

Fidedignidade, validade interna e externa da Prova de Reconhecimento de Palavras – Versão Brasileira

Confiabilidad, validez interna y externa de la Prueba de Reconocimiento de Palabras - Versión Brasileña
Fiabilité, validité interne et externe de l'Épreuve de Reconnaissance de Mots - Version Brésilienne
Reliability, internal and external validity of the Word Recognition Assessment – Brazilian Version

Douglas de Araújo Vilhena¹ e Ângela Maria Vieira Pinheiro¹

1. *Programa de Pós-Graduação em Psicologia: Cognição e Comportamento, Faculdade de Filosofia e Ciências Humanas, Universidade Federal de Minas Gerais, Brasil.*

Resumo

Objetivou-se prover fontes de evidências psicométricas de fidedignidade, validade da estrutura interna e validade externa para a Prova de Reconhecimento de Palavras – Versão Brasileira (PRP-BR). Participaram do estudo piloto da Fase 2, planejada para avaliar a efetiva iconicidade dos desenhos, 114 alunos do segundo ao quinto ano do Ensino Fundamental, sem tempo limite. Na Fase 3, participaram 480 alunos do segundo ao quinto ano, com a aplicação da PRP-BR completa (52 itens) sob tempo limite. Compuseram a Fase 4 os 231 participantes que responderam todos os itens. Na Fase 5, 245 participantes do primeiro ao quinto ano responderam a PRP-BR reduzida (43 itens) sob tempo limite. Aplicou-se a PRP-BR de forma individual nos participantes do primeiro ano e de forma coletiva nos do segundo ao quinto ano. O tempo limite de aplicação para o primeiro ao terceiro ano foi de quatro minutos e para o quarto e quinto ano foi de dois minutos. Na Fase 2, todos os itens atingiram o critério de corte de 80% de acerto, o que sugere suas manutenções para os subsequentes estudos de validação. Ao considerar os resultados da Teoria Clássica dos Testes e da Teoria de Resposta ao Item, foram retirados nove dos 52 itens. Como evidência de adequada fidedignidade, o alfa de Cronbach aumentou progressivamente entre as fases deste estudo. Como evidências de validade interna, verificou-se um significativo efeito de escolaridade tanto com a versão completa quanto a reduzida em todas as amostras. Apresentou-se evidências de validade externa convergente e de critério, com os resultados estando coerentes com a literatura. Conclui-se que a PRP-BR apresentou satisfatórias fontes de evidências psicométricas de fidedignidade, validade da estrutura interna e validade externa.

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

Resumen

El objetivo fue proporcionar fuentes de evidencia psicométrica de confiabilidad, validez de estructura interna y validez externa para la Prueba de Reconocimiento de Palabras - Versión Brasileña (PRP-BR). El estudio piloto de la Fase 2, diseñado para evaluar la iconicidad efectiva de los diseños, participó 114 niños de segundo a quinto año de Educación Primaria, sin límite de tiempo. En la Fase 3, participaron 480 niños de segundo a quinto año, con la aplicación de la PRP-BR completa (52 ítems) con límite de tiempo. La Fase 4 comprendió los 231 participantes que respondieron a todos los ítems. En la Fase 5, 245 participantes del primero al quinto año respondieron la PRP-BR reducida (43 ítems) bajo límite de tiempo. La aplicación fue individualmente a los participantes del primer año y de forma colectiva a los del segundo al quinto año. El tiempo límite de aplicación para el primero al tercer año fue de cuatro minutos y para el cuarto y quinto año fue de dos minutos. En la Fase 2, todos los ítems alcanzaron el criterio de corte del 80% correcto, lo que sugiere su mantenimiento para posteriores estudios de validación. Después de considerar los resultados de la Teoría Clásica de Test y la Teoría de Respuesta al Ítem, nueve de los 52 ítems fueron eliminados. Como evidencia de confiabilidad adecuada, el alfa de Cronbach aumentó progresivamente entre las fases de este estudio. Como evidencia de validez interna, hubo un efecto significativo de la escolaridad tanto con la versión completa como con la reducida en todas las fases. Se presentaron evidencias de validez externa convergente y de criterio, siendo los resultados consistentes con la literatura. Se concluye que la PRP-BR presentó fuentes satisfactorias de evidencia psicométrica de confiabilidad, validez de estructura interna y validez externa.

Palabras clave: evaluación educativa, dificultades de aprendizaje, dislexia, estudio de validación, psicometría.

Artigo recebido: 26/04/2023; Artigo aceito: 28/12/2023.

Correspondências relacionadas a esse artigo devem ser enviadas a Ângela Maria Vieira Pinheiro, Universidade Federal de Minas Gerais, Faculdade de Filosofia e Ciências Humanas –Av. Pres. Antônio Carlos, 6627 -Pampulha, Belo Horizonte–Minas Gerais, MG –Brasil, CEP 31270-901.

E-mail: pinheiroamva@gmail.com

DOI:10.5579/rml.2023.0828

Résumé

L'objectif était de fournir des sources de preuves psychométriques de fiabilité, de validité de structure interne et de validité externe pour le preuve de reconnaissance de mots - version brésilienne (PRP-BR). L'étude pilote de la Phase 2, conçue pour évaluer l'iconicité effective des dessins, participé 114 étudiants de la deuxième à la cinquième année du primaire, sans limite de temps. En Phase 3, 480 étudiants de la deuxième à la cinquième année ont participé, avec l'application du PRP-BR complet (52 items) dans un délai limité. La Phase 4 comprenait les 231 participants qui ont répondu à tous les items. En Phase 5, 245 participants de la première à la cinquième année ont répondu au PRP-BR réduit (43 items) dans un délai limité. La PRP-BR a été appliqué individuellement aux participants de la première année et collectivement à ceux de la deuxième à la cinquième année. Le délai de candidature pour les première à troisième années était de quatre minutes et pour les quatrième et cinquième années, il était de deux minutes. Dans la Phase 2, tous les items ont atteint le critère de 80 % corrects, ce qui suggère leur maintien pour les études de validation ultérieures. Lors de l'examen des résultats de la Théorie Classique des Tests et de la Théorie de la Réponse aux Items, neuf des 52 items ont été supprimés. Comme preuve d'une fiabilité adéquate, l'alpha de Cronbach a progressivement augmenté entre les phases de cette étude. Comme preuve de la validité interne, il y avait un effet significatif de la scolarisation avec la version complète et la version réduite dans toutes les phases. Des preuves de validité externe convergente et de critère ont été présentées, les résultats étant cohérents avec la littérature. Il est conclu que le PRP-BR a présenté satisfaisantes psychométriques sources de fiabilité, de validité de structure interne et de validité externe.

Mots-clés: troubles d'apprentissage, dyslexie, étude de validation, lecture, psychométrie.

Abstract

The objective was to provide psychometric evidence of reliability, internal structure validity and external validity for the Word Recognition Assessment – Brazilian Version (PRP-BR). The pilot study of Phase 2, designed to evaluate the iconicity of the illustrations, participated 114 children from the second to the fifth grades of Elementary School, with no time limit. In Phase 3, 480 students from the second to the fifth grades participated, with the application of the complete PRP-BR (52 items) under a time limit. Phase 4 comprised the 231 participants who responded to all items. In Phase 5, 245 participants from the first to the fifth grades answered the reduced PRP-BR (43 items) under a time limit. The PRP-BR was applied individually to participants in the first grade and collectively to those from the second to the fifth grades. The application time limit for the first to third grades was four minutes and for the fourth and fifth grades it was two minutes. In Phase 2, all items reached the cutoff criterion of 80% accuracy, which suggested their maintenance for subsequent validation studies. Considering the results of the Classical Test Theory and the Item Response Theory, nine of the 52 items were removed. As evidence of adequate reliability, Cronbach's alpha progressively increased between the phases of this study. As evidence of internal validity, there was a significant effect of schooling both with the complete and the reduced version in all samples. Evidence of convergent and criterion external validity were presented, with the results being consistent with the literature. In concluded, the PRP-BR presented satisfactory psychometric evidence of reliability, internal structure validity and external validity.

Keywords: dyslexia, educational measurement, learning disabilities, validation study, psychometrics.

1. INTRODUÇÃO

É um dever urgente de todos e prioridade absoluta do Estado, da Escola, da sociedade e das famílias garantir a alfabetização para todos, com especial atenção para os alunos com dificuldade de leitura e os com dislexia do desenvolvimento, sem deixar ninguém para trás. O rastreio universal, precoce e periódico dos alunos do Ensino Fundamental é importante para assegurar que sejam identificados aqueles que precisam de intervenções educacionais para recuperar a defasagem (Komesidou et al., 2022).

