

PROLEC-T – Prova de Compreensão de Texto: Análise de suas Características Psicométricas

Ângela Maria Vieira Pinheiro¹

Douglas de Araújo Vilhena

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

Marcos Antônio da Cunha Santos

*Departamento de Estatística, Instituto de Ciências Exatas, Universidade Federal
de Minas Gerais, Belo Horizonte, MG, Brasil*

Resumo

A Prova de Compreensão de Texto da Bateria PROLEC (PROLEC-T), amplamente utilizada no Brasil, apresenta amostra limitada e apenas análises descritivas dos escores. Objetivos: conhecer a estrutura interna (adequação dos itens via Teoria de Resposta ao Item) e os índices de validade externa (concorrente) da referida prova. Método: estudo das respostas de 457 alunos (49,7% meninos) do 2º ao 5º ano, aleatoriamente selecionados em 77 salas de aula de oito escolas estaduais, ao PROLEC-T e a três testes de leitura. Resultados: a prova apresenta baixo nível de dificuldade e pouca variabilidade de escores, evidenciando validade interna não satisfatória. Mostra também limitações em relação à validade concorrente, uma vez que as correlações entre os seus escores e o ano escolar, a idade da criança e as notas finais na disciplina Língua Portuguesa foram fracas, embora tenham sido medianas com as medidas de leitura. Conclusão: Em sua forma atual, o PROLEC-T parece ser adequada apenas para uma avaliação informal de crianças com dificuldade de leitura. No entanto, caso seja reformulada, sugere-se a substituição de seus textos por histórias inéditas, em ordem crescente de dificuldade, com apenas perguntas inferenciais.

Palavras-chave: Avaliação de leitura, compreensão de textos, validade instrumentos, teoria de resposta ao item.

PROLEC-T – Text Comprehension Test: Psychometric Properties Analysis

Abstract

The Text Comprehension task (PROLEC-T) of the PROLEC Battery of Evaluation of Reading Processes, widely used in Brazil, has a limited sample and only descriptive analyses. Purpose: verify the internal structure (via the Item Response Theory) and external (concurrent) validity indices of the test. Methods: study of 457 student's responses (49.7% boys), from 2nd to 5th year randomly selected in 77 classrooms in eight state schools, to the PROLEC-T and three reading tests. Results: PROLEC-T has a low level of difficulty and low variability of scores, indicating unsatisfactory internal validity. It also shows limitations on the concurrent validity, since the correlations were weak between its scores and the

¹ Endereço para correspondência: Universidade Federal de Minas Gerais, Faculdade de Filosofia e Ciências Humanas, Programa de Pós-Graduação em Psicologia: Cognição e Comportamento, Av. Antônio Carlos, 6.627, Belo Horizonte, MG, Brasil 31270-901. Fone: +55-31-98894-4600. E-mail: pinheiroamva@gmail.com

school year, age of the child and his or her final grades in Portuguese, in spite of being moderate with reading measures. Conclusion: The test seems to be only suitable for an informal assessment of children with reading difficulties. However, if reformulated, it is suggested to replace the current texts for unpublished stories, increasing order of difficulty, and questions with only inferential questions.

Keywords: Reading evaluation, text comprehension, validity of instruments, item response theory.

PROLEC-T – Prova de Compreensão de Textos: Análisis de sus Características Psicométricas

Resumen

La prueba de Comprensión de textos de la Batería de Evaluación de los Procesos Lectores PROLEC (PROLEC-T), ampliamente utilizada en Brasil, tiene una muestra limitada y sólo el análisis descriptivo. Objetivo: conocer la estructura interna (adaptación de las cuestiones a través de la Teoría de Respuesta al Ítem) y los índices de validez externa (concurrentes) de esa prueba. Método: estudio de respuestas de 457 alumnos (49.7% de chicos) del 2º al 5º grado de la primaria, aleatoriamente seleccionados en 77 aulas de ocho escuelas públicas del Estado, a el PROLEC-T y tres pruebas de lectura. Resultados: La prueba tiene un conjunto de dificultad bajo y una baja variabilidad de las puntuaciones, lo que indica la validez interna no satisfactoria. También muestra limitaciones en la validez concurrente, ya que las correlaciones eran débiles con el año escolar, la edad del niño y con las notas finales en la asignatura Lengua Portuguesa, aunque hayan presentado mediana con las medidas de lectura. Conclusión: el establecimiento de normas para la PROLEC-T no es justificable. De esta manera, parece ser sólo adecuado para una evaluación informal de los niños con dificultades de lectura. Se sugiere la sustitución de los textos por historias inéditas, seguidas solamente de preguntas inferenciales, y orden creciente de dificultad de los textos.

Palabras clave: Evaluación de lectura, comprensión de textos, validez de instrumentos, teoría de respuesta al ítem.

A habilidade de compreensão tanto na audição como na leitura se dá por meio de processos múltiplos e complexos, tais como o pensamento, o raciocínio e a interpretação, que funcionam de forma paralela, interativa e rápida. Na leitura, quando a decodificação se torna automatizada (processo de conversão grafema–fonema preciso e fluente), a compreensão ocorre com quase nenhum custo para o sistema cognitivo (Kintsch & van Dijk, 1978), marcando assim o ponto final da alfabetização. Subjacentes a esse desenvolvimento encontram-se as habilidades centrais da linguagem, como a compreensão oral e a extensão do vocabulário do leitor. A capacidade de usar a sintaxe para antecipar o significado de palavras, a aptidão para monitorar o contexto e fazer inferências com base no conhecimento prévio são outros pré-requisitos para a compreensão (Kamhi, 2009; McGuinness, 2004). Ademais, a

familiaridade com gêneros literários, em particular com a organização dos textos expositivos, em contraste com os narrativos (que são bem conhecidos pelas crianças antes da aprendizagem da leitura), tem sido também citada pela literatura.

O texto narrativo mostra uma estrutura claramente definida, constituída à base de padrões recorrentes, uma vez que deve possuir uma introdução, um clímax e uma conclusão. O discurso desse estilo textual recapitula as experiências passadas, sendo fiel à sequência temporal dos acontecimentos vividos pelo narrador. Objetiva ressaltar fatos relevantes, como por exemplo, uma lição importante aprendida durante a história. Já nos textos expositivos, a sucessão dos eventos tem como prioridade a importância dos fatos para a exposição, independentemente da ordem em que realmente ocorrem. O propósito desse gênero textual é explicar sobre um tópi-

co específico (um conceito ou uma ideia), sobre o qual o leitor possa não ter um conhecimento prévio. Enquanto os textos narrativos são muito presentes no cotidiano das crianças desde muito cedo e requerem para a sua compreensão a utilização de habilidades básicas de reconhecimento de palavras, os expositivos são frequentes no contexto acadêmico, uma vez que apresentam uma linguagem formal e exigem um maior uso de habilidades superiores, como inferência, planejamento e organização, além de entendimento sobre conceitos (Eason, Goldberg, Young, Geist, & Cutting, 2012; Leslie & Caldwell, 2001; Pilonieta & Medina, 2009). Para Williams, Hall, e Lauer (2004), o conhecimento sobre a organização e a estruturação de um texto é importante porque ele auxilia o leitor a organizar seu conteúdo e a construir o significado da informação transmitida.

