

## 1. DADOS DO INSCRITO

**Categoria da Inscrição:** Inovação de Avaliações e Protocolos

**Eixo:** 02 – Avaliação psicométrica

**Grupo:** Laboratório de Processos Cognitivos da Universidade Federal de Minas Gerais  
(LabCog-UFMG)

**Autor 1 (correspondente):** Dr. Douglas de Araújo Vilhena

**Endereço completo:** Av. Antônio Carlos, 6627, Pampulha, Belo Horizonte, MG, CEP:  
31.270-901

**País:** Brasil

**CPF:** 073.436.056-82

**Instituição de Origem:** Laboratório de Processos Cognitivos (LabCog), Programa de Pós-Graduação em Psicologia: Cognição e Comportamento (PsiCogCom), Universidade Federal de Minas Gerais (UFMG)

**E-mail de contato:** [douglasvilhena@ufmg.br](mailto:douglasvilhena@ufmg.br)

**Autora 2:** Dr.<sup>a</sup> Ângela Maria Vieira Pinheiro

**Endereço completo:** Av. Antônio Carlos, 6627, Pampulha, Belo Horizonte, MG, CEP:  
31.270-901   **País:** Brasil

**CPF:** 326.021.466-68

**Instituição de Origem:** Laboratório de Processos Cognitivos (LabCog), Programa de Pós-Graduação em Psicologia: Cognição e Comportamento (PsiCogCom), Universidade Federal de Minas Gerais (UFMG)

**E-mail de contato:** [pinheiroamva@gmail.com](mailto:pinheiroamva@gmail.com)

## 2. DADOS DO TRABALHO

**Título do trabalho:** Evidências psicométricas do Teste de Reconhecimento de Palavras e do Teste de Reconhecimento de Pseudopalavras

**Resumo:** Objetivou-se apresentar uma síntese das evidências psicométricas do Teste de Reconhecimento de Palavras (TRP) e do Teste de Reconhecimento de Pseudopalavras (TRPp). Participaram 598 alunos do segundo ao quinto ano do Ensino Fundamental, 7 a 11 anos, 52% meninos, estratificados em 102 salas de aula de oito Escolas Estaduais. Como evidência de fidedignidade, houve consistência interna alta e equivalência entre as duas metades e as formas paralelas. A validade de conteúdo foi fortalecida com as transcrições fonêmicas e com os significativos efeitos de lexicalidade, frequência, regularidade e extensão. Validade da Estrutura Interna foi verificada com o efeito de escolaridade e curva de distribuição normal. Diferentes validades externas convergente, discriminante e de critério foram verificadas. Concluiu-se que os testes propostos demonstraram evidências psicométricas satisfatórias, o que sustenta o estabelecimento de normas padronizadas de conversão do escore bruto para Percentil.

**Palavras-chave:** avaliação educacional, deficiências da aprendizagem, dislexia, estudo de validação, leitura, psicometria.

### 1. Introdução e justificativa

*O Teste de Reconhecimento de Palavras (TRP) e o Teste de Reconhecimento de Pseudopalavras (TRPp) são instrumentos psicométricos padrão-ouro para avaliar a leitura em voz alta de estudantes do segundo ao quinto ano do Ensino Fundamental*

brasileiro (Pinheiro & Vilhena, 2022; Vilhena & Pinheiro, 2022). Fundamentam-se a partir de um compreensivo referencial teórico sobre o reconhecimento visual de palavras, que considerou os efeitos psicolinguísticos de frequência de ocorrência, de regularidade grafema-fonema, de extensão e de lexicalidade.

Os estudos psicométricos dos instrumentos, que incluem evidências de validade de conteúdo e validade externa (Pinheiro & Vilhena, 2022), validade interna, fidedignidade e normatização (Vilhena & Pinheiro, 2022), complementam os dados de fidedignidade e da estrutura interna encontrados em estudos prévios do TRP e TRPp (Cogo-Moreira *et al.*, 2012; Pinheiro, 2013; Pinheiro *et al.*, 2017; Vilhena *et al.*, 2016; Vilhena & Pinheiro, 2016, 2020). Em conjunto, demonstram que ambos avaliam o construto ‘decodificação grafêmica’.

A criação e validação psicométrica do TRP e do TRPp justifica-se pela importância da identificação precoce dos alunos com dificuldade de leitura e da avaliação diagnóstica da dislexia do desenvolvimento. Justifica-se também pela necessidade de os profissionais terem suas práticas baseadas em evidências, com parâmetros para o acompanhamento das intervenções e que atendam às normas internacionais [*i.e.*, CID-11 (OMS, 2022) e DSM-5 (APA, 2013)] e às nacionais (*ex.*, Lei nº 14.254 de 2021, que dispõe sobre o acompanhamento integral para educandos com dislexia).