Viana e Ribeiro (2007, 2010), na Universidade do Minho em Portugal, elaboraram a Prova de Reconhecimento de Palavras – Versão Portuguesa (PRP-PT) com o objetivo de instrumentalizar profissionais da área de educação e psicólogos, de forma fidedigna e válida, para avaliar a habilidade de reconhecimento de palavras de crianças. As avaliações podem ser individuais ou coletivas, em contexto clínico ou escolar, para os quatro primeiros anos de escolaridade. A PRP-PT é amplamente utilizada no português europeu como uma prova de *screening*, com aplicação em larga escala, para rastrear precocemente os alunos com dificuldade de leitura. A aplicação é simples, devendo o participante ver um desenho à esquerda e reconhecer a palavra correspondente dentre quatro alternativas. Deve-se responder a PRP-PT em silêncio, sob limite de tempo, que pode ser de quatro minutos (1º e 2º ano) ou dois minutos (3º e 4º ano).

A construção da PRP-PT considerou o controle das características psicolinguísticas de frequência de ocorrência,

regularidade, extensão e vizinhança ortográfica (Viana et al., 2013). A PRP-PT foi construída com base no conceito de vizinhança ortográfica, com as alternativas distratoras apresentando a mesma extensão da palavra-alvo e todas as letras nas suas posições sequenciais respectivas, à exceção de uma letra [ex., livro e litro; porco e porto] (Grainger et al., 1989; Paap & Johansen, 1994; Rumelhart & MacClelland, 1982). A facilidade ou dificuldade para reconhecer corretamente a palavra-alvo correspondente ao desenho dependerá da força das inibições exercidas pelas outras palavras distratoras, particularmente pelas vizinhas ortográficas com maior frequência de ocorrência.

Viana et al. (2013) apresentaram as evidências psicométricas de fidedignidade e validade da PRP-PT por meio de dois estudos com amostras distintas em Portugal. No primeiro estudo, participaram 328 alunos do 1º ao 4º ano do primeiro Ciclo do Ensino Básico (59% do sexo masculino) de escolas públicas no norte de Portugal, com a aplicação coletiva da PRP-PT, sem limite de tempo. Já no segundo estudo, participaram 3.131 alunos (51% do sexo masculino) do 1º ($n = 402$), 2º ($n = 932$), 3º ($n = 943$) e 4º ano ($n = 854$) de escolas urbanas e rurais na zona norte, centro e sul de Portugal.

Estudos de adaptação transcultural justificam-se pela importância do Brasil e Portugal, assim como dos demais países de língua portuguesa, possuem instrumentos equivalentes e com evidências psicométricas satisfatórias para avaliar diferentes aspectos cognitivos envolvidos na leitura. A criação de provas para avaliar a leitura de crianças do ensino fundamental é importante para rastrear crianças com dificuldades de leitura, o que permite o encaminhamento

clínico para que essas dificuldades sejam diagnosticadas precocemente e tratadas antes que se tornem mais graves. Também é importante uma prática baseada em evidência sustentada por parâmetros objetivos para avaliar o progresso do aprendizado da leitura. A melhoria do desempenho leitor está diretamente ligada ao bom desempenho escolar, pois a leitura é fundamental para a compreensão de outras áreas do conhecimento, como matemática, ciências e história.

Pinheiro et al. (*in press*) apresentaram a Fase 1 e a Fase 2 do estudo de validação de conteúdo da Prova de Reconhecimento de Palavras – Versão Brasileira (PRP-BR). A Fase 1 consistiu na adaptação transcultural dos itens originais da PRP-PT para a realidade de alunos do Ensino Fundamental brasileiro. Avaliou-se a pertinência dos desenhos ao contexto dos alunos brasileiros, as características psicolinguísticas das palavras-alvo e as alternativas distratoras. As características psicolinguísticas de cada palavra-alvo foram minuciosamente detalhadas para controlar os efeitos de vizinhança ortográfica, frequência de ocorrência, regularidade e de extensão. A Fase 1 resultou na atualização de três ilustrações, exclusão de dois itens e substituição de 24 alternativas dos 40 itens originais da PRP-PT. Com a adição de 14 novas ilustrações, a PRP-BR totalizou 52 desenhos.

Participaram da Fase 2, planejada como estudo piloto para avaliação da iconicidade dos 52 desenhos selecionados na Fase 1, alunos do segundo ao quinto ano do Ensino Fundamental, sem limite de tempo de aplicação. Verificou-se empiricamente que todos os 52 itens atingiram o critério de corte de 80% de acurácia ($M = 98,2 \pm 2,7\%$; $Min = 86,5\%$; $Max = 100\%$), o que sugere a manutenção de todos os itens para os subsequentes estudos psicométricos (Pinheiro et al., *in press*). Dessa forma, a PRP-BR demonstrou adequadas fontes de evidências de validade de conteúdo, com equivalência transcultural e coerência com a fundamentação teórica.

O objetivo do presente estudo foi prover fontes de evidências psicométricas de fidedignidade, validade da estrutura interna e validade externa para a PRP-BR. A partir da amostra da Fase 2, com aplicação sem tempo limite, investigou-se o efeito de escolaridade na acurácia de leitura. Na Fase 3, a PRP-BR foi aplicada com tempo limite para avaliar a fluência de leitura e o efeito de escolaridade. Na Fase 4, analisou-se a fidedignidade dos participantes da Fase 2 e Fase 3 que responderam todos os 52 itens, uma vez que o alpha de Cronbach não pode ser calculado com dados faltantes. Por fim, na Fase 5 aplicou-se a PRP-BR reduzida com tempo limite.

2. MÉTODO

2.1 Participantes

Delineamento transversal para estudo da fidedignidade, da validade da estrutura interna e validade externa da PRP-BR. A Tabela 1 apresenta a frequência de participantes por ano escolar, por escola e por regional na cidade de Belo Horizonte, capital de Minas Gerais, Brasil. A Tabela 2 apresenta as características demográficas dos participantes (sexo e idade), a versão da PRP-BR (número de itens, tempo e modo de aplicação) e o índice de fidedignidade alpha de Cronbach.

Participaram da Fase 2, planejado como estudo piloto para avaliar a efetiva iconicidade dos desenhos, alunos do

segundo ao quinto ano do Ensino Fundamental ($n = 114$; 61% do sexo feminino; faixa etária = 7 a 11 anos; $M = 10,5 \pm 0,6$ anos de idade), de três escolas (duas Estaduais Públicas e uma Particular), em Belo Horizonte, ao final do ano letivo de 2008. Participaram da Fase 3, com a PRP-BR completa aplicada com tempo limite, alunos do segundo ao quinto ano do Ensino Fundamental ($n = 480$; 50,8% do sexo feminino; faixa etária = 6 a 13 anos; $M = 9,4 \pm 1,2$ anos), de oito escolas (três Estaduais e cinco Particulares) em Belo Horizonte, ao final do ano letivo de 2009.

Compuseram o estudo de fidedignidade da Fase 4 todos os 114 participantes da Fase 2 e mais os sujeitos da Fase 3 que responderam todos os 52 itens da PRP-BR dentro do tempo ($n = 197$ participantes), totalizando 231 participantes (57% do sexo feminino; faixa etária = 7 a 12 anos; $M = 9,1 \pm 1,3$ anos). Na Fase 5, participaram do estudo com a PRP-BR reduzida (43 itens) alunos do primeiro ao quinto ano do Ensino Fundamental ($n = 245$; 45% do sexo feminino; faixa etária = 6 a 12 anos; $M = 8,7 \pm 1,5$ anos) de dez escolas (três Estaduais, quatro Municipais e três Particulares) de Belo Horizonte, ao final do ano letivo de 2011.

2.2 Instrumentos

A Prova de Reconhecimento de Palavras – Versão Brasileira (PRP-BR), assim como a versão original (PRP-PT), são instrumentos de avaliação da leitura silenciosa destinados às crianças do Ensino Fundamental (Pinheiro et al., 2024). Tanto a versão da PRP-BR completa com 52 itens (aplicada na Fase 2 e Fase 3) quanto a reduzida com 43 itens (aplicada na Fase 5) são fidedignas e válidas para avaliar a acurácia (precisão) e a fluência (velocidade) do construto ‘reconhecimento de palavras’. Cada item é composto por um desenho à esquerda seguido de quatro palavras como alternativas (isoladas, reais, letra minúscula) (Figura 1). Apenas uma das alternativas é a resposta correta (palavra-alvo) e as outras três alternativas são distratoras. A PRP-BR foi apresentada no formato caneta e papel, impressa frente e verso em três folhas de tamanho A4. Os desenhos e as palavras foram dispostos dentro de uma tabela, com bordas pretas, com igual distribuição da dimensão vertical (2 cm) e horizontal (3 cm).

