

Desenvolvimento do Subteste de Leitura do Teste de Desempenho Escolar II

Marcia de Lima Athayde¹
Euclides José de Mendonça Filho¹
Rochele Paz Fonseca²
Lilian Milnitsky Stein²
Claudia Hofbeinzi Giacomoni¹

¹Universidade Federal do Rio Grande do Sul, Porto Alegre, RS

²Pontifícia Universidade Católica do Rio Grande do Sul, Porto Alegre, RS

Resumo

Este estudo teve como objetivo o desenvolvimento do Subteste de Leitura da segunda edição do Teste de Desempenho Escolar (TDE-II). Este estudo envolveu a seleção de itens, análise de juízes e a validação que contou com a participação de 661 estudantes de 1º a 9º ano do ensino fundamental. O subteste foi dividido em duas versões: a versão A para estudantes de 1º a 4º ano e a versão B para 5º a 9º ano. A análise fatorial exploratória indicou que ambas as versões são unidimensionais com um fator explicando 57% da versão A e um fator explicando 38% da variância compartilhada pelos itens da versão B. Análises via Teoria de Resposta ao Item indicou níveis de discriminação e dificuldade adequados para mensurar o construto. Os resultados revelaram evidências relacionadas ao conteúdo e à estrutura interna do teste. Em relação à fidedignidade, as versões do Subteste de Leitura cobrem de forma precisa as habilidades, sugerindo que o subteste apresenta boas propriedades psicométricas para avaliação da leitura.

Palavras-chaves: neuropsicologia cognitiva, leitura, psicometria, avaliação

Development of the Reading Subtest of the School Performance Test II

Abstract

This study aimed at the development of the Reading Subtest of the second edition of the School Performance Test (TDE-II). This study consisted in items selection, analysis by experts, and validation in which a sample of 661 students from the first to the ninth grades of Elementary School participated. The Subtest was divided in two versions: version A for students from the first to the fourth grades and version B for students from the fifth to ninth grades. Exploratory factor analysis indicated unidimensionality with one factor explaining 57% of the shared items variance for version A and 38% for version B. Item Response Theory analysis indicated appropriate levels of discrimination and difficulty for measuring the construct. Results suggest sources of validity in relation to test content, and internal structure. In terms of reliability, both versions covered the range of ability with high precision, thus suggesting that the Subtest presents good psychometric properties for reading evaluation.

Keywords: cognitive neuropsychology; reading; psychometric; evaluation

Desarrollo del Sub-test de Lectura del Test de Desempeño Escolar II

Resumen

Este estudio tuvo como objetivo desarrollar el Sub-test de Lectura de la segunda edición del Test de Desempeño escolar (TDE-II). El estudio consistió en la selección de ítems, análisis de jueces y una validación que contó con la participación de 661 alumnos de primer a noveno año de Enseñanza Primaria. El Sub-test se dividió en dos versiones: la versión A para estudiantes de primer a cuarto año y la versión B para alumnos de quinto a noveno año. El análisis factorial exploratorio indicó que las dos versiones son unidimensionales con un factor explicando el 57% de la versión A y 38% de la varianza compartida por los ítems de la versión B. Análisis vía Teoría de Respuesta al Ítem (TRI) indicaron niveles de discriminación y dificultad adecuados para medir el constructo. Los resultados mostraron evidencias relacionadas a contenido y a estructura interna del test. Con relación a la fidedignidad, las versiones del Sub-test de Lectura abarcan de manera precisa las habilidades, sugiriendo que el Sub-test presenta buenas propiedades psicométricas para evaluación de la lectura.

Palabras clave: neuropsicología cognitiva; lectura; psicometría; evaluación

Introdução

A leitura de palavras eficiente é crucial para uma adequada compreensão textual (Corso & Salles, 2009; Scliar & Cabral, 2013; Kim, Park, & Wagner, 2014), bem como para diversas atividades do cotidiano, dentre elas o

estudo de disciplinas escolares, como História, Ciências e Geografia (Ferreira & Leal, 2007). Dessa forma, tanto o ensino da habilidade de leitura é fundamental quanto sua avaliação em cada ano escolar (Ferreira & Leal, 2007).

A avaliação da leitura tem-se realizado por tarefas que envolvem:

- a) leitura de palavras, como a Avaliação de Leitura de Palavras Isoladas (Salles, 2002, 2005) e o Teste de Competência de Leitura Silenciosa de Palavras (Capovilla & Capovilla, 2001; Capovilla & Capovilla, 2004);
- b) palavras e pseudopalavras, como é o caso da Avaliação da leitura oral de palavras/pseudopalavras isoladas – LPI (normas de desempenho em Salles, Piccolo, & Zamo, Toazza, 2013) e do Teste de Competência de Leitura de Palavras e Pseudopalavras (TCLPP) (Seabra & Capovilla, 2010);
- c) compreensão textual, como a Avaliação da compreensão leitora: textos expositivos (Saraiva, Moojen, & Munarski, 2009) e a Avaliação da compreensão leitora por meio de reconto e questionário (normas de desempenho em Corso, Piccolo, Miná e Salles, 2015), dentre outras;
- d) avaliações mais globais, como as Provas de Avaliação dos Processos de Leitura – PROLEC (Cuetos, Rodrigues, & Ruano, 2012), as quais envolvem identificação de letras, leitura de palavras e pseudopalavras, compreensão textual, dentre outros processos de leitura; e a Provas de Habilidades Metalinguísticas e de Leitura (Prohmele) (Cunha & Capellini, 2009), a qual engloba habilidades metalinguísticas relacionadas com a aprendizagem da leitura.

Nesse cenário, tem-se o Subteste de Leitura do Teste de Desempenho Escolar (TDE) (Stein, 1994). O TDE possui o diferencial de ser o único teste de natureza de *screening* que apresenta normas para a população brasileira e evidências psicométricas de validade e precisão (Lúcio, Pinheiro, & Nascimento, 2009; Knijnik et al., 2014; Athayde et al., 2014). Realiza uma avaliação ampla da aprendizagem, fornecendo de forma objetiva uma avaliação das capacidades fundamentais para o desempenho escolar (leitura, escrita e aritmética).

O Subteste de Leitura do TDE é composto por 70 palavras, as quais são apresentadas sem contexto, e tem como objetivo avaliar de forma breve a capacidade de ler palavras, não avaliando os demais aspectos envolvidos na leitura, tal como a compreensão leitora.

Identificou-se a necessidade de atualizar o Subteste de Leitura, em virtude de: a) desatualização das normas – tendo em vista que o TDE original foi criado e normatizado há 21 anos, ele encontra-se defasado. Ademais, a recomendação do meio científico é de que os testes estejam atualizados em relação às normas, à precisão e à validade (resolução 02/2003 do Conselho

Federal de Psicologia; *International Test Commission*, 2000; *American Psychological Association* (APA) (Adams, 2000); b) mudança no currículo escolar do ensino fundamental (modificação prevista na Lei nº 10.172, de 9 de janeiro de 2001) – a redistribuição dos conteúdos em nove anos escolares fez com que o TDE ficasse desatualizado no que se refere aos dados normativos para cada ano escolar; c) lacunas quanto a características psicométricas apropriadas, uma vez que o subteste original carecia de itens que avaliassem adequadamente sujeitos com diferentes níveis de habilidade, bem como não apresentava os itens seguindo uma ordem crescente de dificuldade (ver estudos de Athayde et al., 2014 e Lúcio, Pinheiro, & Nascimento, 2009); d) carência de instrumentos que avaliem de forma breve e ampla o desempenho na habilidade de leitura – o Subteste de Leitura original do TDE (Stein, 1994) era o único instrumento com esse objetivo (triagem psicopedagógica), validado e normatizado para a população brasileira; e) sua relevância em pesquisas realizadas em todo o país e em diferentes áreas da ciência (ver o estudo de Knijnik et al., 2013).