## **2. Problema e Objetivo**

Todas as crianças têm o direito de aprender a ler, sendo que a leitura proficiente é uma das principais ferramentas para a redução das desigualdades sociais. Isso torna fundamental a identificação precoce das crianças com dificuldades na alfabetização ou com transtorno específico de aprendizagem da leitura.

A variável mais utilizada na literatura é a acurácia, também referida como exatidão e precisão, que é calculada como a porcentagem de palavras lidas corretamente. Um problema psicométrico com os instrumentos que medem apenas a acurácia é que estão sujeitos ao efeito de teto, conforme frequentemente se observa (Vilhena & Pinheiro, 2022), especialmente nas ortografias transparentes, em que a acurácia de leitura frequentemente atinge seu máximo após poucos meses de instrução formal (Cossu *et al.*, 1995; Landerl & Wimmer, 2008).

Por essa razão, a utilização do tempo de processamento, em adição à acurácia como escores de um teste, faz especial sentido nas ortografias transparentes porque, em contraste com as opacas, a leitura nos estágios iniciais de aprendizagem, assim como as dificuldades nessa habilidade, caracterizam-se por lentidão no reconhecimento de palavras, mesmo se acompanhado por alto índice de precisão (Marinelli *et al.*, 2016; Wimmer, 1993; de Jong & van der Leij, 2003).

Levando em consideração esse achado, o TRP e o TRPp apresentam como escores a acurácia e o tempo de processamento (taxa de acurácia), sendo a combinação destas duas variáveis um dos diferenciais dos testes. Com base nos resultados dessas variáveis, o presente trabalho tem o objetivo de apresentar uma síntese das fontes de evidências psicométricas do TRP e do TRPp no que concerne a: (a) fidedignidade; (b) validade de conteúdo; (c) validade da estrutura interna; (d) validade externa, e (e) normatização.

### **3. Método dos estudos psicométricos**

Estudo com delineamento transversal, quantitativo, comparativo e populacional. As definições contemporâneas de validade estabelecidas pelo *Standards for Educational and Psychological Testing* foram levadas em consideração (AERA *et al.*, 2014), assim

como as demandas do Comitê de Ética em Pesquisa da Universidade Federal de Minas Gerais (CAAE: 17754514.6.0000.5149).

### *3.1.1. Participantes*

Participaram da pesquisa 598 alunos voluntários do segundo ao quinto ano do Ensino Fundamental, 52,3% do sexo masculino, de 7 a 11 anos ( $M = 9,6 \pm 1,1$  anos), todos falantes nativos do Português Brasileiro, com a média da capacidade cognitiva geral classificada como média superior (Percentil 76).

Foram sorteadas oito escolas estaduais, estratificadas em diferentes regionais da cidade de Belo Horizonte, Minas Gerais. Em cada uma das 102 salas de aula, apenas seis alunos foram selecionados, de forma aleatória (sorteio a partir da lista de presença) para responder a uma bateria de avaliação cognitiva. Para o estudo da validade convergente e discriminante, parte dos professores ( $n = 72$ ) completaram uma escala de leitura (EACOL) e de comportamentos externalizantes (SDQ) para cada um dos alunos participantes ( $n = 452$ ).

### *3.1.2. Instrumentos*

**Teste de Reconhecimento de Palavras (TRP):** avalia a leitura de palavras isoladas (Pinheiro & Vilhena, 2023). É composto por 88 palavras selecionadas por Pinheiro (2013), sendo quatro de treino e 84 de teste. As palavras variam com relação à frequência de ocorrência (Pinheiro, 2015), com número iguais de itens de alta e de baixa frequência. Em cada um destes níveis, as palavras variam ainda em regularidade grafema–fonema e em comprimento (quatro a oito letras). O TRP possui três formas paralelas, cujas versões variam apenas em relação à ordem das palavras (Versão A, Versão B e Versão C). As versões do TRP são apresentadas em cartões plastificados,

cada um contendo 11 linhas, com seis a nove palavras por linha, impressos em tinta preta em papel branco A4, em fonte Arial, tamanho 14. Pinheiro e Vilhena (2022) e Vilhena e Pinheiro (2022) proveram as principais fontes de evidências psicométricas. O score é dado em: (a) Acurácia: porcentagem de palavras lidas corretamente (ex., 75 palavras corretas correspondem a uma acurácia de 89%); e (b) Taxa de acurácia: número total de palavras lidas corretamente, multiplicado por sessenta e dividido pelo tempo total em segundos (ex., 75 palavras corretas em 120 segundos correspondem à taxa de acurácia de 37,5 palavras por minuto).