A PRP-BR completa (impressa em seis páginas) e a reduzida (impressa em cinco páginas) possuem características psicolinguísticas equivalentes. As palavras-alvo de ambas as versões apresentam de quatro a oito letras ($M = 5,6 \pm 1,1$ letras). Na PRP-BR completa (52 itens), metade dos desenhos remetem para palavras com duas sílabas e a outra metade para três sílabas. 37 palavras-alvo possuem regularidade grafema-fonema previsíveis (71%) e 15 imprevisíveis (29%). 15 itens possuem todos os três distratores com vizinhança ortográfica; 21 itens possuem dois vizinhos ortográficos e um aproximado; 15 itens com um vizinho ortográfico e dois aproximados; e um item sem vizinho ortográfico e três aproximados.

Tabela 1.

Frequência e porcentagem de participantes nos estudos da Fase 2 (n = 114), Fase 3 (n = 480), Fase 4 (n = 231) e Fase 5 (n = 245) por escola, por regional e por ano escolar em Belo Horizonte, MG, Brasil

Estudo	Escola	Regional	Ano escolar					Total	
			1°	2°	3°	4°	5°	n	%
Fase 2	Estadual 1	Noroeste	0	2	8	6	4	20	18
	Estadual 2	Pampulha	0	8	10	10	10	38	33
	Particular 1	Pampulha	0	12	20	13	11	56	49
	Total	n	0	22	38	29	25	114	100
		%	0	19	33	25	22		100
Fase 3	Estadual 1	Pampulha	0	12	28	19	26	85	18
	Estadual 2	Leste	0	10	14	25	29	78	16
	Estadual 3	Noroeste	0	4	8	15	15	42	9
	Particular 1	Pampulha	0	12	24	22	14	72	15
	Particular 2	Pampulha	0	7	0	0	0	7	1
	Particular 3	Leste	0	12	20	17	26	75	16
	Particular 4	Leste	0	28	0	19	27	74	15
	Particular 5	Leste	0	10	10	17	10	47	10
	Total	n	0	95	104	134	147	480	100
	%	0	20	22	28	31		100	
Fase 4	Total	n	0	42	86	31	72	231	100
		%	0	18	37	13	31		100
Fase 5	Estadual 1	Pampulha	5	2	13	8	0	28	11
	Estadual 2	Noroeste	9	7	5	0	0	21	9
	Estadual 3	Nordeste	10	11	16	8	0	45	18
	Municipal 1	Centro-Sul	5	6	9	7	7	34	14
	Municipal 2	Oeste	0	0	0	11	4	15	6
	Municipal 3	Venda Nova	1	4	9	7	5	26	11
	Municipal 4	Leste	10	12	9	11	0	42	17
	Particular 1	Pampulha	0	5	0	4	0	9	4
	Particular 2	Oeste	2	3	3	4	0	12	5
	Particular 3	Centro-Sul	0	2	7	3	1	13	5
	Total	n	42	52	71	63	17	245	100
		%	17	21	29	26	7		100



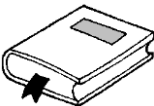
Tabela 2.

Características demográficas dos participantes (sexo e idade), versão da PRP-BR (número de itens, tempo e modo de aplicação) e índice de fidedignidade (α)

Variável descritiva	Fase 2: Iconicidade	Fase 3: Completa	Fase 4: Fidedignidade	Fase 5: Reduzida
Participantes (n)	114	480	231	245
Sexo feminino (%)	60,5	50,8	57,1	44,9
Faixa etária (anos)	6 a 11	6 a 13	7 a 12	6 a 12
Anos de Idade (M ± SD)	10,5 ± 0,6	9,4 ± 1,2	9,1 ± 1,3	8,7 ± 1,5
Número de itens	52	52	52 e 43	43
Ano escolar: tempo de aplicação (min)	2° ao 5° ano: sem limite	2° e 3° ano: 4 4° e 5° ano: 2	2° ao 5° ano	1° e 3° ano: 4 4° e 5° ano: 2
Aplicação	Coletiva	Coletiva	Fase 2 + Fase 3	1° ano: Individual 2° ao 5°: Coletivo
Alpha de Cronbach (α)	0,55	0,61	52 itens: 0,70	43 itens: 0,71
Acurácia (%)	Média	98,2	64,2	54,5
	Desvio Padrão	2,7	23,9	28,2
	Mínimo	86,5	12,0	2,3
	Máximo	100	100	100

Figura 1.

Exemplos dos três itens de treino da PRP-BR

a)		pombal	pompa	pomba	lomba
b)		potro	porto	tosco	porco
c)		litro	livro	libra	livre

Nota. Figura 1a: POMBA; Figura 1b: PORCO; Figura 1c: LIVRO.

Com relação à frequência de ocorrências em livros didáticos escolares, as 52 palavras-alvo possuem no mínimo cinco (*i.e.*, sereia) e no máximo 1544 ocorrências por milhão (*i.e.*, bola), sendo 19 palavras-alvo classificadas como de alta frequência, 21 como de média frequência e seis como de baixa frequência ($M = 255 \pm 289$ ocorrências por milhão) (Pinheiro, 2015). Em apenas oito dos 52 itens, ou 12 das 156 alternativas, as palavras-alvo apresentaram uma frequência menor do que pelo menos uma das alternativas (*ex.*, ‘Ponte’ possui menor frequência do que as alternativas ‘Fonte’ e ‘Ponta’).

Na PRP-BR reduzida (43 itens), 19 desenhos remetem para palavras com duas sílabas e 24 para três sílabas. Quanto a regularidade grafema-fonema, 31 palavras-alvo possuem correspondências previsíveis (72%) e 12 imprevisíveis (28%). 13 itens possuem todos os três distratores com vizinhança ortográfica; 17 itens possuem dois vizinhos ortográficos e um aproximado; 12 itens com um vizinho ortográfico e dois aproximados; e um item sem vizinho ortográfico e três aproximados. As palavras-alvo possuem de 24 a 1037 ocorrências por milhão, sendo 13 palavras-alvo classificadas como de alta frequência, 19 como de média frequência e 11 como de baixa frequência ($M = 217 \pm 227$ ocorrências). Em oito dos 43 itens, ou 12 das 129 alternativas, a palavra-alvo apresentou frequência menor do que pelo menos uma das alternativas.

Para a correção do escore bruto da PRP-BR, é atribuído um ponto por cada item respondido corretamente e zero para os com erro ou omitidos. O escore bruto é convertido em: (1) Acurácia: porcentagem de itens respondidos corretamente (*ex.*, 40 itens corretos correspondem à acurácia de 77% na versão com 52 itens e 93% na versão com 43 itens); e (2) Taxa de Acurácia: número total de itens respondidos corretamente, multiplicado por sessenta e dividido pelo tempo total em segundos (*ex.*, 36 itens corretos em 120 segundos correspondem à Taxa de Acurácia de 18 itens por minuto).

2.3 Procedimentos

Os procedimentos da Fase 1 para estudo de iconicidade foram apresentados separadamente (Pinheiro et al., *in press*). Em suma, o estudo de validação de conteúdo da

Fase 1 consistiu em: (1) averiguar a iconicidade dos desenhos para a realidade de alunos do Ensino Fundamental brasileiro e (2) realizar a seleção das alternativas com base na vizinhança ortográfica em relação à palavra-alvo.

A seleção das escolas foi feita com base nos dados fornecidos pela Secretaria de Educação do Estado de Minas Gerais. As amostragens das escolas da Fase 2 e da Fase 3 deram-se por conveniência de deslocamento dos aplicadores da pesquisa. Na Fase 5, a amostragem foi estratificada em seis regionais da cidade. Os participantes que compuseram as amostras foram sorteados, a partir da lista de presença, em diferentes turmas de cada ano escolar. Todas as escolas selecionadas concordaram em participar da pesquisa por meio da assinatura do termo de anuência institucional.

A PRP-BR foi aplicada de forma individual nos participantes do primeiro ano e de forma coletiva nos do segundo ao quinto ano do Ensino Fundamental. Permitiu-se apenas um participante por mesa. Orientou-se que seria realizada uma prova de leitura fácil e em silêncio. Após a distribuição da folha de resposta, os participantes foram orientados para que preenchessem com uma caneta o cabeçalho e para não virar a página. Iniciou-se a PRP-BR com os três exemplos da primeira página (Figura 1). A tarefa é olhar para o desenho e reconhecer a palavra correspondente dentre as quatro alternativas ao lado. À esquerda no primeiro exemplo, há um desenho de uma POMBA (Figura 1a). Orientou-se a marcar à direita uma cruz em cima onde está escrito POMBA. No segundo desenho, um PORCO, deve-se cruzar a palavra PORCO (Figura 1b). No terceiro exemplo, um LIVRO, marca-se a palavra LIVRO (Figura 1c). Orientou-se que não é relevante que os traços estivessem perfeitos, pois só é preciso indicar a palavra que correspondesse ao desenho. Exemplificou-se no quadro como se faz uma cruz com rapidez.