Uma avaliação da compreensão de leitura envolve o teste de todos esses fatores listados, isolados ou em combinação, o que torna muito complexa a quantificação desse construto. Além disso, sobre as capacidades medidas, enquanto muitos instrumentos são fortemente influenciados pelas habilidades de compreensão oral, outros são influenciados pela decodificação. Keenan, Betjemann e Olson (2008), ampliando a pesquisa de Nation e Snowling (1997) sobre a apreciação de dois testes de compreensão de leitura britânicos, analisaram cinco instrumentos comumente usados na pesquisa e na clínica nos Estados Unidos. As medidas investigadas variavam em formato (leitura silenciosa e leitura oral), em tamanho (de leitura de sentenças individuais a passagens longas) e nas tarefas utilizadas para testar a compreensão [seleção da figura que melhor representa o significado da sentença lida, preenchimento de lacunas de palavras omitidas de um texto, reconto de passagens lidas, questões de múltipla escolha ou respostas abertas a perguntas literais (explicitadas no texto) e/ou inferenciais (deduzidas a partir das ideias sugeridas no texto)]. Ambos estudos demonstraram que os instrumentos analisados medem habilidades de compreensão diferentes e também a mesma habilidade de maneiras distintas. Ademais, Keenan et al. (2008) evidenciaram que o mesmo

instrumento pode avaliar capacidades diferentes dependendo do nível de desenvolvimento da leitura. Há ainda a possibilidade levantada por Keenan et al. (2014) de que um indivíduo seja considerado como tendo baixas habilidades em compreensão quando na verdade não as possui e vice-versa. Daí a recomendação de que a avaliação de facetas específicas da compreensão e das fases de domínio dessa habilidade seja conduzida por meio de múltiplas medidas e em diferentes períodos do desenvolvimento da leitura.

Esse pequeno relato indica o quão complexo é a questão da avaliação da compreensão da leitura e o cuidado que se deve ter ao elaborar e utilizar um instrumento para esse fim, o que se torna especialmente relevante no Brasil, onde há poucos testes na área. Um desses instrumentos é a Prova de Compreensão de Texto, que compõe a bateria denominada Provas de Avaliação dos Processos de Leitura (PROLEC; Capellini, Oliveira, & Cuetos, 2012a). Essa bateria é a mais completa publicada no Brasil para avaliar a capacidade de leitura de crianças do Ensino Fundamental. As PROLEC são amplamente utilizadas na clínica fonoaudiológica e na psicopedagógica para o rastreamento de distúrbios específicos de aprendizagem, em especial a dislexia, e também como um guia para orientar programas de recuperação. Além dessa utilidade clínica, segundo seus autores, é também recomendada para pesquisadores da área da saúde e da educação, sendo já utilizada em diferentes estudos (Germano, Reilhac, Capellini, & Valdois, 2014; Macedo, Santos, Oliveira, & Martins-Reis, 2015; Machado & Almeida, 2014; Nalom, Soares, & Cárnio, 2015).

As PROLEC são uma adaptação da 5ª. edição da *Bateria de Evaluación de los Procesos Lectores* (Cuetos, Rodrigues, & Ruano, 2004), cuja primeira edição foi desenvolvida na Espanha há 20 anos (Cuetos, Rodrigues, & Ruano, 1996). A bateria é composta de nove provas de aplicação individual que avaliam, em escolares de seis a doze anos, quatro processos de leitura: (a) reconhecimento de letras; (b) reconhecimento de palavras; (c) processos sintáticos; e (d) processos semânticos. Essa 5ª. edição, além do cômputo dos acertos, passou a registrar

o tempo de realização das provas, porém essa análise cronológica não foi adicionada à versão brasileira.

A Prova de Compreensão de Texto (PROLEC-T) [seguindo o nome dessa prova já utilizado na literatura (Nalom et al., 2015)], a última prova das PROLEC, é considerada uma das mais complexas da bateria e a que melhor discrimina os bons e maus leitores (Capellini, Oliveira, & Cuetos, 2012b; Cuetos et al., 2004). A versão espanhola original, constitui-se de quatro histórias, desconhecidas pelas crianças, que variam em igual número em discurso narrativo e expositivo. Em cada uma dessas categorias textuais há uma história curta (com aproximadamente 90 palavras) e outra média (com mais de 130 palavras), sobre as quais as crianças devem responder perguntas que avaliam a sua compreensão de leitura. São quatro perguntas abertas para cada história, todas de natureza inferencial, o que significa que é necessário que as informações sejam deduzidas a partir de ideias implícitas no texto.

A adaptação brasileira da PROLEC-T se difere da versão original por apresentar histórias conhecidas pelas crianças, curtas (com número de palavras variando de 49 a 59) e por avaliar a compreensão por meio de perguntas literais. Este tipo de pergunta, por demandar como resposta informações explicitadas no texto, exige apenas uma busca em sua microestrutura (nível local). Em contraste, as perguntas inferenciais requerem processos cognitivos mais sofisticados, tornando-as mais difíceis de serem respondidas do que as literais (Buarque, Higino, Miranda, Dubeux, & Pedrosa, 1992; Oakhill, 1984).

O primeiro teste da adaptação brasileira das PROLEC foi realizado por Oliveira e Capellini (2010). A bateria foi aplicada em uma amostra de 262 alunos da 1ª a 4ª série (atuais 2º ao 5º ano) do Ensino Fundamental de duas escolas (uma pública municipal e a outra particular) da cidade de Marília, São Paulo. Constatou-se que os alunos da primeira escola ($n = 140$) tenderam a obter escores (médias aritméticas) inferiores aos da escola particular ($n = 140$). Esse estudo teve duas consequências importantes: (a) a recomendação da utilização do instrumento, considerado uma adequada adaptação à realidade brasileira,

na avaliação dos processos de leitura de nossas crianças; e (b) a transformação de seus resultados em normas no manual da bateria publicada no Brasil em 2010 e 2012, com a colaboração do primeiro autor da versão espanhola original do instrumento. Assim, os escores dos alunos da escola municipal e particular do estudo de Oliveira e Capellini (2010) foram transformados em normas populacionais. No entanto, é digno de nota que a forma de seleção da amostra de escolas e de crianças não é apresentada no estudo de referência e nem no manual do instrumento. A omissão dessas informações aliada ao número insuficiente de escolas e à baixa média de crianças por ano escolar (aproximadamente 66), certamente dá margem ao questionamento da validade do instrumento e de suas normas, sem nem mesmo entrar no mérito do escopo da generalização de tais normas.

O presente trabalho objetiva discutir a validade da adaptação brasileira da PROLEC-T, elaborada para testar os processos semânticos, via análise de validação externa e Teoria de Resposta ao Item (TRI) em uma amostra de alunos de escolas estaduais do estado de Minas Gerais. Objetiva também ponderar a relevância da normatização do instrumento publicada no manual (Capellini et al., 2012a) e de subsequentes normas regionais.

Método

Participantes

Conforme a listagem das escolas cadastradas em Belo Horizonte, fornecida pela Superintendência Regional de Ensino, foram sorteadas oito escolas estaduais, distribuídas em diferentes regionais da cidade, não havendo recusa de nenhuma delas em participar da pesquisa. Embora se tenha respeitado a proporção de tipos de escolas (estadual, municipal e particular), somente a Secretaria Estadual de Educação autorizou a realização da pesquisa.