**Teste de Reconhecimento de Pseudopalavras (TRPp):** avalia a leitura de pseudopalavras isoladas (Pinheiro & Vilhena, 2023). É composto por 88 pseudopalavras (ex., *pederno, clandas, verra*), sendo quatro de treino e 84 de teste, selecionadas por Pinheiro (2013). O TRPp foi construído a partir do TRP, sendo que cada pseudopalavra mantém a mesma estrutura ortográfica e número de letras da palavra de referência. O TRPp possui três formas paralelas, cujas versões variam apenas em relação à ordem das palavras (Versão D, Versão E e Versão F). O score é dado em acurácia e taxa de acurácia.

**Teste de Leitura: Compreensão de Sentenças (TELCS):** avalia a compreensão de leitura de sentenças (Vilhena & Pinheiro, 2022). É composto por 36 sentenças isoladas, sendo sempre omitida a última palavra. São oferecidas cinco alternativas de resposta apresentadas em forma de múltipla escolha, sendo que apenas uma delas dá sentido à sentença.

**Subteste Compreensão de Textos das Provas de Avaliação dos Processos de Leitura (PROLEC-Texto):** avalia os processos semânticos de compreensão literal de texto (Capellini *et al.*, 2012). É composto por quatro textos curtos e quatro questões literais sobre cada um deles (máximo: 16 pontos).

**Escala de Avaliação da Competência em Leitura pelo Professor (EACOL):** avaliação indireta do desempenho em leitura preenchido pelo professor (Pinheiro, & Costa, 2015; Vilhena & Pinheiro, 2016). É composta por descritores operacionais da habilidade de leitura que podem ser reconhecidos pelo professor.

**Teste de Matrizes Progressivas Coloridas de Raven (Raven CPM):** avalia a capacidade cognitiva geral não verbal de crianças entre 5 e 11 anos (Angelini *et al.*, 1999). O construto avaliado é o raciocínio analógico, que é a capacidade de inferir relações entre objetos ou elementos (Pasquali, *et al.*, 2002).

**Questionário de Capacidades e Dificuldades (SDQ):** escala para triar comportamentos externalizantes, embora também avalie o comportamento social positivo, entre a faixa etária de 4 a 16 anos, respondida pelos professores (Fleitlich-Bilyk *et al.*, 2000). O instrumento possui índices adequados de fidedignidade e validade em 21 países, incluindo o Brasil (Saur & Loureiro, 2012).

### 3.1.3. Procedimentos

Todos os instrumentos foram administrados no mesmo dia, em duas sessões, cada uma com duração média de 15 minutos. Na primeira sessão, grupos de até dez crianças foram submetidos coletivamente ao TELCS e ao Raven CPM. Na segunda sessão, cada criança foi individualmente submetida ao TRP e ao TRPp (administrados em sequência, mas em ordem aleatória), seguidos pela PROLEC-Texto. Os professores responderam a EACOL e ao SDQ, para cada aluno, durante o período de uma semana.

Foi requerido aos participantes, individualmente, ler em voz alta o TRP e o TRPp com o maior número de acertos e o mais rápido possível. O teste é fácil de administrar e de pontuar. É iniciado com os quatro itens de treino, seguidos dos 84 itens

de teste, lidos em voz alta pelo participante da esquerda para a direita, de cima para baixo.

#### 4. Resultados dos estudos psicométricos

Como evidência de **fidedignidade**, (1) os coeficientes de consistência interna (*Kuder-Richardson*, *Spearman-Brown* e *Guttman*) do TRP e TRPp apresentaram grau de covariância próximo ao máximo ( $KR-20 = 0,98$ ). (2) A equivalência foi verificada pelo método das duas metades (*split-half*), com os itens foram repartidos em duas pares vs. ímpares, com índices de correlações *Spearman-Brown* e *Guttman* muito altos para o TRP (0,95) e para o TRPp (0,96). (3) Como outra evidência de equivalência, não houve diferença entre as três versões paralelas utilizadas do TRP [Versão A, Versão B e Versão C ( $F_{(2, 478)} = 1,7, p = 0,19$ )] e as três versões do TRPp [Versão D, Versão E e Versão F ( $F_{(2, 478)} = 1,1, p = 0,33$ )].