No prévio estudo de iconicidade da Fase 2, a PRB-BR foi aplicada sem tempo limite. Já na Fase 3 (mesma aplicação da Fase 4) e Fase 5, para avaliar a fluência, o tempo limite de aplicação para o 1º ao 3º ano do Ensino Fundamental foi de quatro minutos e para o 4º e 5º ano foi de dois minutos. Nas aplicações cronometradas, orientou-se que é preciso trabalhar depressa, uma vez que há um tempo limite para responder a prova. Por isso, não foi permitido conversar e nem fazer perguntas depois que a prova começasse. Após dado início, orientou-se para que todos virassem a página e comessem a responder o mais rápido possível, sem esquecer dos dois lados das três folhas da prova.

Todos os participantes forneceram consentimento informado e o Comitê de Ética em Pesquisa da Universidade Federal de Minas Gerais aprovou todos os procedimentos do estudo (Processo nº ETIC 347/04). O projeto de pesquisa foi realizado em total conformidade com a Resolução nº 466/12 do Conselho Nacional de Saúde e do Código de Ética da *World Medical Association* (2008) para pesquisas envolvendo seres humanos.

2.4 Análise de dados

As análises foram realizadas usando o *software* SPSS (IBM, Chicago, Illinois). Na análise da distribuição populacional, os valores de assimetria e curtose foram divididos pelos respectivos erros padrão, utilizando um critério de significância superior a 1,96 (Cramer & Howitt,

2004). Todas as amostragens, da Fase 2 a Fase 5, não apresentaram uma distribuição populacional normal, com assimetria negativa concentrada próxima ao máximo (*i.e.*, 100% de acerto) e uma curtose com achatamento leptocúrtico ($p < 0,001$). Portanto, utilizou-se testes não-paramétricos (*i.e.*, Kruskal-Wallis, Spearman).

A análise da qualidade do item considerou os elementos da Teoria Clássica dos Testes (Lord & Novick, 1968). Para estimar a fidedignidade, utilizou-se o alfa de Cronbach para toda a escala e por ano escolar, as correlações item-total corrigidas e o alfa de Cronbach se o item for excluído. Utilizou-se o seguinte critério para o alfa de Cronbach: excelente $\geq 0,90$; bom $\geq 0,80$; aceitável $\geq 0,70$; questionável $\geq 0,60$; pobre $\geq 0,50$; inaceitável $< 0,50$ (George & Mallery, 2003). Considerou-se aceitável a correlação item-total corrigida superior a 0,2 (Kline, 1993).

As características dos itens foram avaliadas por meio do modelo de três parâmetros da Teoria de Resposta ao Item. Espera-se que os itens apresentem índice de discriminação (parâmetros a) acima de 0,3. O índice de probabilidade de acerto ao acaso (parâmetro c) é calculado em uma escala percentual e o ideal é que seus valores sejam baixos. A significância estatística foi estabelecida em $p < 0,05$.

3. RESULTADOS

3.1 Fase 2 – Estudo piloto de iconicidade

Na Fase 2, verificou-se um significativo efeito de escolaridade na variável acurácia da PRP-BR ($H_{(3, 111)} = 13,7$, $p = 0,003$), mesmo sendo uma aplicação com tempo livre para os participantes responderem todos os 52 itens. A análise *post hoc* revelou que os participantes do segundo ano obtiveram uma menor acurácia quando comparados aos do quarto e quinto ano do Ensino Fundamental (*i.e.*, 2° < 4°–5° ano) ($U_{(1, 113)} = 1413,5$, $p < 0,001$), sem diferença dos participantes do terceiro ano com os demais anos escolares. Como validade externa de critério, não houve diferença significativa entre o desempenho dos participantes das escolas públicas com os da escola particular e nem entre meninos e meninas.

3.2 Fase 3 – Validade da PRP-BR completa (52 itens)

Na Fase 3 ($n = 480$ participantes), com aplicação da PRP-BR com 52 itens e tempo limite, a variável acurácia apresentou uma média geral de $64,2 \pm 23,9\%$, com mínimo de 11,3% e máximo 100% (Tabela 4). Verificou-se, por meio da estatística não paramétrica Kruskal-Wallis, um significativo efeito de escolaridade na variável acurácia ($H_{(3, 477)} = 70,6$; $p < 0,001$) e na taxa de acurácia ($H_{(3, 477)} = 144,9$; $p < 0,001$).

Ao comparar os anos escolares com o mesmo tempo limite de aplicação, os participantes do 2° ano apresentaram uma menor acurácia do que os do 3° ano ($U_{(1,198)} = 6,4$; $p < 0,001$); e os participantes do 4° ano uma menor acurácia do que os do 5° ano ($U_{(1,280)} = 13,1$; $p < 0,001$). O dobro de tempo de aplicação para os participantes do 2° e do 3° ano (quatro minutos), quando comparado ao 4° e 5° ano (dois minutos), resultou em uma maior acurácia ($U_{(1, 479)} = 3,2$, $p = 0,0006$). A correlação de Spearman calculada entre o ano escolar e o percentual de acerto foi igual a 0,243 ($p < 0,001$).

A análise de regressão linear exploratória observou que para cada ano de escolaridade completo é esperado um aumento de 6,5% na habilidade do aluno.

A taxa de acurácia apresentou um significativo e progressivo aumento do número de itens lidos por minuto a cada ano escolar (2° < 3° < 4° < 5° ano; $H_{(3, 404)} = 91,2$; $p < 0,001$). A taxa de acurácia dos alunos do 2° ano foi menor do que a dos 3° ano [2° vs. 3° ano: $U_{(1, 218)} = 3,81$, $p = 0,0002$], que por sua vez obtiveram desempenho inferior aos do 4° ano [3° vs. 4° ano: $U_{(1, 222)} = 8,44$, $p < 0,0001$], que foram inferiores aos alunos do 5° ano [4° vs. 5° ano: $U_{(1, 185)} = 3,72$, $p = 0,0003$].

Para o rastreamento das crianças com dificuldade de leitura, selecionou-se um critério de corte no percentil 50 para indicar uma acurácia igual ou pior do que a média inferior da amostra. Com limite de tempo para a aplicação maior para os alunos mais novos, selecionou-se como critério de corte para o 2° e 3° ano uma acurácia de 56% de acertos; e para o 4° e 5° ano uma acurácia de 51% de acertos.

Como evidência de validade externa convergente, a variável Acurácia da PRP-BR completa apresentou correlação moderada com as notas dos alunos na disciplina de Português ($r = 0,44$) e de Matemática ($r = 0,46$) ($p < 0,0001$). Como validade externa de critério, não houve diferença na acurácia entre os meninos e as meninas. Não foi verificada diferença significativa da porcentagem de acertos ao item entre as palavras-alvo com duas sílabas quando comparada as de três sílabas; entre as palavras com três a nenhuma vizinhança ortográfica; ou entre as palavras com regularidade previsível e imprevisível.

3.3 Fase 4 – Validade interna e fidedignidade da PRP-BR completa e da reduzida

A qualidade estatística de cada item da PRP-BR foi analisada a partir da amostra da Fase 4 ($n = 231$ participantes), composta por apenas os participantes da Fase 2 e da Fase 3 que responderam todos os 52 itens (Tabela 1 e Tabela 2). A Tabela 3 apresenta os índices relativos à Teoria Clássica dos Testes (média da porcentagem de acertos dos itens individuais, os coeficientes de correlação biserial item-total corrigidos e o alpha de Cronbach na ausência do item) e os da Teoria de Resposta ao Item (índices de discriminação, dificuldade e probabilidade de acerto ao acaso e nível de dificuldade).

A amostra da Fase 4 apresentou uma porcentagem média de acertos próxima ao máximo ($97,9 \pm 2,7\%$), com mínimo de 87% (Tabela 4). Como evidência de fidedignidade, o alpha de Cronbach entre os 52 itens foi igual a 0,70, o que demonstra que a PRP-BR apresenta uma consistência interna aceitável. O alpha de Cronbach, ao considerar os quatro anos escolares, situou-se em 0,60, 0,64, 0,77 e 0,86, respectivamente.

Tabela 3.