Em vista dessa necessidade de atualização, o presente estudo tem por objetivo apresentar o processo de desenvolvimento do Subteste de Leitura do TDE-II, (Stein, L.M., Giacomoni, C., Fonseca, R. P., instrumento em preparação, para atualização dos três subtestes que o compõe – escrita, leitura e aritmética), bem como verificar evidências relacionadas ao conteúdo e à estrutura interna do teste, além de aferir parâmetros de fidedignidade. Esse processo tomou como base a versão original do Subteste de Leitura (Stein, 1994), bem como critérios teóricos e empíricos definidos por *experts*. Além disso, seguiu procedimentos rigorosos de construção de testes referentes ao construto, uma vez que, segundo Pasquali (2010), um instrumento psicológico construído com referência a construto parte da teoria psicológica, e não de qualquer dado empírico, sendo esse justamente o procedimento realizado na construção do Subteste de Leitura do TDE II, o qual será visto a seguir.

Método

O método usado para a criação do Subteste de Leitura foi baseado nos procedimentos descritos por Pasquali (2010) para construção de testes referentes a construtos. Nessa sessão, serão descritos os procedimentos realizados no âmbito dos polos teórico, empírico e analítico, sendo que a última etapa do polo analítico, referente à normatização, será apresentada em um estudo futuro.

Polo Teórico - Etapa 1: Seleção dos Itens do Subteste de Leitura do TDE-II

Nessa etapa, alguns itens foram selecionados a partir do Subteste de Leitura original (Stein, 1994), outros foram indicados exclusivamente para a versão preliminar do subteste. Para nortear esse processo de desenvolvimento dos itens, foram estabelecidos critérios empíricos e teóricos.

Em relação aos critérios empíricos, a construção dos itens levou em consideração os achados experimentais dos estudos de Lúcio, Pinheiro e Nascimento (2009), Lúcio e Pinheiro (2014) e Athayde et al. (2014). Segundo essas pesquisas, constatou-se que o Subteste de Leitura original (Stein, 1994) carecia de itens extremamente fáceis e também de itens difíceis, a fim de avaliar apropriadamente estudantes de todos os níveis de habilidade. Portanto, o critério empírico estabelecido foi a necessidade de selecionar itens de nível de dificuldade muito fácil, bem como difíceis.

No que tange aos critérios teóricos, uma junta de juízes, peritos em linguagem escrita, trabalhou na seleção deles. Esta foi composta por dois fonoaudiólogos, sendo um com mestrado em Distúrbios da Comunicação Humana (juiz 1) e outro com graduação também em Psicologia, especialista em linguagem, mestre e doutor em Psicologia do Desenvolvimento e perito em Neuropsicologia (juiz 2). Os juízes peritos consultaram a literatura pertinente para o tema em questão e examinaram outros instrumentos que avaliavam o mesmo construto (ver instrumentos citados na introdução do artigo). Além disso, a escolha dos critérios teóricos levou em consideração o que se pretendia para a versão do Subteste de Leitura do TDE-II: itens que abarcassem a aprendizagem de todo o ensino fundamental de nove anos e que estivessem distribuídos obedecendo a uma ordem crescente de complexidade, de forma que a dificuldade de cada item aumentasse com o avanço da escolaridade.

Dessa forma, os critérios teóricos elegidos para embasar a seleção dos itens foram: 1) o nível de aprendizagem de cada ano escolar; 2) itens com conteúdo não regionalizado; e 3) a seleção de termos não pejorativos. Além desses, alguns critérios psicolinguísticos foram selecionados para nortear a seleção dos itens: grau de concretude e abstração, classificação gramatical, extensão lexical, estrutura silábica e grau de familiaridade das palavras. Para verificar o processo de seleção de itens com base nas variáveis psicolinguísticas, sugere-se consultar o estudo (Informação Retirada pela Revista) (manuscrito em preparação).

Em relação ao nível de aprendizagem de cada ano escolar, a identificação do processo de desenvolvimento da leitura ao longo do ensino fundamental serviu de base para a seleção dos itens. Os itens foram elaborados de acordo com o avanço da aprendizagem escolar, seguindo uma ordem crescente de dificuldade. Em relação à regionalização dos itens, buscou-se eleger itens que fossem familiares para todas as regiões do Brasil. Quanto à utilização de termos não pejorativos, buscou-se selecionar itens que não tivessem uma imagem considerada social ou emocionalmente negativa. A título de exemplo, os itens “durex” e “isqueiro” do Subteste de Leitura original foram excluídos, o primeiro por ser um termo regionalizado e o segundo por ser pejorativo.

Após a definição dos critérios teóricos e empíricos, os juízes 1 e 2 construíram um *pool* de cerca de 100 itens. Esse número grande de itens possibilitaria excluir aqueles que não apresentassem bom comportamento empírico, quando da fase da testagem empírica. Assim, foram selecionados 38 itens da versão original do subteste, sendo que esses foram escolhidos por terem apresentado bons índices de dificuldade e discriminação no estudo de Athayde et al. (2014), bem como atendiam às características psicolinguísticas elegidas. Para a seleção dos novos itens, os juízes usaram como fonte de inspiração: livros de temas diversos e *internet* (112 itens), um banco de dados de fluência verbal livre, oriundo de pesquisas realizadas no Grupo de Pesquisa (Informação Retirada pela Revista) (30 itens) e o glossário do MEC (seis itens).

Em seguida, os juízes 1 e 2 examinaram esse conjunto de 186 itens, buscando eleger os 100 itens mais adequados para comporem a versão preliminar do Subteste de Leitura. Para tanto, os juízes peritos levaram em consideração os critérios teóricos e empíricos mencionados anteriormente. Além disso, foi considerada a originalidade dos itens, buscando-se eliminar itens já encontrados em outros instrumentos de avaliação psicológica, fonoaudiológica ou psicopedagógica: WISC IV (Rueda, Noronha, Sisto, Santos, & Castro, 2012), NEUPSILIN (Fonseca, Salles, & Parente, 2009), PROHMELE (Cunha & Capellini, 2009) e PROLEC (Capellini, Oliveira, & Cuetos, 2010).

Polo Teórico - Etapa 2: Análise Teórica dos Itens do Subteste de Leitura do TDE-II

Com o intuito de verificar evidências relacionadas ao conteúdo do teste, os 100 itens selecionados

na Etapa 1 do presente estudo foram analisados por um juiz perito, com formação em Pedagogia, Fonoaudiologia e Psicopedagogia, além de mestrado em Educação – Psicologia Escolar (juiz 3). Por apresentar vasta experiência em Psicopedagogia Clínica, esse juiz perito foi eleito para averiguar a pertinência de cada item para a avaliação do construto em questão – habilidade de leitura de palavras. Além disso, o juiz 3 deveria estabelecer a ordem de apresentação dos itens, para que seguissem uma ordem crescente de complexidade, alcançando, assim, o critério teórico referente ao nível de aprendizagem de cada ano escolar.

