisticamente significativas entre procedimentos de cotação apenas na monitorização do outono (ET= 2.655, para $p=.033$). Nas subsequentes monitorizações, feitas ao longo do ano letivo, não há diferenças estatisticamente significativas entre os procedimentos de cotação. O teste *pos-hoc Tukey* evidenciou que os procedimentos de cotação que diferem estatisticamente entre si no outono, ou seja, no início do ano letivo, são o procedimento RC-RI e o procedimento RC.

As provas de MBC-Maze tiveram muito boa aceitação por parte dos professores, das famílias e dos alunos envolvidos, tendo todos autorizado e/ou participado. A prova mostrou-se simples e prática na aplicação em grupo. O trabalho de administração das provas demorou em cada turma cerca de 15 minutos, tempo que compreendeu as informações aos alunos sobre o que se iria fazer em seguida, a obtenção do seu consentimento verbal, a sua organização na sala, a leitura das respetivas instruções, a realização das provas e a recolha das mesmas. Este tempo foi diminuindo ao longo do ano letivo e à medida que os alunos se familiarizavam com os procedimentos. Independentemente do procedimento de cotação, cada prova pôde ser corrigida em menos de cinco minutos.

No Quadro 2 apresentam-se as cinco estatísticas de eficiência de diagnóstico, nomeadamente sensibilidade (S), especificidade (E), valor preditivo positivo (VPP), valor preditivo negativo (VPN) e percentagem de classificações corretas (PCC) para o ponto de corte

(PC) relativo ao percentil 20 do início e do final do ano letivo, que foi o considerado para a identificação de alunos em risco, tal como sugerido por (DENO, 1985).

Quadro 2 - Eficácia de diagnóstico da prova MBC-Maze

Procedimento	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

Fonte: Elaborado pelas autoras.

O procedimento de cotação que apresentou uma maior sensibilidade no início do ano foi o RC-½RI. A sensibilidade da prova MBC-Maze, cotada com o procedimento RC-½RI, para prever, no início do ano letivo, se um aluno estará em risco de vir a apresentar uma leitura muito insatisfatória no final do ano letivo, é de 96%. A especificidade da prova, quando cotada com este procedimento de cotação, ou seja, a possibilidade de excluir do grupo de risco os alunos que terão uma leitura satisfatória e, conseqüentemente, uma intervenção desnecessária, é de 38%. No início do ano letivo, este é um dos procedimentos mais eficazes (87% de classificações corretas) para ser utilizado na cotação do instrumento de triagem.

Considerações finais

A utilização de um sistema universal de triagem para a deteção precoce do risco de dislexia é um dos maiores e mais promissores desafios duma reforma da educação. A triagem determina o nível de risco para problemas de leitura em geral e o risco potencial de ter ou desenvolver dislexia. Os procedimentos de triagem para o risco de dislexia devem ser eficientes e baratos e devem ser usados para todos os alunos em sala de aula.

Este estudo testou a administração, a cotação e a adequação técnica de uma prova de MBC-Maze na identificação de alunos em risco de apresentarem dislexia e analisou o efeito de diferentes procedimentos de cotação desta prova. A partir dos resultados obtidos foi possível retirar as conclusões que se apresentam de seguida:

1. Os resultados obtidos na prova Maze são válidos independentemente do procedimento de cotação usado.

No que respeita à validade, estudou-se a validade concorrente. Os coeficientes de correlação da prova Maze com o TIL obtidos na primavera, quando utilizados os cinco procedimentos de cotação, remetem para o facto de os resultados na prova Maze serem válidos, uma vez que todos os referidos coeficientes (entre .776 e .806) são indicativos de associações lineares altas de acordo com Pestana e Gageiro (2008). Trata-se de coeficientes de correlação comparáveis aos obtidos por Tichá, Espin e Wayman (2009) (.82 entre a prova Maze e o MBST e .88 entre a prova Maze e o WJ-III) e por Hale *et al.*, (2011) (.859, entre os resultados em provas Maze realizadas em três minutos e os resultados no WJ-III ACH Broad Reading Cluster Score e .860 entre os resultados nas mesmas provas Maze e os resultados no WCPM), uma vez que correspondem também a associações lineares altas (PESTANA; GAGEIRO, 2008). São semelhantes aos encontrados por Pierce, McMaster e Deno (2010) (.73 e .82) quando ao comparar cada um dos cinco procedimentos de cotação, correlacionaram os respetivos resultados com três medidas de critério (o KTEA-II Letter Word Overall, o KTEA-II Reading Comprehension Overall e Provas de Leitura Oral), e globalmente superiores aos encontrados noutros estudos, nomeadamente por Jenkins e Jewell (1993), que correlacionaram os resultados obtidos em provas Maze (realizadas por alunos do 2º ao 6º anos, num minuto, e cotadas recorrendo ao procedimento de cotação RC3E com os resultados em testes como o Gates-MacGinitie Reading Tests (GM-TR) e o Metropolitan Achievement Tests (MAT), para estudarem a validade concorrente dos resultados obtidos nas provas Maze), e obtiveram correlações entre .65 e .76 com o primeiro e entre .66 e .76 com o segundo.