Descrição dos itens da PRP-BR com parâmetros psicométricos avaliados pela Teoria Clássica dos Testes e Teoria de Resposta ao Item

Item	Palavra	<i>M</i>	Corr	α -item	<i>a</i>	<i>b</i>	<i>c</i>	Dif	Status
1	Banana	98,8	0,03	0,70	0,84	-3,18	0,29	II	Valid
2	Baleia	93,8	0,22	0,70	0,81	-1,99	0,50	III	Valid
3	Camelo	99,5	0,16	0,70	1,16	-3,18	0,26	II	Valid
4	Ponte	91,6	0,15	0,70	0,56	-2,89	0,19	II	Valid
5	Quatro	95,3	0,30	0,59	0,80	-2,4	0,17	III	Valid
6	Jarra	96,0	0,07	0,70	0,97	-2,53	0,21	III	Valid
7	Escova	98,5	0,04	0,70	0,89	-3,43	0,18	II	Valid
8	Casaco	98,0	0,20	0,70	0,75	-3,91	0,21	I	Valid
9	Bota	99,5	0,00	0,70	0,91	-4,27	0,23	I	Valid
10	Morango	93,1	0,29	0,59	0,63	-2,68	0,15	II	Valid
11	Quadro	97,0	0,13	0,70	0,75	-3,85	0,23	I	Valid
12	Laço	89,4	0,33	0,58	0,84	-1,49	0,35	III	Valid
13	Sapo	99,5	0,16	0,70	0,83	-3,76	0,28	I	Valid
14	Botão	97,5	0,10	0,70	0,87	-3,18	0,15	II	Valid
15	Garrafa	97,0	0,36	0,58	0,77	-3,82	0,18	I	Valid
16	Palhaço	99,5	0,16	0,70	1,29	-3,85	0,22	I	Valid
17	Prato	90,1	0,33	0,58	0,82	-1,59	0,20	III	Valid
18	Barraca	96,3	0,18	0,70	0,97	-2,56	0,21	III	Valid
19	Escada	98,6	0,03	0,70	0,42	-3,98	0,23	I	Valid
20	Queijo	97,0	0,04	0,70	0,66	-3,44	0,18	II	Valid
21	Martelo	98,6	0,03	0,70	0,95	-3,03	0,14	II	Valid
22	Pente	97,7	0,00	0,70	0,99	-2,82	0,15	II	Valid
23	Cadeira	96,3	0,18	0,70	0,68	-3,37	0,19	II	Valid
24	Cenoura	99,5	0,16	0,70	1,23	-3,31	0,17	II	Valid
25	Cinto	99,2	0,34	0,58	0,89	-3,64	0,24	I	Valid
26	Braço	99,6	-0,05	0,71					Rej
27	Chave	98,6	0,03	0,70	0,87	-2,88	0,20	II	Valid
28	Panela	99,4	0,22	0,70	1,08	-3,08	0,24	II	Valid
29	Cavalo	96,8	0,13	0,70	0,71	-3,42	0,21	II	Valid
30	Aranha	94,8	0,15	0,70	0,42	-4,15	0,31	I	Valid
31	Prego	98,8	0,00	0,70	0,67	-3,78	0,21	I	Valid
32	Cobra	97,0	0,21	0,70	1,00	-2,7	0,21	II	Valid
33	Galinha	99,1	0,22	0,70	1,32	-3,00	0,25	II	Valid
34	Bandeira	99,4	0,00	0,70	0,97	-3,55	0,27	II	Valid
35	Regador	98,1	0,00	0,70	0,69	-2,29	0,15	III	Valid
36	Barco	98,0	0,22	0,70	1,25	-2,62	0,25	III	Valid
37	Estrela	99,6	-0,05	0,71					Rej
38	Faca	96,8	0,14	0,70	0,79	-2,85	0,20	II	Valid
39	Gato	95,4	0,12	0,70	0,67	-2,67	0,26	II	Valid
40	Cinco	98,2	0,21	0,70	1,20	-2,48	0,25	III	Valid
41	Milho	99,6	-0,05	0,71					Rej
42	Coelho	97,7	0,26	0,70	1,47	-2,16	0,25	III	Valid
43	Galo	99,6	-0,05	0,71					Rej
44	Bola	99,6	-0,05	0,71					Rej
45	Peixe	99,6	-0,05	0,71					Rej
46	Violão	83,8	0,21	0,70	0,72	-1,21	0,26	III	Valid
47	Leão	99,6	-0,05	0,71					Rej
48	Rato	99,6	-0,05	0,71					Rej
49	Piano	99,5	0,22	0,70	1,08	-3,15	0,26	II	Valid
50	Sereia	99,6	-0,05	0,71					Rej
51	Espada	94,2	0,20	0,70	0,82	-2,07	0,26	III	Valid
52	Gaveta	98,6	0,03	0,70	0,78	-3,19	0,26	II	Valid

Legenda. Item: número da ordem de apresentação do item; Palavra: palavra-alvo correspondente ao desenho; *M*: média da porcentagem de participantes que acertaram o item; Corr: correlação entre o acerto do item com o resultado da PRP-BR; α -item: alpha de Cronbach na ausência do item; *a*: índice de discriminação do item; *b*: índice de dificuldade; *c*: probabilidade de acerto ao acaso; Dif: Nível de dificuldade; Status: item validado (Valid) ou rejeitado (Rej) para inclusão na PRP-BR reduzida.

Tabela 4.

Análise descritiva (Média, Desvio Padrão, Mínimo, Máximo) da Acurácia (porcentagem de acertos) e da Taxa de Acurácia (itens por minutos) da PRP-BR aplicada na Fase 2 (n = 114 participantes), na Fase 3 (n = 480) e na Fase 5 (n = 245)

Fase	Variável	Estatística	Ano escolar					Total
			1°	2°	3°	4°	5°	
Fase 2 (52 itens)	Acurácia (%)	Média	-	96,8	98,1	98,8	99,0	98,2
		Desvio Padrão	-	2,7	2,7	2,1	2,8	2,7
		Mínimo	-	92,3	90,4	90,4	86,5	86,5
		Máximo	-	100	100	100	100	100
Fase 3 (52 itens)	Acurácia (%)	Média	-	66,5	78,4	51,9	63,9	64,2
		Desvio Padrão	-	22,6	21,9	21,9	22,1	23,9
		Mínimo	-	17,3	23,1	11,5	17,3	11,3
	Taxa de acurácia (itens por minuto)	Máximo	-	100	100	100	100	100
		Média	-	8,7	10,2	13,5	16,6	12,8
		Desvio Padrão	-	2,9	2,8	5,7	5,8	5,7
Fase 5 (43 itens)	Acurácia (%)	Mínimo	-	2,3	3,0	3,0	4,5	2,3
		Máximo	-	13	13	26	26	26
		Média	35,0	64,8	55,2	57,5	58,0	54,5
		Desvio Padrão	27,3	28,9	28,2	24,3	20,7	28,2
	Taxa de acurácia (itens por minuto)	Mínimo	6,8	6,8	4,5	2,3	25,0	2,3
		Máximo	97,7	100	100	100	95,5	100
		Média	3,8	7,0	11,9	12,4	12,5	9,6
		Desvio Padrão	2,9	3,1	6,0	5,2	4,5	5,8
		Mínimo	0,7	0,7	1,0	0,5	5,4	0,5
		Máximo	10,5	10,8	21,5	21,5	20,5	21,5

Ao analisar os dados de acordo com a Teoria Clássica dos Testes, os coeficientes de correlação corrigida item-total variaram de -0,05 a 0,41 ($M = 0,12 \pm 0,13$). Seis itens (*i.e.*, quatro, prato, laço, cinto, garrafa, banana) obtiveram correlações item-total acima do critério ($r > 0,30$). Nove itens apresentaram valores item-total negativos (*i.e.*, bola, braço, estrela, galo, leão, milho, peixe, rato, sereia), o que indica que os itens possuem problemas, uma vez que quanto maior a habilidade do aluno, menor sua chance de responder corretamente ao item em questão.

O alpha de Cronbach aumentou de 0,70 para 0,71 na ausência de qualquer um dos nove itens com bisseriais menores do que zero. Outras cinco palavras apresentaram bisserial igual 0 (*i.e.*, bota, bandeira, pente, prego, regador), porém suas exclusões não alteraram o alpha de Cronbach. De acordo com o índice de discriminação (parâmetro *a*), da Teoria de Resposta ao Item, foram retirados nove dos 52 itens da prova original (*i.e.*, braço, bola, estrela, galo, leão, milho, peixe, rato, sereia). Uma análise simples do comportamento aponta um alto percentual de acerto (> 99,5%) em todas essas palavras excluídas.

Após essas exclusões, foi possível elaborar uma versão reduzida da PRP-BR composta por 43 itens. O alpha de Cronbach foi classificado como aceitável ($\alpha = 0,71$). A PRP-BR reduzida apresentou alto poder de discriminação *a* ($M = 0,88 \pm 0,23$; $Med = 0,84$; $Mín = 0,42$; $Máx = 1,47$), com todos os itens atingido o critério ($a > 0,30$). O valor médio do índice de dificuldade, parâmetro *b*, foi de $-3,01 \pm 0,72$, com uma variação entre -4,27 e -1,21. O parâmetro *c* apresentou valores baixos e dentro do ideal para todos os itens ($M = 0,22 \pm 0,04$; $Mín = 0,14$; $Max = 0,28$).