Essa análise teórica resultou na seguinte classificação dos itens: 18 itens de 1º ano, 23 itens de 2º ano, 21 itens de 3º ano, 21 itens de 4º ano e 10 itens de 5º ano. Não foram classificados itens como pertencentes aos anos mais avançados, tendo em vista que a habilidade de ler palavras é uma atividade simples para os últimos anos escolares do ensino fundamental. Portanto, presumiu-se que não havia palavras possíveis de serem lidas apenas em anos escolares mais avançados desse nível. Além desses achados, essa última análise de juízes sugeriu a exclusão de sete itens, em virtude de não serem adequados para a avaliação da habilidade em questão. Sendo assim, a versão preliminar do Subteste de Leitura foi composta por 93 itens.

Polo Experimental – Etapa 1: Testagem Empírica da Versão Preliminar do Subteste de Leitura do TDE-II

Com o intuito de verificar evidências de validade relacionada à estrutura interna da versão preliminar do

Subteste de Leitura, foi realizada uma testagem empírica dos itens. O objetivo dessa etapa foi testar e analisar o comportamento empírico dos itens, a fim de verificar os apropriados para a avaliação do construto em questão.

Participantes

Participaram 706 estudantes dos nove anos escolares do ensino fundamental, de escolas públicas e privadas da região metropolitana de Porto Alegre. Foram incluídos na amostra todos os escolares com percentil maior que 5 no Teste Matrizes Progressivas Coloridas de Raven: Escala Especial (Angelini, Alves, Custódio, Duarte, & Duarte, 1999) e no Teste Matrizes Progressivas de Raven: Escala Geral (Vilhena, Guntert, & Tosi, no prelo). Além disso, foram excluídos sujeitos com suspeita de déficit intelectual, auditivo e alterações de fala e linguagem. Dessa forma, foram excluídas $n = 28$ crianças por déficit intelectual (identificada na avaliação do Raven), $n = 10$ devido a alterações de fala ou linguagem, $n = 3$ por serem alunos especiais da escola, $n = 2$ por apresentarem diagnóstico de Autismo e Síndrome de Asperger, $n = 1$ por suspeita de comprometimento intelectual (identificada pelo questionário sociodemográfico e avaliações do TDE), e $n = 1$ por suspeita de deficiência auditiva. Com isso, a amostra foi constituída por $n = 661$ estudantes, com idades entre 6,34 e 17,75 anos ($M = 10,17$; $DP = 2,43$). A Tabela 1 mostra a distribuição dos estudantes por ano escolar, sexo e tipo de escola.

Tabela 1.

Caracterização Sociodemográficas da Amostra

Ano Escolar	Escolas Públicas		Escolas Privadas		Faixa etária	
	Meninas	Meninos	Meninas	Meninos	Média	DP
1º	23	18	18	18	6,91	0,32
2º	24	21	35	31	7,85	0,32
3º	37	19	20	17	8,92	1,21
4º	25	33	19	16	9,88	0,55
5º	33	24	33	26	10,72	1,45
6º	27	22	15	8	11,86	2,12
7º	11	12	10	12	13,18	0,72
8º	13	12	10	7	14,30	0,95
9º	0	0	4	8	14,65	0,50
Total	193	161	164	143	10,17	2,43

Instrumentos

Os escolares foram avaliados por meio da versão preliminar do Subteste de Leitura e um teste de inteligência: o Teste Matrizes Progressivas Coloridas de Raven: Escala Especial (Angelini, Alves, Custódio, Duarte, & Duarte, 1999) foi aplicado em estudantes de até 11 anos e nove meses de idade e o Teste Matrizes Progressivas de Raven: Escala Geral (Vilhena, Guntert, & Tosi, no prelo) em escolares com mais de 11 anos e 10 meses. Além disso, os responsáveis pelos estudantes responderam a um questionário de dados sociodemográficos, conforme versão de Fonseca, Jacobsen e Pureza (2015).

Procedimentos de Coletas de Dados

O projeto dessa pesquisa foi aceito pelo Comitê de Ética da (Informação Retirada pela Revista) (CAAE: 06624312.7.0000.5336). Os estudantes foram autorizados pelos responsáveis a participar do estudo, mediante assinatura do Termo de Consentimento Livre e Esclarecido (TCLE).

Em um primeiro encontro foi realizada a avaliação intelectual, sendo individual com as crianças do 1º ao 3º ano do ensino fundamental, e coletiva, em grupos de oito participantes, nos demais anos escolares. No segundo encontro, foi realizada a avaliação da leitura de palavras isoladas, por meio da versão preliminar do Subteste de Leitura. Crianças do 1º ao 3º ano escolar eram expostas a metade dos itens no primeiro encontro e, no segundo encontro, à outra metade dos itens, a fim de evitar o efeito do cansaço. Para todas as crianças do ensino fundamental, o teste era interrompido após o 10º erro consecutivo. A avaliação era realizada individualmente, em sala silenciosa da própria escola, em horário de aula. A leitura feita pela criança da ficha de estímulos era gravada e suas respostas classificadas como “acerto” ou “erro” pelo examinador. Era considerado “erro” quando o participante não conseguia ler o item, ou o lia de forma incorreta. Em seguida, as gravações da avaliação de leitura eram aferidas por uma equipe treinada para correção dessas. Esses examinadores deveriam ouvir as gravações com fones de ouvido, conferindo as respostas assinaladas pelo avaliador e corrigindo-as, se necessário.

Procedimentos de Análises de Dados

Os procedimentos de análise dos dados contaram com três etapas. Inicialmente, avaliou-se a dimensionalidade do subteste, por meio de análises fatoriais de eixos principais, baseada na Teoria Clássica dos Testes (TCT). Em seguida, foi analisada a consistência interna do instrumento, por meio do coeficiente alfa, o qual

foi utilizado para verificar a precisão do instrumento (Zanon & Hauck Filho, 2015). Posteriormente, a Teoria de Resposta ao Item (TRI) foi utilizada para estimar o nível de habilidade dos sujeitos e os parâmetros de discriminação e dificuldade de cada item. Técnicas paramétricas de TRI foram utilizadas para analisar os dados, em que os pressupostos de monotonicidade (princípio em que a probabilidade de acerto de um determinado item aumenta de acordo com o aumento da habilidade do participante), independência local (os itens do teste não dependem um do outro ou não influenciam um ao outro para obtenção da resposta) e unidimensionalidade (princípio que o conjunto de itens diz respeito a um único traço latente) foram alcançados.

O Subteste de Leitura foi dividido em duas versões, de acordo com os pressupostos teóricos referentes à distribuição dos itens por anos escolares do ensino fundamental. Dessa forma, julgou-se oportuno direcionar os itens referentes ao 1º e ao 2º ano escolar para a versão A do subteste e os itens pertencentes ao 3º, 4º e 5º ano para a versão B. Com essa divisão dos itens, foi possível constatar a unidimensionalidade das versões A e B, o que é condição necessária para o uso da TRI. Portanto, as análises subsequentes tiveram como objetivo identificar as características psicométricas de cada uma das versões, o que será descrito a seguir.