Para a **validade de conteúdo**, foram providas as transcrições fonêmicas de cada palavra e pseudopalavra, que por sua vez foram equilibradas de acordo com quatro características psicolinguísticas: (1) lexicalidade (palavras e pseudopalavras) ( $F_{(2,594)} = 203,7, p < 0,001, d = 1,14$ ); (2) frequência de ocorrência (alta e baixa) [TRP: ( $F_{(1,594)} = 10,5, p = 0,0017, d = 0,70$ )]; (3) extensão (curta, média e longa) [TRP: ( $F_{(1,594)} = 6,0, p = 0,0017, d = 0,64$ ); TRPp: ( $F_{(2,594)} = 6,8, p = 0,011, d = 0,66$ )]; e (4) regularidade (semirregulares e irregulares) [TRP: ( $F_{(1,594)} = 9,8, p = 0,0028, d = 0,72$ )].

Para a **validade da estrutura interna**, considera-se que (1) a seleção das palavras do TRP ocorreu a partir de um banco palavras analisado via Teoria de Resposta ao Item (TRI), para a seleção de palavras com índices de discriminação de médio a alto e três níveis de dificuldade (baixa, média e alta). (2) A variável Taxa de Acurácia demonstrou índices de validade da estrutura interna satisfatórios, com distribuição

normal padrão, simétrica, mesocúrtica e significativa **efeito de escolaridade** para o TRP [ $F_{(3, 594)} = 63,2, MSE = 391,4, p < 0,001$ ] e para o TRPp [ $F_{(4, 594)} = 47,3, MSE = 129,2, p < 0,001$ ], sem sobreposição dos Intervalos de Confiança de 95% entre as médias dos anos escolares.

Para a evidências de **validade externa**, investigou-se diferentes fontes de evidências de validade convergente, discriminante e de critério. Com relação às evidências de **validade externa convergente**, as variáveis Acurácia e Taxa de Acurácia do TRP e do TRPp demonstraram significantes correlações bivariadas ( $p < 0,001$ ), de moderadas a fortes, entre si e com todos os instrumentos de avaliação da (a) habilidade de leitura: TELCS; PROLEC-Texto; Escala de Avaliação da Competência em Leitura pelo Professor (EACOL) (Pinheiro & Costa, 2015; Vilhena & Pinheiro, 2016); (b) capacidade cognitiva: CPM; e (c) dados demográficos.

As correlações mais fortes foram entre o desempenho de leitura de sentenças do TELCS com as Taxas de Acurácia do TRP ( $r = 0,84$ ) e do TRPp ( $r = 0,79$ ), cujas variáveis envolvem a acurácia dentro de um tempo delimitado. Com base nos coeficientes estimados da reta de regressão do gráfico de dispersão, verificou-se relação linear positiva entre a Taxa de Acurácia do TRPp em função do TRP ( $r = 0,90$ ;  $y = 0,46x + 5,86$ ;  $R^2 = 0,6869$ ;  $p < 0,001$ ).

Como **validade externa discriminante**, ambas as variáveis do TRP e do TRPp demonstraram correlações fracas com comportamentos externalizantes avaliados por meio do SDQ ( $r = -0,36$  a  $-0,11$ ).

Evidências de **validade externa de critério** serão providas caso os valores de Acurácia e Taxa de Acurácia forem capazes de prever critérios estabelecidos pela literatura, além dos já demonstrados efeitos psicolinguísticos de lexicalidade,

frequência, regularidade e extensão. Considerou-se as seguintes quatro validades de critério:

(1) Como primeiro critério, a Acurácia e a Taxa de Acurácia do TRP e TRPp devem ser variáveis importantes para predizer o desempenho escolar. Parte dos participantes ( $n = 316$ ) foram classificados por desempenho escolar nas disciplinas Língua Portuguesa e Matemática, de acordo com o Conceito Final (soma das notas dos quatro bimestres letivos), que variou de Conceito A (conceito máximo) a Conceito C (conceito mínimo encontrado). No TRP, para essas disciplinas, os alunos com Conceito A apresentaram escores maiores do que os alunos com Conceito B e Conceito C, tanto para a variável Acurácia ( $F_{(3,312)} > 16,8, p < 0,001$ ), quanto para a Taxa de Acurácia ( $F_{(1,314)} > 10,5, p < 0,001$ ). O mesmo padrão ocorreu com o TRPp (Conceito A > B > C), para ambas as disciplinas, tanto para a variável Acurácia ( $F_{(2,313)} > 30,2, p < 0,001$ ) quanto para a Taxa de Acurácia ( $F_{(2,313)} > 14,2, p < 0,001$ ).

(2) De acordo com o segundo critério, o sexo masculino e o feminino devem apresentar desempenho de leitura equivalentes (Vilhena & Pinheiro, 2020). A análise de variância não demonstrou diferença significativa de sexo para a variável Acurácia e Taxa de Acurácia, tanto para o TRP quanto para o TRPp ( $p > 0,15$ ).