2. No início do ano letivo não é indiferente o procedimento de cotação que é utilizado.

No âmbito da comparação dos diferentes procedimentos de cotação, importa ainda destacar as diferenças estatisticamente significativas entre as médias obtidas quando utilizados os procedimentos de cotação RC-RI e RC, na monitorização do outono, início de um ano letivo e em que os alunos se encontram ainda num período “inicial” da aprendizagem da leitura (embora se trate de um 3º ano de escolaridade). Significa então que neste momento do ano letivo não é indiferente a opção por um destes dois procedimentos de cotação, devendo o professor ponderar a sua escolha, também em função de outras variáveis que interesse considerar. Já nas monitorizações que tiverem lugar no inverno e na primavera pode escolher-

se qualquer um dos procedimentos de cotação, uma vez que as diferenças nos resultados obtidos não são estatisticamente significativas.

3. A sensibilidade da prova MBC-Maze, para predizer no início do ano letivo se um aluno terá possibilidade de vir a apresentar uma leitura muito insatisfatória no final do ano letivo, variou entre 93% e 96%, consoante o procedimento de cotação.

Embora um sistema universal de triagem para deteção do risco de dislexia possa potencialmente ajudar muitos estudantes que se deparam com problemas na aprendizagem da leitura, ela também apresenta riscos e desafios para os sistemas escolares e para os alunos dentro deles.

A implementação de um sistema universal de triagem eficaz para detetar precocemente o risco de dislexia não é uma simples questão de escolher e administrar um teste único para seleccionar crianças. Pelo contrário, ela deverá assentar na utilização de provas de fácil implementação, breve realização e cotação, válidas e fiáveis, que possam ser implementadas pelos professores pelo menos três vezes no ano letivo, a todos os alunos de todos os anos de escolaridade no 1.º Ciclo do Ensino Básico. Um sistema eficaz de triagem para a dislexia acarreta riscos associados a dois tipos de resultados de triagem incorretos: aquele em que o examinador identifica o aluno como estando em risco, verificando-se posteriormente que o mesmo não apresenta resultados abaixo de um determinado valor de referência (erro de identificação de falso positivo); aquele em que o examinador diz que o aluno não está em risco, mas acaba por apresentar posteriormente resultados abaixo de um valor de referência especificado (erro de identificação de falso negativo). O problema da identificação de falsos positivos deve ser enquadrado no contexto dos recursos existentes na escola. Se no processo de triagem universal o examinador identificar muitos falsos positivos, ele pode estar a contribuir para que se desperdicem recursos ao estar a proporcionar intervenção e/ou avaliação adicional a alunos que, na realidade, deles não necessitam. Assim as escolas estarão a gastar preciosos recursos de intervenção com alunos que efetivamente não necessitam de intervenção adicional, o que pode fazer com que os recursos se tornem insuficientes para os alunos que realmente precisam deles. Portanto, a triagem deve procurar identificar corretamente 80% ou mais dos alunos que não estão em risco. No entanto, as consequências da identificação de falsos negativos, que pode significar falta de acesso a intervenções precoces na leitura a alunos que não conseguem ter sucesso sem intervenção, talvez sejam mais graves, devido à necessidade de os estudantes em risco de dislexia receberem intervenção na leitura o mais cedo possível, no período em que a aquisição da leitura é enfatizada.

Assim pode considerar-se que as provas de MBC-Maze revelam boas características técnicas, validade, sensibilidade e especificidade, que tornam segura a sua utilização por parte dos professores na monitorização do progresso dos alunos na leitura. Adicionalmente, no contexto de um Modelo de Resposta à Intervenção, além da triagem universal com recurso a provas de MBC, os investigadores internacionais sugerem que as escolas implementem um apoio multinível, no sentido de, progressivamente, providenciarem intervenções mais intensivas, que de forma direta e explícita trabalhem as necessidades académicas, comportamentais e sociais dos alunos (BRADLEY; DANIELSON; HALLAHAN, 2002).

AGRADECIMENTOS: Esta investigação foi desenvolvida no âmbito de um projeto financiado com uma bolsa da Fundação para a Ciência e a Tecnologia (Portugal).

REFERÊNCIAS

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.; CATTS, 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.

Como referenciar este artigo

VAZ, Paula Marisa Fortunato; MARTINS, Ana Paula Loução. Monitorização com base no currículo na escola inclusiva: adequação técnica das provas maze na triagem do risco 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>

Submetido em: 11/10/2019

Revisões requeridas: 20/11/2019

Aprovado em: 28/12/2019

Publicado em: 30/04/2020