A amostra contou com 457 alunos voluntários (49,7% meninos) do 2º ao 5º ano do Ensino Fundamental. Em cada uma das 77 salas de aula que participaram do estudo, seis alunos foram selecionados, de forma aleatória (por meio de

sorteio a partir da lista de presença), para responder aos instrumentos da pesquisa. Não houve critério para a inclusão e a exclusão de crianças da amostra (cuja análise descritiva é apresentada na Tabela 1) e nenhuma desistência em participar do estudo foi registrada. A média aritmética da capacidade cognitiva geral das crianças, avaliada por meio das Matrizes Progressivas Coloridas de Raven (CPM), é classificada como média superior, uma vez que situou-se no Percentil 73. O termo de consentimento livre e esclarecido foi assinado por todos envolvidos na pesquisa, e o Comitê de Ética da Universidade Federal de Minas Gerais aprovou todos os seus procedimentos (CAAE: 17754514.6.0000.5149).

Instrumentos

As CPM (Angelini, Alves, Custódio, Duarte, & Duarte, 1999) avaliam a capacidade cognitiva geral de crianças de 5 a 11 anos de idade. O teste é constituído por 36 itens, divididos em três conjuntos de 12 itens (A, Ab, B) e organizados de acordo com uma dificuldade crescente intra e interconjuntos. Para ser bem-sucedida, a criança deve completar uma figura exposta na parte superior de uma folha com uma das seis opções impressas abaixo dela, de forma espacial e lógica, o que envolve a compreensão de que as imagens são caracterizadas por suas diferenças, similaridades, identidade, mudança, simetria, orientação.

Na adaptação brasileira da PROLEC-T, as histórias foram retiradas da coleção “Português: uma proposta para o letramento” (Soares, 1999). A prova investiga se o estudante é capaz de extrair o significado do que lê e de integrá-lo ao seu conhecimento (Capellini et al., 2012a). A sua versão original é composta por quatro histórias (que também chamaremos de textos), de aplicação individual e em uma única sessão. Cada texto possui quatro perguntas, sendo 16 o número total de questões do instrumento, com cada uma delas podendo ser classificada como correta (valendo 1 ponto) ou incorreta (valendo zero). De acordo com o manual da PROLEC-T, os Textos 1 (“Cachorro ciclista”) e 2 (“Casa de caixotes”) são expositivos, e os Textos 3 (“A raposa e a uva”) e 4 (“Férias de Patrícia”) são narrativos.

Não há, nesse documento nenhuma informação sobre os tipos de perguntas da prova. No entanto, com base nas definições de perguntas literais e inferenciais aqui apresentadas, conclui-se que todas as perguntas da PROLEC-T são literais.

O Teste de Leitura: Compreensão de Sentenças (TELCS; Vilhena, Sucena, Castro, & Pinheiro, 2016) é composto por 40 sentenças isoladas (de 8 a 20 palavras cada) cuja última palavra é sempre omitida. São oferecidas cinco palavras (estilo múltipla escolha), com apenas uma alternativa que dá sentido à sentença. As alternativas podem possuir entre si semelhanças visuais (ex., mesmo número de letras), semelhanças fonológicas (ex., começarem com o mesmo fonema (aliteração) ou terminarem com a mesma rima) ou semelhanças semânticas (ex., pertencerem a uma mesma categoria semântica, como tipo de profissão). Pontua-se um ponto para a resposta correta e zero para a incorreta. O TELCS apresentou robusta consistência interna, com elevado *alpha* de Cronbach (0,97) e efeitos significativos ($p < 0,01$) de escolaridade ($2^\circ < 3^\circ < 4^\circ < 5^\circ$ ano) e idade ($7 < 8 < 9 < 10 < 11$ anos). Validação concorrente foi provida pelas fortes correlações de Pearson verificadas entre os escores do TELCS com as taxas de acurácia das PLP e das PLPP ($r = 0,84$ e $0,79$) e com um índice geral que agregou todas as medidas de leitura ($r = 0,84$).

A Prova de Leitura de Palavras (PLP) e a Prova de Leitura de Pseudopalavras (PLPP) foram validadas por Cogo-Moreira, Ploubidis, de Avila, de Jesus Mari, e Pinheiro (2012). Cada prova contém 88 itens de teste e 4 de treinamento, que devem ser lidos em voz alta, gerando dois tipos de escores: porcentagens de acertos e taxa de acurácia (número total de itens lidos corretamente por minuto). As palavras variam em termos de frequência de ocorrência, de regularidade grafema–fonema e de número de letras. As pseudopalavras foram construídas a partir das palavras, sendo que cada uma mantém a mesma estrutura ortográfica e número de letras da palavra de referência. Os itens de ambas as listas são impressos em cartões plastificados contendo 12 linhas com aproximadamente sete palavras cada, em fonte Arial tamanho 14. As correlações positivas e significativas dessas provas com o TE-

LCS oferecem índices adicionais de validade a elas, além dos obtidos por Cogo-Moreira et al. (2012). Nesse estudo, as duas provas [aplicadas em uma amostra de 617 alunos, $M = 9,8$ anos, $DP = 1,0$, de 10 escolas públicas, aleatoriamente selecionadas, do ensino fundamental da cidade de São Paulo] apresentaram alta correlação entre si ($r = 0,92, p < 0,001$) e com a acurácia de leitura de texto (PLP: $r = 0,92, p < 0,001$; PLPP: $r = 0,873, p < 0,001$). Finalmente, a análise de regressão, ajustada para os *clusters* das 10 escolas de onde os alunos da amostra eram provenientes, mostrou um efeito de escolaridade significativo tanto na PLP quanto na PLPP, oferecendo assim mais uma validação para as provas.

Procedimentos de Coleta de Dados

Os instrumentos foram aplicados, no mesmo dia, em duas sessões, cada uma com duração média de 15 minutos. Na primeira, aplicaram-se as CPM e o TELCS de forma coletiva, em grupos de no máximo 10 crianças. Já na segunda sessão, aplicaram-se individualmente a PROLEC-T, a PLP e a PLPP.

As CPM foram apresentadas como quebra-cabeças. Os dois primeiros itens foram introduzidos coletivamente e de forma explícita, com os subsequentes itens respondidos sem auxílio. As crianças do 2º ano foram avaliadas individualmente e as do 3º ao 5º ano em grupo. O tempo de aplicação foi livre e não houve casos de duração de mais de 12 minutos.

Para o TELCS, houve uma fase de treino com quatro sentenças. As demais 36 sentenças, lidas de forma silenciosa, foram completadas individualmente, o mais rápido possível, dentro de no máximo 5 minutos.

Em relação à PROLEC-T, de acordo com o manual, o conjunto de histórias é aplicado em uma ordem fixa, após a seguinte instrução: “Eu vou apresentar pequenos textos para que você os leia. Leia com atenção porque depois que você terminar de ler, eu farei perguntas sobre eles”. Mais uma vez, não consta no manual se a leitura solicitada deve ser realizada: (a) em voz alta ou silenciosa; (b) em tempo livre ou com limite; (c) apenas uma vez; e (d) com perguntas após cada texto individual ou após o conjunto de textos.