As estimativas de habilidade dos participantes das duas versões foram computadas usando a função *factor.scores()* do pacote *psych*. Análises de variância do desempenho escolar de acordo com o ano escolar foram realizadas por meio dos testes paramétricos e não paramétricos (Kruskal Wallis), sendo que as análises *post hoc* utilizaram a correção de Bonferroni para o nível de significância.

Para realização das análises utilizaram-se os *softwares Statistical Package for the Social Sciences (SPSS)* e o *Software R (R Core Team, 2014)*, com as funções dos pacotes *psych* (Revelle, 2015), *ltm* (Rizopoulos, 2006) e *mokken* (der Ark, 2007).

Resultados

Polo Analítico – Versão A do Subteste de Leitura do TDE-II: Dimensionalidade, Análise Fatorial, TRI

A versão A do Subteste de Leitura contou inicialmente com 42 itens, cuja avaliação da dimensionalidade foi realizada por meio da análise fatorial exploratória da matriz de correlações policóricas dos itens, utilizando o método de extração de eixos principais (*principal axis factoring*). O teste KMO foi de 0,71 e o teste de esfericidade

de Bartlett foi significativo, $\chi^2(630) = 13321.97$, $p < 0.0001$, permitindo confirmar a adequação da análise fatorial para o tratamento dos dados. Para avaliar a pertinência do item em relação à dimensão avaliada, foram considerados itens que: 1º) possuíam o valor absoluto da carga fatorial principal do item acima de 0,30 para amostras acima de 350 participantes (significância baseada em $p < 0.05$ e poder estatístico de 80%); 2º) apresentaram cargas fatoriais similares em dois ou mais fatores em um mesmo item, sendo a diferença entre valores absolutos das cargas fatoriais dos itens maior do que 0,10; e 3º) tinham uma relação inequívoca com sua dimensão, ou seja, demonstravam convergência teórica com os demais itens do mesmo fator (Hair, Black, Babin, & Anderson, 2010; Laros, 2005; Smith & McCarthy, 1995).

A análise fatorial apontou um fator explicando 57% da variância dos dados e as cargas fatoriais dos itens variaram de 0,38 a 0,94. Vale ressaltar que a análise paralela de Horn também apontou um fator principal, bem como um componente principal (Horn, 1965). Nessa etapa, nenhum item foi eliminado, uma vez que todas as cargas fatoriais obtiveram valores acima de 0,30 e o alfa obtido foi 0,86.

Para realização da análise por meio da TRI, a possibilidade de escalabilidade do teste foi testada por meio do procedimento proposto por Mokken (der Ark, 2007; Rusch, Mair, & Hatzinger, 2013; van der Linden & Hambleton, 1996) e por meio da monotonicidade dos itens, em que os valores acima de 0,30 são considerados satisfatórios para a utilização de modelos paramétricos da TRI (Junker & Sijtsma, 2000). Dessa forma, o conjunto de itens da versão A obteve o valor diagnóstico de 0,36, sendo considerado aceitável.

Quatro modelos de TRI paramétricos foram testados no tratamento dos dados: Rash, modelo de um parâmetro (1PL), dois parâmetros (2PL) e três parâmetros (3PL). Esses quatro modelos foram avaliados por meio de teste de hipótese estatística na comparação de valores de verossimilhança, do critério de informação de Akaike (AIC) e do critério de informação Bayesiano (BIC). O modelo de 2PL (verossimilhança = -5670.72, AIC = 11509.44, BIC = 11886.91) foi significativamente melhor que o modelo Rasch (verossimilhança = -6445.19, AIC = 12974.37, BIC = 13163.11) e que o modelo 1PL (verossimilhança = -5766.01, AIC = 7078.68, BIC = 7286.4). Ao se comparar os modelos 2PL e 3PL (verossimilhança = -5683.18, AIC = 11618.37, BIC = 12184.58), optou-se pelo primeiro, em virtude de ter possuído menores valores de AIC e BIC

e maior verossimilhança (Beaujean, 2014; Field, Miles, & Field, 2012). Em contrapartida, o modelo 3PL apresentou uma solução instável, já que a matrix Hessiana convergiu em valores infinitos produzindo *missingvalues*.

Posteriormente, analisou-se a adequação dos itens ao modelo 2PL, uma vez que a TRI requer uma investigação formal sobre o funcionamento dos itens enquanto medidas de traço latente (Reise, 1990). A função *item.fit* (Rizopoulos, 2006) foi usada para comparar os valores previstos de cada item do modelo 2PL com os valores empíricos coletados por meio do teste quii-quadrado. A partir dessa análise, seis itens não se adequaram ao modelo 2PL, os quais foram excluídos. Salienta-se que a exclusão desses itens não resultou em perda de conteúdo teórico necessário para a avaliação da habilidade, uma vez que na etapa de desenvolvimento do subteste foram levantados um número maior de itens do que o necessário. Com isso, a versão A final do Subteste de Leitura resultou em 35 itens, cujas dificuldades variaram de -1,91 a -0,02, como pode ser observado na Tabela 2.

O modelo 2PL final obteve valores de verossimilhança igual a -5173.576, AIC igual a 10491.15 e BIC de 10814.7. Os escores dos alunos de 1º a 4º ano variaram de -1.82 a 0.98, sendo a média -0,3 ($DP = 0,76$). Nesse intervalo de habilidade, a versão A do Subteste de Leitura possui 93,99% de informação sob a curva de informação do teste (Figura 1).

Foram encontradas diferenças significativas nas estimativas de habilidade de leitura de acordo com o ano escolar, $H(3) = 94,34$, $p < 0,001$. Análises *post hoc* utilizando correção para múltiplas comparações no teste de Wilcoxon mostraram médias significativamente diferentes entre alunos do 1º ano ($M = -1,08$; $DP = 0,68$) e alunos do 2º ano ($M = -0,34$; $DP = 0,78$) com tamanho de efeito d de Cohen igual a 0,99, porém não foram encontradas diferenças significativas entre alunos do 2º ano com alunos do 3º ano ($M = -0,09$, $SD = 0,85$, $d = 0,30$), nem de alunos do 3º ano com alunos de 4º ano ($M = 0,20$, $DP = 0,52$, $d = 0,40$). Todas as comparações com dois anos de diferenças (ex.: alunos de 1º ano comparados com alunos de 3º ano) foram significativas.