3.4 Fase 5 – Fidedignidade e validade da PRP-BR reduzida

No estudo empírico da Fase 5 (n = 245 participantes) foi realizado com aplicação da PRP-BR reduzida (43 itens) e tempo limite para verificar a validade interna e externa do instrumento. A variável acurácia apresentou uma média de $54,5 \pm 28,2\%$, com mínimo de 2,3% (Tabela 4). Na análise dos dados da consistência interna (n = 50 participantes que finalizaram os 43 itens dentro do tempo), a versão reduzida da PRP-BR apresentou um coeficiente alpha de Cronbach muito alto ($\alpha = 0,97$), o que demonstra alta fidedignidade. Ao considerar os cinco anos escolares separadamente, o alpha situou-se em 0,99, 0,98, 0,86, 0,83 e 0,94.

Verificou-se um significativo efeito de escolaridade na variável acurácia ($H_{(4, 241)} = 29,7$; $p < 0,001$) e na taxa de acurácia ($H_{(4, 241)} = 105,1$; $p < 0,001$). Análise *post hoc* identificou que a diferença concentrava-se entre os alunos do primeiro ano em comparação aos demais anos escolares (1° < 2°–5° ano), tanto nas médias da variável acurácia ($U_{(1, 244)} = 6411,5$, $p < 0,001$) quanto da taxa de acurácia ($U_{(1, 244)} = 7002,0$, $p < 0,001$). Para a versão da PRP-BR reduzida, selecionou-se como critério de corte para os alunos do 1° ano uma acurácia de 23% de acertos; e para os do 2° ao 5° ano uma acurácia de 56% de acertos.

Como evidência de validade externa convergente da PRP-BR reduzida, a partir da amostra da Fase 5, as notas dos alunos na disciplina de Português apresentaram correlação moderada com a acurácia ($r = 0,33$) e a taxa de acurácia ($r = 0,42$) ($p < 0,0001$). Não houve diferença na acurácia entre os meninos e as meninas. Não foi verificada diferença significativa entre a porcentagem de participantes que acertaram as palavras com duas sílabas quando comparada as

de três sílabas, ou entre as palavras com três a nenhuma vizinhança ortográfica, ou entre as palavras com diferentes regularidades.

4. DISCUSSÃO

O objetivo do presente estudo foi prover fontes de evidências psicométricas de fidedignidade, validade da estrutura interna e validade externa para a Prova de Reconhecimento de Palavras – Versão Brasileira (PRP-BR), adaptada por Pinheiro (2008) da versão original portuguesa (Viana & Ribeiro, 2007), como resultado da cooperação celebrada entre a Universidade Federal de Minas Gerais e a Universidade do Minho. A partir da amostra da Fase 2, planejada como estudo piloto para avaliação da iconicidade com aplicação sem tempo limite, investigou-se o efeito de escolaridade na acurácia de leitura. Na Fase 3, a PRP-BR foi aplicada com tempo limite para avaliar a fluência de leitura e o efeito de escolaridade. Na Fase 4, analisou-se a fidedignidade dos participantes da Fase 2 e Fase 3 que responderam todos os 52 itens. Por fim, na Fase 5 aplicou-se a PRP-BR reduzida com tempo limite.

É relevante diferenciar a presente PRP-BR de outros instrumentos já disponíveis no Brasil, como o Teste de Reconhecimento de Palavras (TRP) e o Teste de Reconhecimento de Pseudopalavras (TRPp) (Pinheiro & Vilhena, 2022; Vilhena & Pinheiro, 2022) e o Teste de Leitura: Compreensão de Sentenças (TELCS) (Vilhena et al., 2016; Vilhena & Pinheiro, 2020), publicados pela Editora NilaPress em julho de 2023 e em agosto de 2022, respectivamente. A PRP-BR avalia um dos primeiros componentes a consolidar-se na alfabetização, que é o reconhecimento em silêncio de palavras isoladas. De forma distinta, o TRP e o TRPp exigem a leitura em voz alta de uma lista de palavras e de pseudopalavras, respectivamente. Assim, os TRP e TRPp são cognitivamente mais difíceis, pois necessitam da adicional ativação de regiões cerebrais envolvidas na produção acurada e fluente da fala durante a leitura.

A avaliação do reconhecimento de palavras (avaliado tanto pela PRP quanto pelo TRP) pode ser complementada com tarefas mais complexas, como a compreensão leitora. O TELCS é um teste padrão-ouro para avaliar a compreensão de sentenças. O TELCS é uma adaptação transcultural do teste francês L3 de Lobrot, cujo Teste de Idade de Leitura (TIL) é o equivalente em Portugal (Sucena & Castro, 2010). Por ser uma avaliação mais fácil, a aplicação da PRP-BR pode ser realizada já no primeiro ano do Ensino Fundamental, diferente dos TRP/TRPp e do TELCS que são aplicados a partir do segundo ano. Essa característica torna a PRP-BR um instrumento ainda mais importante na triagem precoce das crianças que precisam de maior atenção no processo de aprendizagem da alfabetização. Como vantagem, tanto a PRP e quanto o TELCS podem ser aplicados em pessoas surdas, uma vez que são de autoaplicação e de leitura silenciosa.

4.1 Estudo piloto de iconicidade

A realização do estudo piloto é um dos pré-requisitos da validade de conteúdo transcultural. Na Fase 2, planejou-se

um estudo piloto com 114 participantes do segundo ao quinto ano do Ensino Fundamental para avaliar empiricamente a iconicidade dos 52 desenhos selecionados na Fase 1. Para avaliar a correspondência correta entre o desenho e a palavra-alvo, a aplicação não teve tempo limite.

Todos os participantes avaliados, inclusive os do 2º ano, apresentaram familiaridade com todos os 52 desenhos, com acurácia próxima ao máximo ($M = 98,2 \pm 2,7\%$). Uma vez que a porcentagem mínima de acertos ao item foi de 87% na Fase 2, todos os itens ultrapassaram o critério mínimo de 80% de acertos. Estes resultados sugeriram a manutenção de todos os 52 itens (PRP-BR completa), pois os participantes compreenderam com facilidade a correta correspondência entre a ilustração e a palavra-alvo (Pinheiro et al., *in press*).

Como comparação, no estudo piloto de iconicidade da versão original em Portugal, baseado na resposta de 272 crianças de quatro classes do 1º ciclo do ensino básico, eliminou-se 37 dos 77 itens iniciais que obtiveram acurácia menor do que o critério de 80% de acertos, totalizando 40 itens a versão final da PRP-PT (Viana & Ribeiro, 2010).

4.2 Evidências de fidedignidade

Utilizou-se a Teoria Clássica dos Testes e a Teoria de Resposta ao Item para analisar a qualidade de cada item da PRP-BR. A Teoria Clássica dos Testes é a forma mais conhecida de correção de uma avaliação, onde a habilidade do aluno é medida por meio da nota (percentual de questões corretas) (Lord & Novick, 1968).

Como evidência de fidedignidade, utilizou-se o alfa de Cronbach, que avalia o grau em que os itens de uma matriz de dados estão correlacionados entre si. O alfa de Cronbach, calculado desconsiderando os dados faltantes, aumentou progressivamente entre as fases deste estudo: Fase 2 (0,55), Fase 3 (0,61), Fase 4 (0,70) e Fase 5 (0,97). Esses resultados indicam que a PRP-BR apresenta uma consistência interna adequada. Em Portugal, como evidência de fidedignidade, Viana et al. (2013) identificaram coeficientes de consistência interna elevados tanto no estudo com 328 participantes sem tempo limite ($\alpha = 0,98$) quanto no com 3.131 participantes com tempo limite ($\alpha = 0,96$ a $0,98$).

Na Fase 4, ao considerar os quatro anos escolares separadamente, o alfa de Cronbach situou-se em 0,60, 0,64, 0,77 e 0,86. Na Fase 5, para os cinco anos avaliados, o alfa situou-se em 0,99, 0,98, 0,86, 0,83 e 0,94. Em Portugal, houve similares coeficientes nos quatro anos de escolaridade (0,96; 0,98; 0,97 e 0,98), o que indica que a PRP-PT permite a medida do reconhecimento de palavras de modo igualmente fidedigno nos anos a que se destina (Viana et al., 2013). Esses resultados replicam um estudo prévio em Portugal que encontrou elevados alfa de Cronbach nos quatro anos de escolaridade (0,95, 0,95, 0,93 e 0,84) (Viana & Ribeiro, 2010).

A análise dos itens por meio da Teoria Clássica dos Testes deve ser considerada apenas uma etapa preliminar, uma vez que o alfa de Cronbach tem como premissa que os itens retidos apresentam a mesma importância para o fator. Assim, utilizou-se a Teoria de Resposta ao Item de três parâmetros para analisar a PRP-BR, pois fornece melhores informações a respeito das características dos itens e do nível de informação que agregam ao cálculo da escala de

habilidades. O modelo de três parâmetros é o mais completo e considera não só a dificuldade do item e sua discriminação, como também um terceiro parâmetro relativo à probabilidade de acerto dos indivíduos com baixa habilidade. Este modelo é fundamentado na ideia de que quanto maior a habilidade em leitura do aluno, maior será a chance dele responder corretamente a um item (selecionar a palavra-alvo).