Não há também informação sobre qualquer controle a respeito da familiaridade da criança com os textos do instrumento. Diante dessas limitações, o procedimento adotado na presente pesquisa consistiu em solicitar ao participante que lesse cada uma das histórias silenciosamente, sem limite de tempo, e respondesse de forma oral às perguntas abertas (também feitas oralmente), imediatamente após a leitura de cada texto. Não foi dada à criança a opção de releitura.

As provas PLP e PLPP foram aplicadas em sequência, mas com a ordem de aplicação alterada entre os alunos. Os participantes são solicitados a ler voz alta cada item do cartão de cada prova, começando da primeira linha na direção da esquerda para direita. O tempo de leitura e os erros são registrados pelo aplicador.

Finalmente, o desempenho escolar da amostra foi investigado por meio das notas finais (soma dos quatro bimestres letivos) na disciplina Língua Portuguesa. Apenas quatro escolas permitiram o acesso às notas escolares, o que correspondeu a 316 alunos (69% da amostra total). As notas foram dadas em termos de conceito, que variou de A (conceito máximo) a C (conceito mínimo encontrado).

Análise de Dados

As análises estatísticas foram realizadas por meio do *software* IBM SPSS *Statistics* versão 21.0 e pelo *software* R versão 3.1.2 com uso do pacote LTM. Não se encontrou valores atípicos em nenhum instrumento, ao utilizar a técnica *OutlierLabelingRule* (Tukey, 1977), com valor de g de 2,2 (Hoaglin & Iglewicz, 1987). A averiguação do efeito de escolaridade foi feita via Análise de Variância com um fator (ANOVA) e pelo coeficiente de correlação de Spearman. Para verificar a distribuição dos dados, os valores da assimetria (*skewness*) e da curtose, divididos pelos respectivos valores de desvios padrão, foram avaliados com o critério de significância superior a 1,96 (Cramer & Howitt, 2004). A correlação de Spearman foi também utilizada para averiguar a relação do PROLEC-T com todos os instrumentos de leitura e dados demográficos.

Os parâmetros dos itens da PROLEC-T foram analisados via TRI (Baker & Kim, 2004;

Hambleton, Swaminathan, & Rogers, 1991). Para a obtenção dos índices de discriminação e de dificuldade de cada item, testou-se primeiramente o modelo para respostas dicotômicas (certo/errado) de dois parâmetros (discriminação e dificuldade do item) e depois o de três parâmetros (modelo 3p, que considera adicionalmente probabilidade de acerto ao acaso; Baker & Kim, 2004). A escolha do modelo 2p ou 3p da TRI deve ser realizada considerando-se a natureza das respostas do teste, as evidências da presença de acertos casuais nos dados e a análise da qualidade do ajuste do modelo (Hambleton et al., 1991). Para a análise da validade da estrutura interna, via verificação da unidimensionalidade da PROLEC-T, foi utilizada a análise fatorial a partir da matriz de correlação tetracórica, procedimento indicado para variáveis dicotômicas. Foi adotado o nível de significância de 0,05 para os testes estatísticos.

Resultados

Investigação da Validade da PROLEC-T

A Tabela 1 apresenta as análises descritivas da PROLEC-T publicadas (Oliveira & Capellini, 2010) e as do presente estudo. Neste estudo encontrou-se uma assimetria negativa em todos os anos escolares, o que evidencia que os escores se concentram do lado direito da curva de normalidade, e sugere que as crianças tendem a ter facilidade em responder ao teste. Em relação à curtose, apesar de ocorrer uma distribuição normal (mesocúrtica) no 2º, 3º, 4º e 5º ano escolar, verificou-se uma distribuição leptocúrtica quando os dados são analisados em conjunto, o que demonstra que a população geral tende a obter escores muito próximos (aproximadamente 10 pontos).

Tabela 1.

Desempenho na PROLEC-T: Número de Alunos (n), Médias e Desvio Padrão (DP) da Amostra do Estudo de Referência (Oliveira & Capellini, 2010) e Análise Descritiva da Amostra do Presente Estudo

Ano Escolar	Estudo de Referência				Presente estudo			
	Municipal		Particular		Estadual			
	<i>n</i>	Média (DP)	<i>n</i>	Média (DP)	<i>n</i>	Média (DP)	Assimetria (DP)	Curtose (DP)
2º	24	8,6 (3,2)	37	12,1 (4,1)*	98	8,2 (3,6)	-0,68 (0,24)†	-0,09 (0,48)
3º	33	10,0 (4,3)	34	14,7 (1,1)*	99	9,1 (3,6)	-0,50 (0,24)†	0,07 (0,48)
4º	31	8,6 (3,2)*	34	12,1 (4,1)*	123	10,8 (2,7)	-0,49 (0,22)†	-0,22 (0,43)
5º	34	13,4 (1,8)*	35	15,0 (1,1)*	137	11,1 (2,7)	-0,54 (0,21)†	-0,09 (0,41)
2º-5º	122	–	140	–	457	10,0 (3,3)	-0,76 (0,11)†	0,52 (0,23)†

Notas. * Comparação entre grupos do estudo de referência vs. presente estudo com valor $p < 0,05$; † Divisão da assimetria ou curtose pelo desvio padrão significativo acima de 1,96.

Os escores da PROLEC-T apresentaram fraca correlação (coeficiente de Spearman) com o ano escolar ($\rho = 0,33$), com a idade da criança ($\rho = 0,33$), e com suas notas na disciplina de Língua Portuguesa ($\rho = 0,34$), com nenhuma correlação com o sexo. As correlações com as medidas de leitura variaram de $\rho = 0,34$ a 0,52 (Tabela 2). Sobre o efeito de escolaridade, com

o objetivo de localizá-lo, utilizou-se adicionalmente a ANOVA (corrigida por Bonferroni): os escores das crianças do 2º e do 3º, assim como os das crianças do 4º e do 5º ano não foram significativamente distintos. Pode-se esquematizar essa relação da seguinte forma: $[(2^\circ = 3^\circ) < (4^\circ = 5^\circ)]$.

Tabela 2
Coefficiente de Correlação de Spearman entre o PROLEC-T e Dados Demográficos, Educacionais e Avaliações de Leitura

Variáveis	PROLEC-T	
Escolaridade	0,33*	
Idade	0,33*	
Nota em Língua Portuguesa	0,34*	
TELCS	0,52*	
PLP	% de Acertos	0,43*
	Taxa de acurácia	0,46*
PLPP	% de Acertos	0,34*
	Taxa de acurácia	0,42*

Nota. * Nível de significância: $p < 0,001$.