Polo Analítico – Versão B do Subteste de Leitura do TDE-II: Dimensionalidade, Análise Fatorial, TRI

A versão B do Subteste de Leitura foi composta inicialmente por 51 itens e a análise fatorial apontou uma dimensão explicando 38% de variância dos dados ($KMO = 0,9$, teste de esfericidade de Bartlett $\chi^2(1275) = 7689.24$, $p < 0,0001$, alfa = 0,89). A análise paralela

Tabela 2.
Dificuldade e Discriminação dos Itens da Versão A do TDE Leitura

Item	Dificuldade	Discriminação	Item	Dificuldade	Discriminação
L36	-1,91	1,28	L33	-0,83	3,25
L40	-1,46	2,42	L39	-0,83	2,03
L17	-1,21	2,68	L31	-0,82	3,75
L6	-1,12	4,19	L23	-0,82	2,69
L38	-1,07	3,42	L18	-0,77	3,08
L12	-1,07	2,72	L24	-0,76	3,21
L29	-1,06	3,80	L21	-0,72	3,17
L7	-1,04	3,02	L3	-0,67	3,43
L35	-1,04	1,69	L42	-0,65	2,66
L16	-1,04	3,90	L30	-0,64	3,16
L10	-1,02	3,68	L25	-0,57	3,57
L19	-1,00	2,85	L32	-0,53	2,26
L9	-0,98	4,34	L13	-0,45	2,80
L14	-0,97	4,02	L26	-0,41	4,58
L22	-0,94	2,90	L20	-0,33	2,86
L37	-0,92	3,02	L41	-0,21	2,42
L34	-0,92	2,31	L28	-0,02	1,86
L27	-0,89	3,08			

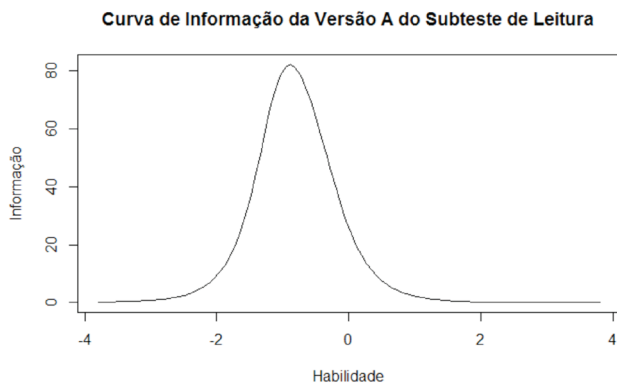


Figura 1. Curva de informação da versão A do subteste de Leitura.

de Horn também apontou um fator principal, bem como um componente principal (Horn, 1965). Nessa etapa, nove itens foram eliminados por conterem baixas cargas fatoriais (abaixo de 0,3). Ao final da análise, os valores das cargas fatoriais variaram de 0,40 a 0,83. A possibilidade de análise dos dados por meio da TRI foi corroborada por meio da obtenção de valores

aceitáveis de escalabilidade e monotonicidade (der Ark, 2007; Rusch, Mair, & Hatzinger, 2013; van der Linden & Hambleton, 1996), todos acima de 0,3.

Seguindo os mesmos procedimentos descritos para a comparação de modelos de TRI da versão A, na versão B o modelo 2PL (verossimilhança = -10004.77, AIC = 20177.54, BIC = 20555.02), mostrou-se significativamente mais satisfatório, quando comparado com o modelo Rasch (verossimilhança = -10205.52, AIC = 20495.05, e BIC = 20683.79) e o modelo 1PL (verossimilhança = -10157.75, AIC = 20401.50, BIC = 20594.73). Mesmo não encontrando diferenças significativas entre o modelo 2PL e o 3PL (verossimilhança = -9991.98, AIC = 20235.96, e BIC = 20802.17), optou-se pelo modelo 2PL, seguindo o princípio de parcimônia em ciência, em que se privilegia a explicação e o modelo mais simples, com menor número de parâmetros (Pasquali & Primi, 2003).

A investigação de adequação do conjunto de itens ao modelo 2PL por meio da função *item.fit* (Rizopoulos, 2006) apontou índices apropriados para todo o conjunto de itens. Nessa etapa, a versão B do Subteste de Leitura contava com 42 itens. Porém, identificou-se

a necessidade de diminuir a quantidade de itens, levando-se em consideração que o TDE se destina a ser um instrumento de rastreamento. Para tanto, realizou-se nova revisão dos itens, a fim de eliminar aqueles que se apresentavam menos relevantes para a avaliação do construto. A eliminação desses itens foi realizada seguindo critérios empíricos e teóricos. Primeiramente, identificaram-se os itens que estavam mensurando o mesmo nível de habilidade (itens com níveis de dificuldade muito próximos). Após, esses itens redundantes passaram por nova análise teórica, a qual foi realizada pelos juízes 1 e 2. Os juízes deveriam analisar os pares de itens com dificuldades semelhantes, optando pela permanência de um em detrimento do outro. Para isso, levaram em consideração as características teóricas dos itens, bem como os valores de discriminação.

Como resultado dessa última análise, nove itens foram excluídos. Com isso, o modelo 2PL final contou com 33 itens, cujas dificuldades variaram de -2,503 a 1,332 (Tabela 3). O valor de verossimilhança ficou igual a -7421.424, AIC igual a 14966.85 e BIC igual a 15245.46.

Os escores dos alunos de 5º a 9º ano variaram de -1,74 a 1,97, sendo a média igual a 0,44 ($DP = 0,74$).

Nesse intervalo de habilidade, a versão B do Subteste de Leitura possui 64,5% de informação sob a curva de informação do teste (Figura 2).

De forma semelhante à versão A, as estimativas de habilidade da versão B do subteste de leitura também teve relacionamento significativo com o ano escolar dos participantes, $F(4,282) = 9,65$, $p < 0,001$. As análises *post hoc* mostraram diferenças significativas entre alunos

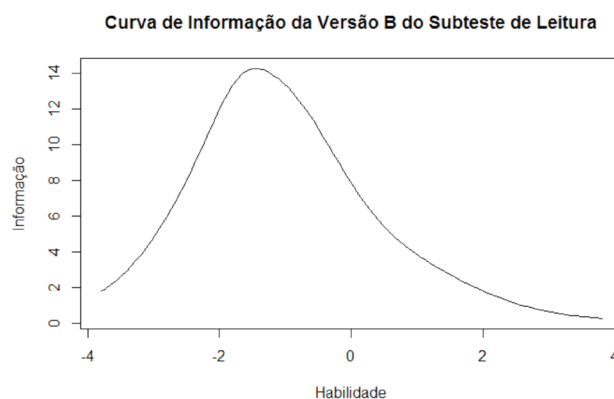


Figura 2. Curva de informação da versão B do subteste de Leitura.

Tabela 3.

Dificuldade e Discriminação dos Itens da Versão B do Subteste de Leitura

Item	Dificuldade	Discriminação	Item	Dificuldade	Discriminação
L64	-2,50	1,67	L91	-0,91	0,83
L54	-2,32	1,80	L78	-0,90	1,00
L80	-2,18	1,93	L63	-0,70	0,95
L53	-2,08	1,47	L58	-0,64	1,13
L44	-1,98	1,46	L57	-0,59	1,75
L90	-1,80	1,37	L81	-0,56	1,79
L70	-1,70	3,10	L60	-0,44	2,11
L83	-1,58	1,13	L52	-0,40	0,90
L65	-1,57	2,23	L84	-0,36	1,11
L67	-1,50	1,40	L86	-0,33	1,03
L50	-1,30	1,16	L92	-0,18	1,58
L76	-1,29	1,13	L93	0,08	1,21
L72	-1,29	1,28	L51	0,19	0,98
L47	-1,12	2,19	L85	0,56	1,38
L77	-1,04	1,19	L69	1,30	1,74
L66	-0,95	1,60	L62	1,33	0,96
L43	-0,92	2,43			

do 5º ano ($M = 0,17$, $DP = 0,74$) com alunos do 6º ano ($M = 0,49$, $DP = 0,66$) $d = 0,30$, 7º ($M = 0,66$, $DP = 0,68$) $d = 0,77$; 8º ($M = 0,72$, $DP = 0,68$) $d = 0,85$ e 9º ano ($M = 1,05$, $DP = 0,62$) $d = 1,17$.