Para a análise da capacidade de discriminação do item, parâmetro *a*, utilizou-se a correlação bisserial, que mede a associação entre a habilidade (traço latente) e o resultado do teste (escore bruto total). De acordo com parâmetro *a*, foram retirados nove dos 52 itens da PRP-BR (*i.e.*, braço, bola, estrela, galo, leão, milho, peixe, rato, sereia), pois apresentaram o bisserial menor do que zero, indicando que quanto maior a habilidade do aluno menor sua chance de responder corretamente ao item em questão. Ou seja, estas nove palavras apresentam baixa capacidade de discriminação em relação ao resultado da PRP-BR.

Dessa forma, foi possível elaborar uma versão reduzida da PRP-BR, composta por 43 itens. A PRP-BR reduzida apresentou alto poder de discriminação *a* ($M = 0,88 \pm 0,23$), com todos os itens acima do critério de corte de 0,30, o que demonstra que medem adequadamente o traço latente desejado e possui capacidade de discriminação. O índice de dificuldade, parâmetro *b*, permitiu o agrupamento dos itens em três níveis de dificuldade. O parâmetro *c*, que mede a chance dos sujeitos com baixa habilidade acertarem um item com dificuldade maior que sua habilidade, apresentou valores dentro do ideal. Os resultados das análises da Teoria de Resposta ao Item, ao considerar que a PRP-BR dá mais informação para as faixas de habilidades mais baixas, reforça a validade de que a prova é um bom instrumento para avaliar o reconhecimento de palavras em sujeitos com dificuldade de leitura ou que se encontram nos estágios iniciais de alfabetização.

4.3 Efeito de escolaridade

Como evidências de validade interna para a PRP-BR, verificou-se um significativo efeito de escolaridade tanto com a versão completa quanto a reduzida em todas as amostras. Quanto maior o grau de escolarização do aluno, maior será a sua habilidade de reconhecimento de palavras. Na Fase 2, apesar da PRP-BR ter sido aplicada sem tempo limite, os participantes do segundo ano obtiveram uma menor acurácia quando comparados aos do quarto e quinto ano do Ensino Fundamental (*i.e.*, $2^\circ < 4^\circ - 5^\circ$ ano).

Na Fase 3, ao comparar os anos escolares com o mesmo tempo limite de aplicação, os participantes do 2º ano apresentaram uma menor acurácia do que os do 3º ano; e os do 4º ano uma menor acurácia do que os do 5º ano. O maior tempo de aplicação para os participantes do 2º e do 3º ano (quatro minutos), quando comparado ao 4º e 5º ano (dois minutos), resultou em uma maior acurácia para as crianças dos primeiros anos escolares [$(2^\circ < 3^\circ) > (4^\circ < 5^\circ$ ano)]. Ao considerar que houve diferença no limite de tempo de aplicação da PRP-BR, a variável Taxa de Acurácia permitiu verificar um gradual aumento da acurácia ($2^\circ < 3^\circ < 4^\circ < 5^\circ$ ano). Na Fase 5, verificou-se uma diferença entre os alunos do primeiro ano, quando comparado aos demais anos

escolares, nas médias da variável acurácia e da taxa de acurácia ($1^\circ < 2^\circ - 5^\circ$ ano).

Resultados similares aos da PRP-BR já haviam sido encontrados em Portugal. Viana et al. (2013) também verificaram um significativo efeito de escolaridade para a PRP-PT, com os participantes do primeiro ano apresentando um desempenho menor, com tamanho de efeito elevado, em comparação ao segundo ano ($1^\circ < 2^\circ$ ano, $t = 31,98$, $p = 0,001$). Já os participantes do terceiro ano apresentaram diferença significativa e com tamanho de efeito moderado em comparação aos do quarto ano ($3^\circ < 4^\circ$ ano, $t = 12,1$, $p = 0,001$).

Criou-se critério de cortes a partir dos resultados da Fase 2 a Fase 5. Ao considerar a PRP-BR sem tempo limite de aplicação (Fase 2), é esperado que as crianças do 2º ao 5º ano, ao final do ano letivo, obtenham uma acurácia no reconhecimento de palavras acima de 96%. Ao sumarizar os critérios de corte da PRP-BR com limite de tempo (Fase 3 e Fase 5), pode-se considerar que é esperado que as crianças do 2º ao 5º ano, ao final do ano letivo, apresentem uma acurácia no reconhecimento de palavras acima de 60% dentro do tempo, o que equivale a 31 itens na versão completa e 26 itens na versão reduzida da prova.

4.4 Validade externa

Como validade externa de critério, os resultados encontrados nos estudos psicométricos da Fase 2 a Fase 5 da PRP-BR devem estar coerentes com a literatura. Os resultados revelaram a importância de se avaliar a fluência de leitura e não somente a acurácia. Na Fase 2, ao não limitar o tempo de aplicação, a acurácia ficou próxima ao máximo ($M = 98,2 \pm 2,7\%$). Pinheiro e Vilhena (2022) revisaram diferentes instrumentos que avaliam a acurácia do reconhecimento de palavras e confirmaram que as provas de leitura de palavras isoladas são fáceis de serem lidas no português, alcançando 85% de acertos já a partir do segundo ano do Ensino Fundamental.

Esse resultado de alta acurácia já no início da alfabetização está de acordo com a literatura, pois o português brasileiro possui uma ortografia transparente no sentido da leitura (relação grafema–fonema) e uma estrutura silábica simples (Scliar-Cabral, 2023; Scliar-Cabral & Heinig, 2023). Scliar-Cabral (2019) explica que o leitor hábil evoca, rapidamente, por meio da decodificação de um pequeno número de grafemas, a representação fonológica da palavra no léxico fonológico que ele ‘escuta’ em sua fala interior (rota fonológica), a qual é utilizada no acesso de seu significado na memória semântica. Assim, uma vez desconsiderada a fluência de leitura, as crianças, inclusive as com dificuldades de leitura, podem atingir altas porcentagens de acurácia após poucos meses de instrução formal (Pinheiro & Vilhena, 2022; Vilhena & Pinheiro, 2022).

Como outra evidência de validade externa de critério para a PRP-BR, ao considerar os resultados da Fase 2 a Fase 5, tanto a variável Acurácia quanto a Taxa de Acurácia não apresentaram diferença significativa entre meninos e meninas, o que corrobora a literatura (Lúcio et al., 2018; Pinheiro & Vilhena, 2022; Vilhena & Pinheiro, 2020). Em Portugal, Viana et al. (2013) também não verificaram diferença de

desempenho na PRP-PT em função do sexo nos quatro anos analisados.

Como evidência de validade externa convergente, as acurácias de leitura na PRP-BR versão completa ($r = 0,44$) e versão reduzida ($r = 0,33$) apresentaram correlação significativa e moderada com as notas dos alunos dada pelos professores na disciplina de Português. Em dois estudos em Portugal, a PRP-PT apresentou correlações de moderadas a fortes com uma avaliação da leitura realizada pelos professores que variaram entre 0,62 e 0,84 (Viana e Ribeiro, 2010) e de 0,63 a 0,77 (1º ano: $r = 0,69$; 2º ano: $r = 0,77$; 3º ano: $r = 0,63$; 4º ano: $r = 0,67$; $p < 0,001$) (Viana et al., 2013). Em Portugal, como fonte de validade externa convergente, a PRP-PT apresentou correlação positiva e significativa, em cada ano escolar, com o Teste de Idade de Leitura (TIL) (Sucena & Castro, 2010) [Equivalente ao Teste de Leitura: Compreensão de Sentenças (TELCS) no Brasil (Vilhena et al., 2016)] (2º ano: $r = 0,63$; 3º ano: 0,34; 4º ano: 0,38; $p < 0,001$) (Viana et al., 2013). Embora ambos os instrumentos avaliem a leitura silenciosa, o PRP avalia acurácia e fluência de leitura, enquanto o TIL (ou TELCS) combina as variáveis acurácia, fluência e compreensão. Segundo Viana et al., esses valores de correlação estão dentro da amplitude de 0,42 a 0,68 encontrada em estudos sobre fluência e compreensão de leitura. Viana et al. argumentam que as correlações entre medidas de leitura são mais elevadas nos anos iniciais de escolaridade, perdendo o poder explicativo à medida que os processos de decodificação são automatizados.