Análise via Teoria de Resposta ao Item (TRI)

Na fase de análise prévia, foi verificado que o percentual de acertos dos itens mais difíceis foi igual ou próximo a zero para o grupo de menor pontuação total. A proporção de acertos foi também próxima de zero, em todos os itens, para os grupos com faixas de pontuação total na região inferior da escala. Estas características são evidências da ausência de respostas ao acaso. Para verificar a adequação do ajuste dos dados ao modelo 2p, foi efetuada uma comparação com o ajuste ao modelo 3p, por meio da verificação dos valores do critério de informação Bayesiano (BIC) e critério de informação de Akaike (AIC). O modelo 2p para respostas dicotômicas (BIC = 7315,2; AIC = 7183,8) mostrou-se mais adequado para ajuste aos dados que o de três parâmetros (BIC = 7411,4; AIC = 7214,3), por apresentar menores valores nos dois critérios. Esses resultados estão de acordo com a natureza da prova, que contém perguntas abertas e respostas verbais livres (tabuladas como certo ou errado), cuja probabilidade de respostas ao acaso não é relevante. Desta forma, devido à natureza dicotômica das respostas dos itens da PROLEC-T e à análise prévia dos dados, a probabilidade de acerto ao acaso foi considerada desprezível e a análise foi efetuada com base nos parâmetros de discriminação e dificuldade dos itens.

Para a análise da validade da estrutura interna e do modelo da TRI a ser utilizado, foi verificado se houve a presença de uma única dimensão latente. No caso de testes com respostas dicotômicas, o número de dimensões pode ser avaliado por meio dos autovalores da matriz de correlações tetracóricas. Os seis primeiros autovalores encontrados na PROLEC-T foram: 6,53; 1,36; 1,16; 1,08; 0,95; 0,79. A predominância do primeiro autovalor em relação aos demais é forte evidência de unidimensionalidade da prova. Além disso, em teste estatístico baseado na matriz de correlação tetracórica, foi verificado que a hipótese nula de unidimensionalidade é adequada ($valor\ p = 0,158$). De acordo com esses resultados, a PROLEC-T pode ser considerada um instrumento com um único fator dominante. A Tabela 3 expõe os valores dos parâmetros de discriminação e dificuldade dos itens obtidos após o ajuste da matriz de respostas ao modelo da TRI.

Tabela 3
Parâmetros de Dificuldade e de Discriminação dos Itens do PROLEC-T

	Questão	Dificuldade	Discriminação
Texto 1	I1.1	-1,737	1,944
	I1.2	0,141	0,699
	I1.3	0,496	1,196
	I1.4	0,498	0,723
Texto 2	I2.5	-0,546	1,407
	I2.6	0,197	0,915
	I2.7	1,002	1,546
	I2.8	1,723	0,867
Texto 3	I3.9	-0,613	1,876
	I3.10	-1,974	1,460
	I3.11	-0,653	1,895
Texto 4	I3.12	-1,335	2,340
	I4.13	-1,731	1,137
	I4.14	-1,582	1,508
	I4.15	-1,144	1,292
	I4.16	-0,479	0,954

As curvas informação, do modelo TRI dos itens que compõem a PROLEC-T, podem ser vistas na Figura 1. Esse tipo de curva indica as regiões em que a medida revela maior preci-

são. De modo geral, um teste estima com maior precisão a habilidade ou proficiência na região próxima ao pico da curva de informação (Baker & Kim, 2004). A primeira pergunta do Texto 1

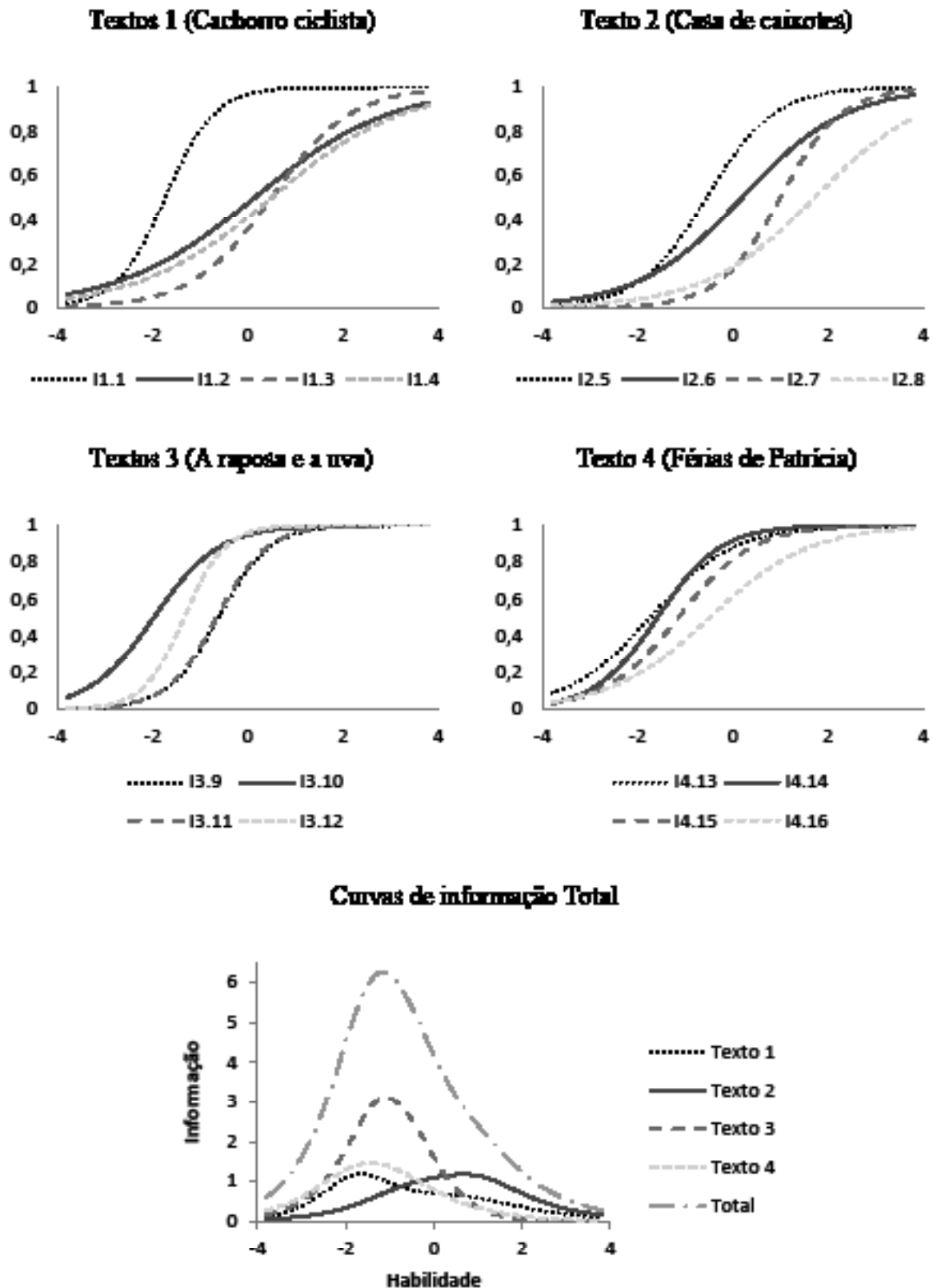


Figura 1. Curvas de informação características de cada item, de cada texto e do escore total do PROLEC-T.

(item I1.1) mostra valor muito baixo no parâmetro de dificuldade do item, com probabilidade de acerto, no ponto médio da escala, até duas vezes e meia maior do que as outras três subseqüentes questões, que possuem um nível de dificuldade homogêneo entre si. O Texto 2 se caracteriza por um conjunto de itens com maiores parâmetros de dificuldade e um item particularmente mais difícil (item I2.8). O Texto 3 contém o item de menor dificuldade (item I3.10). Já o Texto 4 apresenta boa homogeneidade entre os itens, com uma leve tendência de dificuldade crescente em toda a região da escala.