(3) A terceira evidência de validade de critério foi verificada ao ser possível encontrar, por meio da comparação dos resultados do TRP com do TRPp, participantes com déficits específicos. O perfil de desempenho dos leitores, quanto ao contraste no reconhecimento de palavras e de pseudopalavras, foi enquadrado em quatro grupos: (a) 80% com *desempenho competente*; (b) 15% com *déficit misto*; (c) 3% com *déficit específico no processamento de palavras*; e (d) 2% de *déficit específico no processamento de pseudopalavras*.

(4) A quarta evidência de validade de critério será provida caso os valores de Acurácia forem coerentes com a literatura. As médias de Acurácia do TRP (92%) e do TRPp (83%) apresentaram valores coerentes aos encontrados em distintos instrumentos que avaliam o mesmo construto.

Uma vez que o Teste de Reconhecimento de Palavras (TRP) e o Teste de Reconhecimento de Pseudopalavras (TRPp) demonstraram fontes de evidências psicométricas satisfatórias, foi possível sustentar o **estabelecimento de normas padronizadas** para o segundo ao quinto ano do Ensino Fundamental ( $N = 598$  participantes), com os escores brutos das variáveis Acurácia e Taxa de Acurácia convertidos para o Percentil correspondente. Foram criados pontos de corte para a classificação do desempenho de leitura entre transtorno (Percentil 0 a 7), abaixo da média (Percentil 8 a 25), média (Percentil 26 a 74), acima da média (Percentil 75 a 94) e alto (Percentil 95 a 99) ( $p < 0,001$ ).

## 5. Discussão

Cumpriu-se o objetivo de apresentar as fontes de evidências psicométricas do Teste de Reconhecimento de Palavras (TRP) e do Teste de Reconhecimento de Pseudopalavras (TRPp) no que concerne a: (a) fidedignidade; (b) validade de conteúdo; (c) validade da estrutura interna; (d) validade externa, e (e) normatização. Ambos os instrumentos apresentaram evidências psicométricas satisfatórias, destacando-se a variável Taxa de Acurácia, para avaliar a decodificação grafêmica de alunos do segundo ao quinto ano do Ensino Fundamental brasileiro.

O estabelecimento de normas representativas da população escolar contribui com a avaliação nas áreas da Psicologia escolar e educacional no Brasil. O TRP e o TRPp primam-se devido a concepção teórica, que permite o teste de efeitos

psicolinguísticos clássicos encontrados na literatura e replicados tanto nas ortografias opacas (Provazza *et al.* 2019; Schuster *et al.*, 2016) quanto nas ortografias transparentes (Carreiras *et al.*, 2006; Pinheiro, 1995, 1999; Wimmer *et al.*, 2010).

Os resultados do estudo de **fidedignidade**, com **consistência interna** alta e grau de covariância próximo ao máximo, são uma evidência forte de que os itens do TRP e do TRPp medem o mesmo construto (*i.e.*, decodificação grafêmica). Os resultados das análises pelo método das duas metades demonstraram que o mesmo indivíduo obteve um escore médio equivalente nas duas metades (itens pares *vs.* ímpares) do TRP e do TRP, o que demonstra a homogeneidade dos instrumentos. Houve **equivalência** entre as três formas paralelas do TRP (Versão A, Versão B, Versão C) e as três do TRPp (Versão D, Versão E, Versão F), o que permite alternar, sem prejuízo psicométrico, a apresentação das versões para evitar o efeito de ordem dos itens.

Para a **validade de conteúdo**, foram apresentadas a transcrição fonêmica de todos os itens e a descrição detalhada das definições operacionais e constitutivas dos itens do TRP e do TRPp. O **efeito de lexicalidade** foi demonstrado por meio do contraste, significativo e com tamanho de efeito forte, entre a leitura do conjunto de palavras com o de pseudopalavras. Isso demonstra que a lexicalidade é um fator facilitador da Acurácia e da Taxa de Acurácia, com aumento da precisão e redução do tempo de leitura. A leitura de pseudopalavras, por ser realizada de forma serial e sem o apoio do léxico fonológico, é mais imprecisa e morosa do que a leitura de palavras (Coltheart *et al.*, 2001).

Para o TRP, foram também encontrados os **efeitos de frequência** de ocorrência, de **regularidade** e de **extensão** que, respectivamente, revelaram vantagem em termos de acurácia para as palavras frequentes, regulares e curtas em comparação às não frequentes, irregulares e longas.