Discussão

O presente estudo descreveu o processo de desenvolvimento do Subteste de Leitura do TDE-II. Essa construção se deu seguindo criteriosamente os passos recomendados por Pasquali (2010), para a criação de instrumentos psicológicos referentes a construtos. Além disso, o Subteste apresentou evidências relacionadas ao conteúdo e à estrutura interna do teste, bem como altos parâmetros de fidedignidade.

Em relação ao rigor conferido no desenvolvimento do subteste, Lúcio, Moura, Nascimento e Pinheiro (2012) referem que são raros os instrumentos de avaliação da leitura cuja construção seguiu as etapas propostas pela psicometria. Segundo os autores, o que mais se verifica são trabalhos publicados em revistas e livros, cuja avaliação se dá a partir de listas experimentais de palavras. Os autores ressaltam ainda que a construção de testes de avaliação da leitura deve ser construída tanto com base em critérios psicolinguísticos, quando com dados empíricos, o que vai justamente ao encontro do presente estudo.

Quanto à natureza do instrumento, pode-se considerar que o Subteste de Leitura do TDE-II mantém o caráter de triagem, sendo esse um dos objetivos iniciais para a construção do instrumento. Segundo Fonseca et al. (2012), um teste é considerado um *screening* ou triagem quando o tempo de aplicação varia de cinco a 15 minutos. Pasquali (2010) refere que quando um instrumento tem como objetivo uma investigação mais rápida, procurando-se situar o avaliando em uma dada categoria, este pode ser considerado uma triagem.

Portanto, o subteste permite identificar se o estudante apresenta defasagem na habilidade em questão e, segundo Stein (1994), a partir dessa constatação, o examinador poderá aplicar uma avaliação mais pormenorizada de leitura, buscando compreender a dificuldade da criança. Em relação a esse ponto, Pasquali (2010) refere que, normalmente, após a avaliação por meio de um instrumento de triagem, segue-se um estudo mais aprofundado dos sujeitos que foram alotados em uma categoria deficitária.

Apesar de avaliar a habilidade de ler palavras de modo breve, o Subteste de Leitura do TDE-II apresenta-se mais aprimorado do que sua versão original,

uma vez que traz itens mais adequados para a avaliação da habilidade de leitura. Estudos anteriores (Lúcio, Pinheiro, & Nascimento, 2009; Athayde et al., 2014; Lúcio & Pinheiro, 2014) corroboram essa constatação, já que verificaram que o Subteste de Leitura original não estava discriminando adequadamente os sujeitos ao longo do ensino fundamental. Já a versão atualizada do subteste demonstrou avaliar apropriadamente estudantes ao longo dos nove anos escolares. Isso pôde ser observado a partir das evidências relacionadas ao conteúdo e à estrutura interna do teste e também pela curva de informação das versões A e B do subteste, as quais demonstram estar avaliando apropriadamente sujeitos de todo o *continuum* de habilidade.

No que tange às etapas do processo de desenvolvimento do subteste, o mesmo cumpriu com rigor a exigências da APA e da *International Test Commission*. Pasquali (2010) refere que algumas condições são necessárias para que um teste apresente evidências de validade de conteúdo: definição do conteúdo, explicitação dos processos psicológicos a serem avaliados e determinação da proporção relativa de representação no teste de cada tópico de conteúdo. É possível observar, na etapa 1 do polo teórico, que esse procedimento foi cuidadosamente realizado, uma vez que foram definidos critérios teóricos e empíricos que nortearam a construção dos itens.

Ainda em relação à etapa 1 do polo teórico, observa-se que o presente estudo buscou alcançar o “critério de amplitude” proposto por Pasquali (2010). Segundo esse critério, o instrumento deve conseguir discriminar sujeitos de diferentes níveis de magnitude do traço latente. Esse critério foi atendido, na medida em que se selecionaram itens de diferentes níveis de complexidade, a partir da manipulação dos critérios teóricos elegidos (ver o estudo de Athayde et al., manuscrito em preparação).

Em relação ao nível de dificuldade dos itens, na segunda etapa do polo teórico, os itens passaram por análise teórica buscando atingir o “critério de relevância” mencionado por Pasquali (2010). Segundo o autor, os itens não devem insinuar atributo diferente do pretendido. Para Pasquali (2010), essa análise teórica dos itens, realizada por juízes especialistas, confere validade de conteúdo ao teste.

Por fim, ressalta-se que esse estudo teve o diferencial de ter sido desenvolvido fazendo-se uso da Teoria Clássica dos Testes (ICT) e de modelo moderno – Teoria de Resposta ao Item (TRI). A análise fatorial permitiu a constatação do critério de relevância o que,

segundo Pasquali (2010), somente pode ser obtido pela análise fatorial. Por meio da análise fatorial foi possível assegurar que os itens eram excelentes representantes do traço latente. Ademais, a unidimensionalidade encontrada em cada versão do instrumento indica que os itens estão medindo um mesmo construto, o que também confere validade ao instrumento (Pasquali, 2010, *American Education Research Association [AERA], American Psychological Association [APA] & Nacional Council on Measurement in Education [NCME]*, 2014). Já a TRI contribuiu para a investigação dos níveis de dificuldade e discriminação dos itens. Segundo Pasquali (2010), qualquer instrumento psicológico deveria utilizar a TRI para aferir a dificuldade e discriminação dos itens, tendo em vista que ela é uma técnica mais avançada, nesse quesito, do que a TCT. Portanto, a precisão da medida obtida ao longo de diversos níveis da habilidade leitura, indica uma avaliação adequada da amostra que compôs os nove anos escolares conferindo, assim, parâmetro apropriados de fidedignidade (AERA, APA, & NCME, 2014).

Considerações Finais

Considerando-se que a principal característica do Subteste de Leitura do TDE II é realizar um *screening* da habilidade de ler palavras, outros aspectos envolvidos na leitura não foram incluídos no instrumento (compreensão textual, por exemplo). A avaliação por meio desse subteste permite identificar sujeitos que tenham dificuldade na habilidade de leitura de palavras. Caso o sujeito apresente déficit nesse quesito, sugere-se realizar uma avaliação pormenorizada da área defasada.

O processo de desenvolvimento do subteste fez uso tanto do modelo clássico de construção de instrumentos quanto de modelos modernos. O método utilizado para a elaboração do Subteste de Leitura do TDE-II seguiu os passos postulados pelo modelo clássico. Já a análise dos itens foi realizada tanto pela TCT quanto pelo modelo moderno, ao analisar os níveis de dificuldade e discriminação por meio da TRI.

Espera-se que o Subteste de Leitura do TDE-II possa ser amplamente utilizado por professores, para avaliação do desempenho em leitura de seus alunos, bem como para avaliar o método de ensino. Ademais, almeja-se que psicólogos, fonoaudiólogos e psicopedagogos também façam uso clínico do subteste, a fim de identificar estudantes que possam apresentar defasagem na habilidade de leitura de palavras. Uma limitação do estudo a ser considerada é o fato de que todos os

participantes foram recrutados na região metropolitana de Porto Alegre, no Rio Grande do Sul. Portanto, para pesquisas futuras sugere-se a testagem empírica desse subteste com estudantes de outras regiões brasileiras, bem como a investigação de outras evidências de validade, como as validades convergente e divergente, além do estudo de padronização e normatização.