A Figura 1 também mostra a curva de informação de cada texto e a curva de informação total da prova. É possível verificar que os Textos 3 e 4 apresentam tendência a concentrar informação na região da escala de habilidades médias e abaixo da média (região [-2, 0]), com pouca precisão para medir habilidades acima da média. Adicionalmente, a curva de informação total da prova, que é a soma das informações individuais dos itens, mostra uma tendência de concentrar informação na região [-2, 0]. As curvas características de cada texto e os valores da Tabela 3 confirmam, pela dificuldade média do conjunto, a ordenação: Texto 3, Texto 4, Texto 1 e Texto 2, por ordem crescente de dificuldade.

Discussão

O presente estudo, na busca de instrumentos válidos e normatizados para avaliar os processos de compreensão de leitura, analisou a Prova de Compreensão de Texto das PROLEC, aqui também chamada de PROLEC-T. A tarefa de estabelecimento de validade e de normas, além de sua relevância teórica e prática, é especialmente importante mediante as inúmeras limitações dos testes disponíveis para a avaliação dos processos básicos de leitura (Noronha, Primi, & Alchieri, 2004; Ottati & Noronha, 2013) e a carência de instrumentos especificamente voltados para a compreensão de texto.

Esse estado da arte na área, no Brasil, dificulta a pesquisa básica e a construção de novas ferramentas que, por sua vez, necessitam em seu processo de validação serem correlacionadas

com outras que medem o mesmo construto. Ademais, o maior prejuízo ocorre na clínica, uma vez que a falta de testes validados induz a procedimentos intuitivos e inadequados, que desrespeitam a tendência atual de uma *prática baseada em evidências* que, entre outros propósitos, recomenda a integração entre a experiência do profissional e o conhecimento cientificamente comprovado de maneira a tornar o exercício clínico o mais objetivo possível, conferindo às avaliações e intervenções terapêuticas eficácia e segurança (Daher, 2006; El Dib & Atallah, 2006).

O corrente estudo tem o mérito de ser o primeiro a submeter a PROLEC-T a um processo de busca por suas fontes de evidências de validade, o que foi feito de acordo com os *Standards for Educational and Psychological Testing* (American Educational Research Association, American Psychological Association, National Council on Measurement in Education, & Joint Committee on Standards for Educational and Psychological Testing, 1999). Isso significa que, partindo de amostras de escolas e de crianças selecionadas de forma randômica e com equivalente número de meninos e meninas, o exame da validação da prova ocorreu via diversas fontes, combinando o procedimento tradicional de validação de instrumentos por meio de validade da estrutura interna e externa, com a utilização da TRI. No entanto, é importante ressaltar que, por ser a referida amostra oriunda apenas de escolas estaduais da cidade de Belo Horizonte, as conclusões do presente estudo são limitadas, não podendo ser generalizadas às populações fora dessa região.

Assim, na população considerada, foi possível verificar fortes evidências de que a PROLEC-T apresenta baixo nível de dificuldade e, revela, portanto, pouca variabilidade de escores, o que produz um limitado efeito de escolaridade e baixas correlações de seus escores com o ano escolar e a idade da criança, evidenciando não possuir validades interna e externa satisfatórias. Quanto à sua validação concorrente, os resultados do instrumento não foram melhores no que se refere à sua correlação com as notas das crianças na disciplina de Língua Portuguesa, embora tenha apresentado corre-

lação mediana com as porcentagens de acertos das provas de reconhecimento de palavras e de pseudopalavras e com o TELCS.

Reforçando esses achados, vale citar que, em um estudo recente (Nalom et al., 2015), conduzido com crianças de escolas públicas da cidade de São Paulo, as provas da PROLEC-T e uma Prova de Compreensão de Sentenças, da própria PROLEC, não se correlacionaram. Esse resultado é surpreendente, uma vez que as duas provas se encontram na mesma categoria, ou seja, testam os processos semânticos, que subjazem aos processos de compreensão.

Tomando agora a validade da PROLEC-T investigada via TRI, o primeiro resultado mostra que a prova é unidimensional. Possuir um único fator dominante é congruente com o fato, já levantado, de o instrumento possuir apenas perguntas literais. No entanto, uma vez que o tipo de texto narrativo *versus* expositivo foi um dos fatores controlados, era de se esperar a presença de fatores relevantes para cada tipo de texto.

Com relação à dificuldade dos itens e das histórias, os itens I1.1 (“O que o cachorro sabe fazer?”), I3.10 (“O que a raposa ficou com vontade de fazer quando viu as uvas?”) e 3.12 (“A raposa conseguiu apanhar as uvas?”) demonstraram menor dificuldade entre as 16 perguntas da PROLEC-T. No item I3.10, por exemplo, é possível inferir a resposta correta sem ter lido o texto, uma vez que a resposta é “Apanhá-las para comer”. Em conjunto, esse tipo de perguntas e a familiaridade com a história “A raposa e a uva” podem ter sido a razão de ela ter sido identificada na análise da TRI como a mais fácil. Já no que se refere à discriminação dos itens, é surpreendente que apenas dois itens mostraram poder de discriminação para os leitores com habilidade acima da média. Mesmo assim, um deles, o item I4.16, representa a única pergunta que não segue a ordem cronológica de apresentação dos fatos no texto, uma vez que deveria estar na posição I4.14.

A explicação para esses resultados pode ser encontrada na escolha dos textos e na formulação das perguntas sobre eles. A esse respeito, está explícito no manual da versão original das PROLEC que os textos devam ser desconhecidos

para as crianças e que as perguntas sobre eles devam exigir a utilização de inferências sobre seus conteúdos. Com isso, evita-se que o desempenho da criança seja influenciado por seus conhecimentos prévios e pela utilização da memória de longo prazo, não lhe exigindo a demonstração da capacidade de dedução de informação a partir do conteúdo das histórias. A adaptação brasileira ao utilizar o texto “A raposa e as uvas”, muito conhecido por nossas crianças, e perguntas literais, se distancia da versão tomada como referência. Assim, o fato desse texto ter sido o mais fácil dos quatro textos da prova não é surpreendente.

Ademais, o significativo papel das inferências na compreensão de textos é consistentemente afirmado na literatura [ver Spinillo (2008) para um resumo dos estudos na área]. A ausência do controle da familiaridade de textos e da utilização de perguntas inferenciais é uma importante limitação da PROLEC-T, uma vez que as perguntas literais apenas demandam que o leitor resgate de sua memória as informações explícitas no texto. Assim, não é exigido que ele tenha um papel ativo e que integre a sua experiência de vida com as novas informações textuais. A habilidade de construir inferências foi considerada o fator determinante na discriminação entre crianças jovens que se diferenciavam apenas na capacidade de compreensão, mesmo após o controle do fator memória, revelando ser as perguntas que demandam inferências muito mais complexas do que as que exigem apenas a identificação de informações já apresentadas no texto (Oakhill, 1984).