Em relação às evidências de **validade baseada na estrutura interna**, tanto a variável Acurácia quanto a Taxa de Acurácia apresentaram **efeito de escolaridade** significativa, conseguiram distinguir os leitores por ano de escolar ( $2^{\circ} < 3^{\circ} < 4^{\circ} < 5^{\circ}$  ano). A Taxa de Acurácia possui uma curva padrão de distribuição normal, sendo simétrica (implica que a média, a mediana e a moda são coincidentes) e mesocúrtica (grau de achatamento normal), sem sobreposição dos Intervalos de Confiança entre os anos escolares.

Evidências de **validade externa** incluíram as validações convergente, discriminante e de critério. Para **validade externa convergente**, como esperado, TRP e TRPP apresentaram fortes correlações entre si, o que demonstra grau de dependência estatística linear entre as variáveis dos testes, pois apresentam substratos cognitivos compartilhados. Da mesma forma, o TRP e o TRPp apresentaram correlações significantes ( $p < 0,001$ ) de moderadas a fortes com diferentes instrumentos de leitura de comparação.

A Taxa de Acurácia do TRP e do TRPp apresentou correlação forte com o Teste de Leitura: Compreensão de Sentenças (TELCS) ( $r = 0,84$  e  $0,79$ ). Correlações fracas a moderadas foram encontradas entre a Acurácia (TRP e TRPp) e a compreensão de textos, avaliada por meio de perguntas literais da PROLEC-Texto ( $r = 0,43$  e  $0,34$ ) (Pinheiro *et al.*, 2017). Fornecem evidências de **validade discriminante** para o TRP e o TRPp as correlações fracas ( $r = -0,36$  a  $-0,11$ ) entre a Acurácia e a Taxa de Acurácia e os comportamentos externalizantes, investigados via SDQ, mostrando que os testes avaliam construtos distintos.

Quatro diferentes evidências de **validade de critério** foram providas para o TRP e o TRPp, demonstrando que os instrumentos são adequados para predizer o desempenho de leitura e escolar. Primeiramente, o TRP e o TRPp são bons preditores

do Conceito Final recebido pelos escolares ao término do ano letivo nas disciplinas Língua Portuguesa e Matemática. Este resultado revela que o grau de eficiência de leitura é refletido nas disciplinas escolares, com os alunos com desempenho acadêmico fraco (Conceito B e C) apresentando, em média, dificuldade na Acurácia e na Taxa de Acurácia.

Como segunda evidência de validade de critério, tanto a variável Acurácia quanto a Taxa de Acurácia do TRP e do TRPp não apresentaram diferença significativa de sexo, o que corrobora a literatura (Athayde *et al.*, 2014; Lúcio *et al.*, 2018; Vilhena & Pinheiro, 2020).

A terceira evidência de validade externa de critério revela a importância do uso em conjunto do TRP e do TRPp para verificar déficits específicos no desempenho de leitura. Foi possível verificar quatro grupos: (a) desempenho competente (80%); (b) déficit misto (15%); (c) déficit específico no processamento de palavras (3%); e (d) déficit específico no processamento de pseudopalavras (2%).

Como quarto critério, a média elevada de Acurácia do TRP (92%) está coerente com a literatura (Capellini *et al.*, 2012; Justi & Justi, 2009; Lúcio *et al.*, 2009; Salles *et al.*, 2013; Stivanin & Scheuer, 2007). De forma geral, as provas de leitura de palavras isoladas são fáceis de serem lidas, com crianças típicas alcançando altos índices de Acurácia, coerente com a transparência ortográfica para a leitura do português brasileiro.

As normatizações dos escores do TRP e do TRPp foram apresentadas, uma vez que os instrumentos mostraram fontes de evidências de fidedignidade e validade satisfatórias para alunos do segundo ao quinto ano do Ensino Fundamental. Uma força do estudo é a amostragem aleatória, estratificada e com tamanho estatístico adequado. Nesse sentido, foi possível estabelecer normas de conversão das variáveis Acurácia e da

Taxa de Acurácia para o seu correspondente Percentil e classificação do desempenho de leitura.

## 6. Conclusão dos estudos psicométricos

O TRP e o TRPp são testes padrão-ouro para a avaliação da decodificação grafêmica. Ambos os instrumentos apresentaram fontes de evidências psicométricas satisfatórias, para a avaliação das habilidades de leitura de crianças que estejam estudando do segundo ao quinto ano do Ensino Fundamental, no que concerne a: (a) fidedignidade (equivalência e consistência interna); (b) validade de conteúdo; (c) validade da estrutura interna; (d) validade externa (convergente, discriminante, critério); e (e) normatização.