Referências

- Adams, K. M. (2000). Practical and ethical issues pertaining to test revisions. *Psychological Assessment*, 12(3), 281-286. doi: 10.1037//1040-3590.12.3.281
- American Education Research Association, American Psychological Association & Nacional Council on Measurement in Education. (2014). *Standards for Educational and Psychological Testing*. Washington: AERA, APA, NCME.
- Angelini, A. L., Alves, I. C. B., Custódio, E. M., Duarte, W. F., & Duarte, J. L. M. (1999). *Manual matrizes progressivas coloridas de Raven: Escala especial*. São Paulo: Centro Editor de Testes e Pesquisas em Psicologia.
- Athayde, M. L., Giacomoni, C. H., Zanon, C., & Stein, L. M. (2014). Evidências de validade do Subteste de Leitura do Teste de Desempenho Escolar. *Psicologia: Teoria e Prática*, 16(2), 131-140. doi: 10.15348/1980-6906/psicologia.v16n2p131-140
- Beaujean, A. A. (2014). *Latent variable modeling using R: A Step-by-Step Guide*. New York: Taylor & Francis.
- Capellini S. A., Oliveira A. M., & Cuetos F. (2010). *PROLEC: Provas de avaliação dos processos de leitura*. São Paulo: Casa do Psicólogo.
- Capovilla, A., & Capovilla, F. (2004). *Alfabetização: Método fônico*. São Paulo: Memnon.
- Capovilla, F., & Capovilla, A. (2001). Compreendendo a natureza dos problemas de aquisição de leitura e escrita: Mapeando o envolvimento de distúrbios cognitivos de discriminação fonológica, velocidade de processamento e memória fonológica. *Cadernos de Psicopedagogia*, 1(1), 14-37.
- Conselho Federal de Psicologia (2003). *Resolução CFP Nº 002/2003*. Recuperado de http://site.cfp.org.br/wp-content/uploads/2003/03/resolucao2003_02_Anexo.pdf na data 02/12/2015
- Corso, H. V., & Salles, J. F. (2009). Relação entre leitura de palavras isoladas e compreensão de leitura *Psico-USF, Bragança Paulista, v. 24, n. 2, p. 245-257, abr./jun. 2019*

- textual em crianças. *Letras de Hoje*, 44, 28-35. Recuperado de <http://revistaseletronicas.pucrs.br/ojs/index.php/fale/article/viewFile/5761/4181>
- Corso, H. V., Piccolo, L. R., Mina, C. S., & Salles, J. F. (2015). Normas de desempenho em compreensão de leitura textual para crianças de 1º ano a 6ª série. *Psico*, 46, 68-78. doi: 10.15448/1980-8623.2015.1.16900
- Cuetos, F., Rodrigues, B., Ruano, E. (2010). Adaptado para o português por Capellini, A., Oliveira, A. M., & Cuetos, F. *PROLEC: Provas de Avaliação dos Processos de Leitura*. 1 edição. São Paulo: Casa do Psicólogo.
- Cunha, V. L. O., Capellini, S. A. (2009). Prohmele: Provas de habilidades metalinguísticas e de leitura. Rio de Janeiro: Revinter.
- Der Ark, L. V. (2007). Mokken scale analysis in R. *Journal of Statistical Software*, 20(11).
- Ferreira, A. T. B., & Leal, T. F. (2007). Avaliação na escola e ensino da língua portuguesa: Introdução ao tema. (pp. 11-26). Em B. Marcurschi & L. Suassuna. *Avaliação em língua portuguesa: Contribuições para a prática pedagógica*. 1ª edição, 1 reimpressa. Belo Horizonte: Autêntica.
- Field, A., Miles, J., & Field, Z. (2012). *Discovering statistics using R*. London: Sage.
- Fonseca, R. P., Jacobsen, G. M., & Pureza, J. R. (2015). O que um bom teste neuropsicológico deve ter? Em J. F. Salles, V. G. Haase & L. F. Malloy-Diniz *Neuropsicologia do desenvolvimento: Infância e adolescência*. Porto Alegre: Artmed.
- Fonseca, R. P., Zimmermann, N., Pawlowski, J., Oliveira, C. R., Gindri, G. Scherer, L. C., Rodrigues, J. C., & Parente, M. A. M. P. (2012). Métodos em avaliação neuropsicológica. Em J. Landeira-Fernandez & Sérgio S. Fukusima. (Ed.). *Métodos em neurociência*. São Paulo: Manole.
- Fonseca, R. P., Salles, J. F., & Parente, M. A. M. P. (2009). *Instrumento de Avaliação Neuropsicológica Breve NeuPsiLin*. São Paulo, SP: Vector.
- Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2010). *Multivariate Data Analysis*. Pearson Prentice Hall.
- Horn, J. (1965). A rationale and test for the number of factors in factor analysis. *Psychometrika*, 30(2), 179-185. Recuperado de <http://link.springer.com/article/10.1007/BF02289447>
- International Test Commission. (2000). *Diretrizes para o uso de testes* (Instituto Brasileiro de Avaliação Psicológica, Trad., 2003). Recuperado de https://www.intestcom.org/files/guideline_test_use_portuguese_brazil.pdf
- International Test Commission. (2000). *Diretrizes para o uso de testes* (Instituto Brasileiro de Avaliação Psicológica, Trad., 2003). Recuperado de https://www.intestcom.org/files/guideline_test_use_portuguese_brazil.pdf
- Junker, B., & Sijtsma, K. (2000). Latent and manifest monotonicity in item response models. *Applied Psychological Measurement*, 24(1), 65-81. Recuperado de <http://apm.sagepub.com/content/24/1/65.short>
- Kim, Y. S., Park, C. H., & Wagner, R. K. (2014). Is oral/text reading fluency a “bridge” to reading comprehension? *Reading and Writing*, 27, 79-99. doi: 10.1007/s11145-013-9434-7
- Knijnik, L. F., Giacomoni, C., & Stein, L. M. (2013). Teste de desempenho escolar: Um estudo de levantamento. *Psico-USF*, 18(3), 407-416. doi: 10.1590/S1413-82712013000300007
- Knijnik, L. F., Zanon, C., Giacomoni, C. H., & Stein, L. M. (2014). Avaliação dos subtestes de Leitura e Escrita do Teste de Desempenho Escolar através da Teoria de Resposta ao Item. *Psicologia: Reflexão e Crítica*, 27(3), 481-490. doi: 10.1590/1678-7153.201427308
- Laros, J. A. (2005). O uso da análise fatorial: Algumas diretrizes para pesquisadores. Em L. Pasquali (Ed.), *Análise fatorial para pesquisadores* (pp. 163-184). Brasília: LabPAM.
- Lúcio, P. S., & Pinheiro, A. M. V. (2014). Novos estudos psicométricos para o subteste de leitura do teste de desempenho escolar. *Temas em Psicologia*, 22(1), 109-119. doi: 10.9788/TP2014.1-09
- Lúcio, P. S., Moura, R. J., Nascimento, E., & Vieira, A. M. P. (2012). Construção de uma tarefa de leitura em voz alta de palavras: Análise psicométrica dos itens. *Psicologia: Reflexão e Crítica*, 25(4), 662-670. doi: 10.1590/S0102-79722012000400005
- Lúcio, P. S., Pinheiro, A. M. V., & Nascimento, E. (2009). O impacto da mudança no critério de acerto na