Outras diferenças entre as duas versões, já assinaladas, referem-se ao menor tamanho das histórias e à ausência de tempo de realização dos itens da prova na versão adaptada. Essas limitações podem ser tomadas como outras causas da baixa validade da PROLEC-T. Finalmente, caso os autores do instrumento decidam atualizá-lo, sugere-se a utilização de textos com conteúdo novos, com tamanhos equivalentes à versão original espanhola e com questões inferenciais. Durante a seleção dos textos e das perguntas, deve-se inicialmente buscar evidências de validade baseadas no conteúdo, como a consulta a juízes especialistas, para julgar a relação da pro-

va com as facetas do construto compreensão de texto. Além disso, a sequência dos textos e das perguntas deve obedecer a uma ordem crescente de dificuldade e que seja aplicada com o controle do tempo de execução.

O controle do tempo é importante porque a velocidade é um dos componentes da leitura fluente, que, segundo Wolf e Katzir-Cohen (2001), é um conceito nada trivial. Isso porque a fluência ocorre como resultado do desenvolvimento do processamento eficaz e rápido em todos os componentes da leitura. Ou seja, ela é obtida por meio da sincronização de cada processo (conhecimentos fonológico, ortográfico, semântico e morfosintático) e de cada unidade (letra, padrão ortográfico, palavra, sentença e passagem) subjacente à leitura, cujo conjunto é coordenado pelas funções executivas (Berninger, Abbott, Billingsley, & Nagy, 2001), sendo uma delas a antecipação do conteúdo, que torna a leitura fluente (Wood, Flowers, & Grigorenko, 2001).

A relação entre a fluência de leitura e a compreensão tem sido documentada por muitos estudos, sendo consenso que a fluência permite a liberação de recursos – atenção, memória de trabalho – para serem direcionados para a compreensão (Adlof, Catts, & Little, 2006). Com a acurácia no reconhecimento de palavras, a atenção se desloca dos processos de decodificação (de nível mais baixo) para os de compreensão (de nível mais alto; LaBerge & Samuels, 1974). Dinâmica semelhante ocorre com a memória de trabalho. Por ela ser limitada, quando recrutada para a decodificação, deixa de atender a compreensão (Van Dyke & Shankweiler, 2013). Então, uma leitura imprecisa e lenta se torna um fator de risco para a compreensão.

No que refere à normatização da prova, pode-se afirmar que ela não seguiu diretrizes nacionais e internacionais (*American Educational Research Association et al.*, 1999; Hambleton, 2009; Resolução CFP nº 002, 2003). Por exemplo, de acordo com os critérios de Avaliação da Qualidade de Testes Psicológicos da Resolução do Conselho Federal de Psicologia Nº 002/2003, uma amostra de um estudo de normatização deve ser descrita “de maneira clara e exaustiva, preferencialmente comparando com estimativas nacionais, possibilitando o julgamento do nível

de representatividade do grupo de referência usado para a transformação dos escores”. Quanto ao número de sujeitos, exceto para os testes projetivos, a resolução considera 300 participantes como um número razoável.

Assim, tudo o que pode ser dito sobre a “normatização” da PROLEC-T é que ela oferece apenas “médias locais preliminares”², o que parece suficiente, considerando os índices de validade do instrumento. Todos esses resultados sobre a prova em pauta são muito preocupantes por duas razões:

1. seu uso frequente em estudos sobre a análise do processamento semântico de crianças do 2º ao 5º ano do Ensino Fundamental [no estudo supracitado (Nalom et al., 2015), a PROLEC-T foi inclusive considerado “padrão-ouro”]; e,
2. o fato de ela ser um instrumento comercializado por uma editora de testes de reputação nacional.

Em síntese, a PROLEC-T apresenta limitações para se testar o ganho da habilidade de compreensão de leitura ao longo do desenvolvimento. Dessa forma, não se recomenda a sua utilização para o enquadramento do nível de desempenho do aluno de acordo com o seu ano escolar, uma vez que o efeito de escolarização foi descontínuo e com fraca correlação com a idade da criança e com as suas notas na disciplina de Língua Portuguesa. Igualmente, considerando a fraca evidência de validade de construto (ex., ausência de perguntas inferenciais e de consulta a juízes) e as análises da TRI, a normatização da PROLEC-T não é recomendada. A prova, no formato atual, permite distinguir somente as crianças com atraso de leitura das com desempenho típico (Nalom et al., 2015; Oliveira & Capellini, 2010). Caso ela seja reformulada, sugere-se que os textos sejam substituídos por histórias inéditas, que todas as perguntas sejam inferenciais, e que a ordem dos textos e das perguntas seja alterada de maneira a apresentar uma

² Em comunicação pessoal, a autora da versão brasileira do PROLEC nos informou que a adaptação do instrumento foi estritamente regional, devendo ser utilizada com cuidado em relação aos outros Estados Brasileiros.

disposição crescente de dificuldade. O controle da extensão do texto deve ser também uma variável a ser controlada.

Referências

- Adlof, S. M., Catts, H. W., & Little, T. D. (2006). Should the simple view of reading include a fluency component? *Reading and Writing, 19*(9), 933-958. doi:10.1007/s11145-006-9024-z
- American Educational Research Association, American Psychological Association, National Council on Measurement in Education, & Joint Committee on Standards for Educational and Psychological Testing. (1999). *Standards for educational and psychological testing*. Washington, DC: American Educational Research Association.
- 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*. São Paulo, SP: Centro Editor de Testes e Pesquisa em Psicologia.
- Baker, F. B., & Kim, S. H. (2004). *Item Response Theory: Parameter estimation techniques, second edition* (2nd ed.). New York: Marcel Dekker.
- Berninger, V. W., Abbott, R. D., Billingsley, F., & Nagy, W. (2001). Processes underlying timing and fluency of reading: Efficiency, automaticity, coordination, and morphological awareness. In M. Wolf (Ed.), *Dyslexia, fluency, and the brain* (pp. xxv). Timonium, MD: York Press.
- Buarque, L. L., Higino, Z. M., Miranda, E. M., Dubeux, M. H., & Pedrosa, I. (1992). *Avaliação do desempenho da rede pública escolar do Estado de Pernambuco na área de linguagem* (Vol. 5). São Paulo, SP: Fundação Carlos Chagas.
- Capellini, S. A., Oliveira, A. M., & Cuetos, F. (2012a). *PROLEC: Provas de Avaliação dos Processos de Leitura* (S. A. Capellini, A. M. Oliveira, & F. Cuetos, Trans, 2. ed.). São Paulo, SP: Casa do Psicólogo.
- Capellini, S. A., Oliveira, A. M., & Cuetos, F. (2012b). *PROLEC: Provas de Avaliação dos Processos de Leitura - Manual* (S. A. Capellini, A. M. Oliveira, & F. Cuetos, Trans, 2. ed.). São Paulo, SP: Casa do Psicólogo.
- Cogo-Moreira, H., Ploubidis, G. B., de Avila, C. R., de Jesus Mari, J., & Pinheiro, A. M. (2012). EACOL (Scale of Evaluation of Reading Competence by the Teacher): Evidence of concurrent and discriminant validity. *Neuropsychiatric Disease and Treatment, 8*, 443-454. doi:10.2147/NDT.S36196
- Cramer, D., & Howitt, D. (2004). *The Sage dictionary of statistics: A practical resource for students in the social sciences*. Thousand Oaks, CA: Sage.
- Cuetos, F., Rodrigues, B., & Ruano, E. (1996). *Evaluación de los procesos lectores PROLEC*. Madrid: TEA Ediciones.
- Cuetos, F., Rodrigues, B., & Ruano, E. (2004). *Evaluación de los procesos lectores* (5. ed. rev.). Madrid: TEA Ediciones.
- Daher, W. (2006). Medicina baseada em evidências. *Revista Brasileira de Cirurgia Cardiovascular, 21*, III-IV.
- Eason, S. H., Goldberg, L. F., Young, K. M., Geist, M. C., & Cutting, L. E. (2012). Reader-text interactions: How differential text and question types influence cognitive skills needed for reading comprehension. *Journal of Educational Psychology, 104*(3), 515-528. doi:10.1037/a0027182
- El Dib, R. P., & Atallah, A. N. (2006). Fonoaudiologia baseada em evidências e o Centro Cochrane do Brasil. *Sao Paulo Medical Journal, 124*(2), 51-54.
- Germano, G. D., Reilhac, C., Capellini, S. A., & Valdois, S. (2014). The phonological and visual basis of developmental dyslexia in Brazilian Portuguese reading children. *Frontiers in Psychology, 5*, 1169. doi:10.3389/fpsyg.2014.01169
- Hambleton, R. K. (2009). *International Test Commission Guidelines for Test Adaptation, second edition*. Paper presented at the 11th European Congress of Psychology, Oslo.
- Hambleton, R. K., Swaminathan, H., & Rogers, H. J. (1991). *Fundamentals of item response theory*. Newbury Park, CA: Sage.
- Hoaglin, D. C., & Iglewicz, B. (1987). Fine-tuning some resistant rules for outlier labeling. *Journal of the American Statistical Association, 82*(400), 1147-1149. doi:10.1080/01621459.1987.10478551
- Kamhi, A. G. (2009). The case for the narrow view of reading. *Language, Speech, and Hearing Services in Schools, 40*(2), 174-177. doi:10.1044/0161-1461(2009/08-0068)
- Keenan, J. M., Betjemann, R. S., & Olson, R. K. (2008). Reading comprehension tests vary in the skills they assess: Differential dependence on decoding and oral comprehension.