## 7. Referências

- American Educational Research Association – AERA, American Psychological Association – APA, & National Council on Measurement in Education – NCME. (2014). *Standards for educational and psychological testing*. Washington, DC.
- Angelini, A. L., Alves, I. C. B., Custódio, E. M., Duarte, W. F., & Duarte, J. L. M. (1999). *Matrizes Progressivas Coloridas de Raven: Escala Especial*. CETEPP.
- APA (2013). *Diagnostic and statistical manual of mental disorders: DSM-5* (5th ed.). American Psychiatric Association.
- Athayde, M. L., Giacomoni, C. H., Zanon, C., & Stein, L. M. (2014). Evidências de validade do subteste de leitura do teste de desempenho escolar. *Revista Psicologia: Teoria e Prática*, 16(2), 131-140.
- Capellini, S. A., Oliveira, A. M., & Cuetos, F. (2012). *PROLEC: Provas de Avaliação dos Processos de Leitura* (2. ed.). Casa do Psicólogo.

- Carreiras, M., Mechelli, A., & Price, C. J. (2006). Effect of word and syllable frequency on activation during lexical decision and reading aloud. *Hum Brain Mapp*, 27, 963-972. <https://doi.org/10.1002/hbm.20236>
- Cogo-Moreira, H., Ploubidis, G., De Avila, C., Mari, J., & Pinheiro, A. M. V. (2012). EACOL (Scale of Evaluation of Reading Competency by the Teacher): Evidence of concurrent and discriminant validity. *Neuropsych. Dises. Treat.*, 8(1), 443–454.
- Coltheart, M., Rastle, K., Perry, C., Langdon, R., & Ziegler, J. (2001). DRC: a dual route cascaded model of visual word recognition and reading aloud. *Psychological Review*, 108(1), 204–256. <https://doi.org/10.1037/0033-295X.108.1.204>
- Cossu, G., Giuliotta, M., & Marshall, J. C. (1995). Acquisition of reading and written spelling in transparent orthography. Two non-parallel processes? *Reading and Writing*, 7, 9–22. <https://doi.org/10.1007/BF01026945>
- de Jong, P. F., & van der Leij, A. (2003). Developmental changes in the manifestation of a phonological deficit in dyslexic children learning to read a regular orthography. *Journal of Educational Psychology*, 95(1), 22–40.
- Fleitlich-Bilyk, B., Cortázar, P. G., & Goodman, R. (2000). Questionário de Capacidades e Dificuldades (SDQ) [Strengths and Difficulties Questionnaire (SDQ)]. *Infanto: Revista de Neuropsiquiatria da Infância e Adolescência*, 8(1), 44-50.
- Justi, C. N. G., & Justi, F. R. R. (2009). Os efeitos de lexicalidade, frequência e regularidade na leitura de crianças falantes do português brasileiro. *Psicologia: Reflexão e Crítica*, 22(2), 163-172.
- Landerl, K., & Wimmer, H. (2008). Development of Word Reading Fluency and Spelling in a Consistent Orthography: An 8-Year Follow-Up. *Journal of Educational Psychology*, 100(1), 150-161. <https://doi.org/10.1037/0022-0663.100.1.150>

- Lúcio, P. S., Moreira, H. C., Kida, A. S. B., Carvalho, C. A. F., Pinheiro, Â. M., Mari, J. J., & Avila, C. R. B. (2018). Word Decoding Task: Item Analysis by IRT and Within-Group Norms. *Psicologia: Teoria e Pesquisa*, 34.
- Lúcio, P. S., Pinheiro, Â. 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.
- Marinelli, C. V., Romani, C., Burani, C., McGowan, V. A., & Zoccolotti, P. (2016). Costs and benefits of orthographic inconsistency in reading: evidence from a cross-linguistic comparison. *PLoS One*, 11(6), e0157457.
- OMS. (2022). *CID-11: Classificação estatística internacional de doenças e problemas relacionados à saúde* (11ª ed.). Genebra: OMS.
- Pasquali, L., Wechsler, S. M., & Bensusan, E. (2002). Matrizes Progressivas do Raven Infantil: um estudo de validação para o Brasil. *Avaliação Psicológica*, 1(2), 95–110.
- Pinheiro, Â. M. V. (1995). Reading and spelling development in Brazilian Portuguese. *Reading & Writing*, 7(1), 111-138.
- Pinheiro, Â. M. V. (1999). Cognitive assessment of competent and impaired reading in Scottish and Brazilian children. *Reading & Writing*, 11, 175-211.
- Pinheiro, Â. M. V. (2013). *Prova de Leitura e de Escrita de palavras e de pseudopalavras*. Relatório Técnico Final. FAPEMIG. Nº: APQ-01914-09.
- Pinheiro, Â. M. V. (2015). Frequency of Occurrence of Words in Textbooks Exposed to Brazilian children in the Early Years of Elementary School. Updated database from 1996. *CHILDES*. <http://chilDES.talkbank.org/derived>
- Pinheiro, Â. M. V., & Costa, A. E. B. (2015). EACOL – Escala de Avaliação da Competência em Leitura Pelo Professor: Construção por meio de Critérios e de Concordância entre Juízes. *Psicologia: Reflexão e Crítica*, 28(1), 1–10.