- distribuição dos escores do subteste de leitura do teste de desempenho escolar. *Psicologia em Estudo*, 14(3), 593-601. Recuperado de <http://www.scielo.br/pdf/pe/v14n3/v14n3a21>
- Pasquali, L. (2010). *Instrumentação psicológica: Fundamentos e práticas*. Porto Alegre: Artmed, 2010. 560 p.
- Pasquali, L., & Primi, R. (2003). Fundamentos da Teoria da Resposta ao Item – TRI. *Avaliação Psicológica*, 2(2), 99-110. Recuperado de http://pepsic.bvsalud.org/scielo.php?script=sci_arttext&pid=S1677-04712003000200002
- Reise, S. P. (1990). A Comparison of Item- and Person-Fit Methods of Assessing Model-Data Fit in IRT. *Applied Psychological Measurement*, 14(2), 127-137. doi: 10.1177/014662169001400202
- Revelle, A. W., & Revelle, M. W. (2015). Package “psych.”
- Revelle, W. (2018) psych: Procedures for Personality and Psychological Research, Northwestern University, Evanston, Illinois, USA, <https://CRAN.R-project.org/package=psych> Version = 1.8.12.
- Rizopoulos, D. (2006). ltm: An R package for latent variable modeling and Item Response Theory analyses. *Journal of Statistical Software*, 17(5), 1-25. doi: 10.18637/jss.v017.i05
- Rueda, F. J. M., Noronha, A. P. P., Sisto, F. F., Santos, A. A. A., & Castro, N. R. (2012). *Escala de Inteligência Wechsler para Crianças – WISC-IV*. São Paulo: Casa do Psicólogo.
- Rusch, T., Mair, P., & Hatzinger, R. (2013). Psychometrics with R : A review of CRAN packages for Item Response Theory. Discussion paper series of the center for empirical research methods.
- Salles, J. F., & Parente, M. A. P. P. (2002). Relação entre os processos cognitivos envolvidos na leitura de palavras e as habilidades de consciência fonológica em escolares. *Pró-Fono Revista de Atualização Científica*, 14(2), 141-286. Recuperado de <http://www.ufrgs.br/neurocog/ambulatorio/tarefa-de-leitura-de-palavras-isoladas>
- Salles, J. F. (2005). Habilidades e dificuldades de leitura e escrita de crianças de 2ª série: A neuropsicológica cognitiva. Porto Alegre: Universidade Federal do Rio Grande do Sul.
- Salles, J. F., Piccolo, L. R., Zamo, R., & Toazza, R. (2013). Normas de desempenho em tarefa de leitura de palavras/pseudopalavras isoladas (LPI) para crianças de 1º ano a 6ª série. *Estudos e Pesquisas em Psicologia*, 13, 1-10. Recuperado de <http://www.e-publicacoes.uerj.br/index.php/revispsi/article/view/8416/6249>
- Saraiva, R. A., Moojen, S. M. P., & Munarski, R. (2009). *Avaliação da compreensão leitora de textos expositivos: Para fonoaudiólogos e psicopedagogos*. São Paulo: Casa do Psicólogo.
- Scliar-Cabral, L. (2013). A desmistificação do método global. *Letras de Hoje*, 48(1), 6-11. Recuperado de <file:///C:/Users/Marcia/Desktop/12142-51127-1-PB.pdf>
- Seabra, A. G., & Capovilla, F. C. (2010). Teste de Competência de Leitura de Palavras e Pseudopalavras (TCLPP). São Paulo: Memnon.
- Smith, G. T., & McCarthy, D. M. (1995). Methodological considerations in the refinement of clinical assessment instruments. *Psychological Assessment*, 7, 300-308. doi: 10.1037/1040-3590.7.3.300
- Stein, L. M. (1994). *TDE - Teste de Desempenho Escolar: Manual para aplicação e interpretação*. São Paulo: Casa do Psicólogo.
- Van der Linden, W. J., & Hambleton, R. K. (1996). Handbook of modern Item Response Theory. (W. J. van der Linden & R. K. Hambleton, Eds.). Springer.
- Vilhena C., Guntert. I. B., & Tosi, A. D., (in press). Teste Matrizes Progressivas de Raven. São Paulo: Casa do Psicólogo.

Recebido em: 08/01/2018

Reformulado em: 15/08/2018

Aprovado em: 17/08/2018

Sobre os autores:

Marcia de Lima Athayde é graduada em Fonoaudiologia pela Universidade Federal de Santa Maria, mestre em Distúrbios da Comunicação Humana pela UFSM e doutora em Psicologia pela Pontifícia Universidade Católica do Rio Grande do Sul. Fonoaudióloga da Universidade Federal do Rio Grande do Sul.

ORCID: <https://orcid.org/0000-0001-6855-7575>

Email: marcialadoc@gmail.com

Euclides José de Mendonça Filho é graduado em Psicologia pela Universidade Federal da Bahia (UFBA), mestre e doutorando pela Universidade Federal do Rio Grande do Sul (UFRGS), pesquisador do Grupo de Estudos, Aplicação e Pesquisa em Avaliação Psicológica (GEAPAP – UFRGS) e pesquisador visitante no Silveira Lab (McGill University- Montreal-QC)

ORCID: <https://orcid.org/0000-0003-4098-8247>

Email: euclidesmendonca.f@gmail.com

Rochele Paz Fonseca é graduado em Fonoaudiologia pela Universidade Luterana do Brasil e em Psicologia pela Universidade Federal do Rio Grande do Sul (UFRGS), mestre e doutora em Psicologia do Desenvolvimento pela UFRGS, com pós-doutorado em Clínica e Neurociências pela PUC-Rio, em Medicina pela UFRJ e em Ciências Biomédicas pela Universidade de Montreal e professora Adjunta da Faculdade de Psicologia e do Programa de Pós-Graduação da PUCRS.

ORCID: <https://orcid.org/0000-0001-6767-8439>

Email: rochele.fonseca@gmail.com

Lilian Milnitsky Stein é graduada em Psicologia pela Universidade Federal do Rio Grande do Sul (UFRGS), especialista em Psicologia Escolar pela PUCRS, mestre em Ciência Cognitiva Aplicada pelo Instituto Ontário para Estudos em Educação (Canadá), doutora em Psicologia Cognitiva pela Universidade do Arizona (EUA), com pós-doutorado na Universidade de Barcelona (Espanha) e professora titular do Programa de Pós-Graduação em Psicologia da PUCRS.

ORCID: <https://orcid.org/0000-0002-2314-7369>

Email: stein.lilian@gmail.com

Claudia Hofheinz Giacomoni é graduada em Psicologia pela Pontifícia Universidade Católica do Rio Grande do Sul (PUCRS), mestre em Psicologia do Desenvolvimento pela Universidade Federal do Rio Grande do Sul (UFRGS), doutora em Psicologia pela UFRGS, professora do Departamento de Psicologia do Desenvolvimento e da Personalidade e do Programa de Pós-graduação em Psicologia da UFRGS e coordenadora do NEPP (Núcleo de Estudos em Psicologia Positiva).

ORCID: <https://orcid.org/0000-0002-9183-8591>

Email: giacomonich@gmail.com

Contato com os autores:

Rua Ramiro Barcelos, sala 224, Santa Cecília

Porto Alegre-RS, Brasil

CEP: 90035-003