- Scientific Studies of Reading*, 12(3), 281-300. doi:10.1080/10888430802132279
- Keenan, J. M., Hua, A. N., Meenan, C. E., Pennington, B. F., Willcutt, E., & Olson, R. K. (2014). Issues in identifying poor comprehenders. *L'Année Psychologique*, 114(4), 753-777. doi:10.4074/S0003503314004072
- Kintsch, W., & van Dijk, T. A. (1978). Toward a model of text comprehension and production. *Psychological Review*, 85(5), 363-394.
- LaBerge, D., & Samuels, S. J. (1974). Toward a theory of automatic information processing in reading. *Cognitive Psychology*, 6(2), 293-323. doi:http://dx.doi.org/10.1016/0010-0285(74)90015-2
- Leslie, L., & Caldwell, J. (2001). *Qualitative reading inventory*, 3. New York: Longman.
- Macedo, A. A., Santos, J. N., Oliveira, A. G., & Martins-Reis, V. O. (2015). Programa fonoaudiológico de promoção do letramento (pfpl): Eficácia na compreensão de leitura em escolares. *Distúrbios da Comunicação*, 27(2), 248-255.
- Machado, A. C., & Almeida, M. A. (2014). O modelo RTI - Resposta à intervenção como proposta inclusiva para escolares com dificuldades em leitura e escrita. *Revista Psicopedagogia*, 31, 130-143.
- McGuinness, D. (2004). *Early reading instruction: What Science really tells us about how to teach reading*. Cambridge, MA: Massachusetts Institute of Technology.
- Nalom, A. F. O., Soares, A. J. C., & Cárnio, M. S. (2015). The relevance of receptive vocabulary in reading comprehension. *CoDAS*, 27, 333-338.
- Nation, K., & Snowling, M. (1997). Assessing reading difficulties: The validity and utility of current measures of reading skill. *British Journal of Educational Psychology*, 67(Pt 3), 359-370.
- Noronha, A. P. P., Primi, R., & Alchieri, J. C. (2004). Parâmetros psicométricos: Uma análise de testes psicológicos comercializados no Brasil. *Psicologia: Ciência e Profissão*, 24, 88-99.
- Oakhill, J. (1984). Inferential and memory skills in children's comprehension of stories. *British Journal of Educational Psychology*, 54(1), 31-39. doi:10.1111/j.2044-8279.1984.tb00842.x
- Oliveira, A. M., & Capellini, S. A. (2010). Desempenho de escolares na adaptação brasileira da avaliação dos processos de leitura. *Pró-Fono Revista de Atualização Científica*, 22, 555-560.
- Ottati, F., & Noronha, A. P. P. (2013). Parâmetros psicométricos de instrumentos de interesse profissional. *Estudos e Pesquisas em Psicologia*, 3(2), 22-31. doi:10.12957/epp.2003.7787
- Pilonieta, P., & Medina, A. L. (2009). Reciprocal teaching for the primary grades: "We Can Do It, Too!". *The Reading Teacher*, 63(2), 120-129.
- Resolução CFP nº 002, de 24 de março de 2003*. (2003). Define e regulamenta o uso, a elaboração e a comercialização de testes psicológicos e revoga a Resolução CFP nº 025/2001.
- Soares, M. B. (1999). *Português: Uma proposta para o letramento*. São Paulo, SP: Moderna.
- Spinillo, A. G. (2008). O leitor e o texto: Desenvolvendo a compreensão de textos na sala de aula. *Interamerican Journal of Psychology*, 42, 29-40.
- Tukey, J. W. (1977). *Exploratory Data Analysis*. Reading, MA: Addison-Wesley.
- Van Dyke, J. A., & Shankweiler, D. P. (2013). From verbal efficiency theory to lexical quality: the role of memory processes in reading comprehension. In M. A. Britt, S. R. Goldman, & J.-F. Rouet (Eds.), *Reading: From words to multiple texts* (pp. xiv). New York: Routledge.
- Vilhena, D. A., Sucena, A., Castro, S. L., & Pinheiro, A. M. V. (2016). Reading Test—Sentence Comprehension: An adapted version of Lobrot's Lecture 3 Test for Brazilian Portuguese. *Dyslexia*, 22(1), 47-63. doi:10.1002/dys.1521
- Williams, J. P., Hall, K. M., & Lauer, K. D. (2004). Teaching expository text structure to young at-risk learners: Building the basics of comprehension instruction. *Exceptionality*, 12(3), 129-144. doi:10.1207/s15327035ex1203_2
- Wolf, M., & Katzir-Cohen, T. (2001). Reading fluency and its intervention. *Scientific Studies of Reading*, 5(3), 211-239. doi:10.1207/S1532799XSSR0503_2
- Wood, F., Flowers, L., & Grigorenko, E. (2001). On the functional neuroanatomy of fluency, or why walking is just as important to reading as talking is. In M. Wolf (Ed.), *Dyslexia, fluency, and the brain* (pp. xxv). Timonium, MD: York Press.

Recebido: 16/11/2015
 1ª revisão: 21/02/2016
 2ª revisão: 23/03/2016
 3ª revisão: 1º/05/2016
 Aceite final: 19/05/2016