- Pinheiro, Â. M. V., & Vilhena, D. A. (2022). Teste de reconhecimento de palavras e pseudopalavras: validades de conteúdo e externa. *Signo*, 47(88), 147-164.
- Pinheiro, Â. M. V., & Vilhena, D. A. (2023). Teste de Reconhecimento de Palavras e Teste de Reconhecimento de Pseudopalavras: *Manual Técnico*. NILAPRESS Editora.
- Pinheiro, Â. M. V., Vilhena, D. A., & Santos, M. A. C. (2017). PROLEC – Provas de Avaliação dos Processos de Leitura: análise de suas características psicométricas. *Trends in Psychology [Temas em Psicologia]*, 25(3), 1067-1080.
- Provazza, S., Giofrè, D., Adams, A. M., & Roberts, D. J. (2019). The Clock Counts - Length Effects in English Dyslexic Readers. *Frontiers in psychology*, 10, 2495.
- Salles, J. F., Piccolo, L. R., Zamo, R. S., & Toazza, R. (2013). Normas de desempenho em tarefa de leitura de palavras/pseudopalavras isoladas (LPI) para crianças de 1º ano a 7º ano. *Estudos e Pesquisas em Psicologia*, 13(2), 397-419.
- Saur, A. M., & Loureiro, S. R. (2012). Psychometric Properties of the Strengths and Difficulties Questionnaire: a literature review. *Estudos de Psicologia*, 29(4), 619–629.
- Schuster, S., Hawelka, S., Hutzler, F., Kronbichler, M., & Richlan, F. (2016). Words in Context: The Effects of Length, Frequency, and Predictability on Brain Responses During Natural Reading. *Cerebral cortex*, 26(10), 3889–3904.
- Stivanin, L., & Scheuer, C. I. (2007). Tempo de latência para a leitura: influência da frequência da palavra escrita e da escolarização. *Rev Soc Bras Fono*, 12(3), 206-13.
- Vilhena, D. A., & Pinheiro, A. M. V. (2016). Revised version of the Scale of Evaluation of Reading Competence by the Teacher: final validation and standardization. *Universitas Psychologica*, 15(4), 1-13.
- Vilhena, D. A., & Pinheiro, Â. M. V. (2020). Fidedignidade, Validade e Normatização do Teste de Leitura: Compreensão de Sentenças. *Psic: Teoria e Pesquisa*, 36, e36325.

- Vilhena, D. A., & Pinheiro, Â. M. V. (2022a). *Teste de Leitura: Compreensão de Sentenças – TELCS: Manual Técnico*. NILAPRESS Editora.
- Vilhena, D. A., & Pinheiro, Â. M. V. (2022b). Fidedignidade, validade interna e normatização dos testes de reconhecimento de palavras e de pseudopalavras. *Neuropsicologia Latinoamericana*, 14(2), 1–10.
- Vilhena, D. A., Sucena, A., Castro, S. L., & Pinheiro, Â. M. V. (2016). Reading Test-Sentence Comprehension: An Adapted Version of Lobrot's Lecture 3 Test for Brazilian Portuguese. *Dyslexia*, 22(1), 47-63. <http://dx.doi.org/10.1002/dys.1521>
- Wimmer, H. (1993). Characteristics of developmental dyslexia in a regular writing system. *Applied Psycholinguistics*, 14(1), 1-33.
- Wimmer, H., Schurz, M., Sturm, D., Richlan, F., Klackl, J., Kronbichler, M., & Ladurner, G. (2010). A dual-route perspective on poor reading in a regular orthography: An fMRI study. *Cortex*, 46(10), 1284-1298.

## 8. Memorial

Os estudos psicométricos do Teste de Reconhecimento de Palavras e do Teste de Reconhecimento de Pseudopalavras foram realizados em parte com o apoio da Fundação de Amparo à Pesquisa do Estado de Minas Gerais (FAPEMIG, processo APQ-01914-09), do Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq, processo 134357/2013-2) e da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Código de Financiamento 001. Destacamos as anuências institucionais e os apoios logísticos dados pela Superintendência Regional de Educação Metropolitana A da Secretaria de Estado de Educação de Minas Gerais e pelas diretorias das escolas.