Como limitação do estudo, é importante que sejam feitos estudos normativos para criar critérios de corte da capacidade de reconhecimento de palavras para diferentes regiões do Brasil, para pessoas com surdez e para grupos clínicos. Outra limitação é não ser possível desconsiderar o erro da própria medida na Teoria de Resposta ao Item, que pode ter levado a exclusão errada de nove dos 52 itens da PRP-BR. O participante pode marcar erroneamente a palavra correspondente ao desenho, mesmo sabendo a alternativa correta. Alguns alunos apresentaram um comportamento definido como 'perfeccionista', com uma demasiada preocupação na estética da marcação do X na resposta correta, em detrimento de responder com velocidade. Houve um possível efeito de ordem, com os itens posicionados ao final da aplicação dos 52 itens sendo justamente os que apresentaram os piores índices psicométricos, o que pode sugerir que alguns participantes podem ter acelerado a resposta para terminar todos os itens dentro do tempo, em detrimento da acurácia. Também não se pode descartar o erro do pesquisador na transposição das respostas dos participantes para a planilha de Excel, apesar das exaustivas revisões da integridade do banco de dados.

Futuros estudos poderão investigar as fontes de validade externa convergente entre a PRP com o TRP, TRPp e o TELCS. Apesar das aplicações da PRP-PT e da PRP-BR terem sido planejadas no formato papel e caneta, é plenamente possível digitalizar a prova para ser aplicada em dispositivos digitais. Ao invés de marcar o papel com a caneta, a seleção da alternativa correta poderá ser feita tanto com o clique do *mouse* nos computadores quanto com o toque do dedo na tela nos *tablets*. A randomização digital da ordem de apresentação dos estímulos poderá controlar o efeito de ordem do item. Essa digitalização da aplicação demandará novos estudos psicométricos para verificar se a alteração da

dimensão dos desenhos e a disposição das alternativas alteram a fidedignidade da prova.

A aplicação digital da PRP-BR poderá ter como vantagem a redução do custo impressão das folhas de resposta (três folhas tanto para a impressão das seis páginas da versão completa quanto das cinco páginas da versão reduzida), apresentação randomizada dos itens, não precisar virar as páginas das folhas de resposta, a interrupção automática da aplicação no exato limite de tempo, a agilidade do participante em marcar a alternativa, a mensuração do tempo de reação da resposta ao item e do tempo total de leitura, a conversão do escore bruto nas variáveis acurácia e taxa de acurácia e a geração automática do relatório individualizado. Assim, espera-se que tanto a versão impressa quanto a digital sejam fáceis e rápidas de serem aplicadas, com resposta intuitiva para as crianças.

5. CONCLUSÃO

A PRP-BR avalia a capacidade da criança de reconhecer palavras isoladas de forma acurada e fluente, que são dois componentes importantes para o acesso ao sistema semântico e para uma leitura proficiente. O presente estudo alcançou o objetivo de prover fontes satisfatórias de evidências psicométricas de fidedignidade, validade da estrutura interna e validade externa para a PRP-BR. No primeiro estudo, a Fase 1 de validação de conteúdo apresentou a análise da iconicidade dos desenhos, as características psicolinguísticas das palavras-alvo e as alternativas distratoras (Pinheiro et al., *in press*). No estudo piloto da Fase 2, planejado para avaliar a efetiva iconicidade dos desenhos, todos os itens atingiram o critério de corte de 80% de acurácia, o que sugeriu suas manutenções para os subsequentes estudos de validação. A partir dos dados da Fase 4, ao considerar os resultados da Teoria Clássica dos Testes e da Teoria de Resposta ao Item, foram retirados nove dos 52 itens. Como evidência de adequada fidedignidade, o alfa de Cronbach aumentou progressivamente entre as fases deste estudo. Como evidências de validade da estrutura interna, verificou-se um significativo efeito de escolaridade tanto com a versão completa (Fase 2 e Fase 3) quanto a reduzida (Fase 5). As evidências de validade externa convergente e de critério apresentaram-se coerentes com a literatura. Conclui-se que a PRP-BR é um instrumento com evidências psicométricas satisfatórias para avaliar o construto reconhecimento de palavras.

Referências

- Cramer, D., & Howitt, D. L. (2004). *The SAGE Dictionary of Statistics: A Practical Resource for Students in Social Sciences*. SAGE Publications Ltd. <https://doi.org/10.4135/9780857020123>
- George, D., & Mallery, P. (2003). *SPSS for Windows step by step: A simple guide and reference 11.0 update* (4th ed.). Allyn & Bacon.
- Grainger, J., O'Reagan, K., Jacobs, A., & Segui, J. (1989). On the role of competing word units in visual word recognition: the neighborhood frequency effect. *Perception & Psychophysics*, 45(3), 189-195. <https://doi.org/10.3758/bf03210696>
- Kline, P. (1993). *Handbook of psychological testing*. Routledge.

- Komesidou, R., Feller, M. J., Wolter, J. A., Ricketts, J., Rasner, M. G., Putman, C. A., & Hogan, T. P. (2022). Educators' perceptions of barriers and facilitators to the implementation of screeners for developmental language disorder and dyslexia. *Journal of Research in Reading*, 45, 277-298. <https://doi.org/10.1111/1467-9817.12381>
- Lord, F. M., & Novick, M. R. (1968). *Statistical Theories of Mental Test Scores*. Addison-Wesley, Menlo Park.
- 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, e3437. <https://dx.doi.org/10.1590/0102.3772e3437>
- Paap, K. R., & Johansen, L. S. (1994). The case of the vanishing frequency effect: a retest of the verification model. *Journal of Experimental Psychology: Human Perception & Performance*, 20(6), 1129-1157. <https://doi.org/10.1037/0096-1523.20.6.1129>
- Pinheiro, Â. M. V. (2008). Prova de Reconhecimento de Palavras para crianças – Versão Brasileira (PRP-BR). *Relatório de pesquisa apresentado ao Comitê de Internacionalização da Diretoria de Relações Internacionais da Universidade Federal de Minas Gerais*. Universidade Federal de Minas Gerais.
- 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 – Child Language Data Exchange System. <http://childes.talkbank.org/derived>
- 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. <https://doi.org/10.17058/signo.v47i88.17396>
- Pinheiro, Â. M. V., Vilhena, D. A., Ribeiro, I. S., & Viana, F. L. P. (in press). Adaptação transcultural e validade de conteúdo da Prova de Reconhecimento de Palavras. Manuscrito submetido na *Revista Estudos de Psicologia (Campinas)*.
- Rumelhart, D. E., & Maclelland, J. L. (1982). An interactive model of effects in letter perception: Part 2. The contextual enhancement effect and some tests and extensions of the model. *Psychological Review*, 89(1), 60-94. <https://doi.org/10.1037/0033-295X.89.1.60>
- Scliar-Cabral, L. (2019). Precisamos de mais de 50.000 palavras no léxico ortográfico para ler? *Letras de Hoje*, 54(2), 122-131. <http://dx.doi.org/10.15448/1984-7726.2019.2.34510>
- Scliar-Cabral, L. (2023). *Princípios do sistema alfabético do português brasileiro*. 2ª ed. Editora Lili, 296p.
- Scliar-Cabral, L., & Heinig, O. L. O. M. (2023). Principais contribuições da linguística para a alfabetização. *Fórum Linguístico*, 20(1), 8716-8726. <https://doi.org/10.5007/1984-8412.2023.e87147>
- Sucena, A., & Castro, S. L. (2010). *Aprender a Ler e Avaliar a Leitura. O TIL: Teste de Idade de Leitura* (2nd ed.). Almedina.
- Viana, F. L., & Ribeiro, I. (2007). *P.R.P. - Prova de Reconhecimento de Palavras. Contributos para a validação de um instrumento de avaliação de leitura*. Trabalho apresentado no III Congresso Brasileiro de Avaliação Psicológica: Formas e Contextos, João Pessoa, PB, Brasil.
- Viana, F. L., & Ribeiro, I. (2010). *PRP – Prova de Reconhecimento de Palavras*. Edições Cegoc-Tea.
- Viana, F. L., Ribeiro, I. S., Maia, J., & Santos, S. (2013). Propriedades psicométricas da Prova de Reconhecimento de Palavras. *Psicologia: Reflexão e Crítica*, 26(2), 231-240. <https://doi.org/10.1590/S0102-79722013000200003>
- Vilhena, D. A., & Pinheiro, Â. M. V. (2020). Reliability, Validity and Standardization of the Reading Test: Sentence Comprehension. *Psicologia: Teoria e Pesquisa*, 36, e36325. <https://doi.org/10.1590/0102.3772e36325>
- Vilhena, D. A., & Pinheiro, Â. M. V. (2022). Fidedignidade, validade interna e normatização dos Testes de Reconhecimento de Palavras e de Pseudopalavras. *Neuropsicologia Latinoamericana*, 14(2), 1-10. <https://doi.org/10.5579/nl.2022.0742>
- 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. <https://doi.org/10.1002/dys.1521>
- World Medical Association (2008). *Declaration of Helsinki. Ethical principles for medical research involving human subjects*. 59th WMA General Assembly, Seoul.