

# Construção de uma Prova para Avaliação da Compreensão Leitora no Ensino Fundamental: Estudo Piloto

**Patrícia Silva Lucio<sup>1</sup>**

*Departamento de Psicologia e Psicanálise da Universidade Estadual de Londrina, Londrina, PR, Brasil*

*Programa de Pós-Graduação em Psiquiatria e Psicologia Médica da Universidade Federal de São Paulo, São Paulo, SP, Brasil*

**Adriana de Souza Batista Kida**

*Departamento de Fonoaudiologia da Universidade Federal de São Paulo, São Paulo, SP, Brasil*

**Carolina Alves Ferreira de Carvalho**

*Universidade Federal de São Paulo, São Paulo, SP, Brasil*

**Hugo Cogo-Moreira**

*Departamento de Psiquiatria da Universidade Federal de São Paulo, São Paulo, SP, Brasil*

**Clara Regina Brandão de Ávila**

*Departamento de Fonoaudiologia da Universidade Federal de São Paulo, São Paulo, SP, Brasil*

## Resumo

A compreensão da leitura é o processo pelo qual se extrai significados e se constrói uma representação mental do conteúdo lido, o que requer o bom funcionamento de habilidades cognitivas e linguísticas. No Brasil, há uma carência de instrumentos de avaliação da compreensão leitora que apresentem estudos de validade, fidedignidade e indicadores da qualidade psicométrica dos itens. O presente estudo piloto relata alguns procedimentos psicométricos da construção de uma Prova de Avaliação da Compreensão Leitora para crianças. Alunos do 3º ao 5º ano do Ensino Fundamental de uma escola pública da cidade de São Paulo ( $N = 81$ ) leram sete textos (narrativos e expositivos) e responderam a questões abertas relacionadas. Para a análise dos dados, consideraram-se: (a) os efeitos do texto utilizado, dos processos cognitivos envolvidos na resolução das questões e do ano escolar na precisão das respostas (ANOVAS para medidas repetidas); (b) os parâmetros psicométricos de discriminação e dificuldade dos itens (Teoria de Resposta ao Item). Conforme esperado, as ANOVAS mostraram que a precisão foi menor para os textos mais complexos, para as questões que exigiam uso de inferências, elaboração ou construção de uma representação textual e para as crianças mais jovens. A análise dos itens mostrou que a maioria das questões apresentou índices baixos de discriminação e funcionaram mais para níveis baixos de habilidade. Os resultados indicam que em geral os itens criados mostraram-se promissores para a avaliação da compreensão leitora na faixa etária investigada, mas também apontam a necessidade de ajustes em alguns textos e questões.

**Palavras-chave:** Compreensão da leitura, linguagem escrita, leitura, dificuldade, discriminação.

---

<sup>1</sup> Endereço para correspondência: Universidade Estadual de Londrina, Centro de Ciências Biológicas, Departamento de Psicologia e Psicanálise, Rodovia Celso Garcia Cid, PRr 445 Km 380, Campus Universitário, Caixa-postal: 10011, Londrina, PR, Brasil 86057-970. E-mail: pslucio@gmail.com, adrianabatista@gmail.com, carolcarvalho\_fono@yahoo.com.br, hugocogobr@gmail.com e clara.rbrandao@gmail.com  
Agência de financiamento: Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP).

## Construction of a Task for Assessing Reading Comprehension in Elementary School: Pilot Study

### Abstract

Reading comprehension is the process of extracting meanings and building a mental representation of the read content, which requires the proper functioning of cognitive and linguistic skills. In Brazil, there is a lack of instruments for the assessment of reading comprehension that report validity, reliability, and analysis of the items. This pilot study reports some psychometric procedures of a Task of Reading Comprehension for Children. Students ( $N = 81$ ) from third to fifth year of a public school in the city of São Paulo read seven texts (narrative or expository) and answered open questions. It was analyzed: (a) three main effects on accuracy, namely, the text, the cognitive process used to solve the questions, and the school grade (ANOVA for repeated measures); (b) the psychometric parameters of difficulty and discrimination (Item Response Theory). As expected, the ANOVA showed that the accuracy was lower: (a) for the more complex texts; (b) for three kind of questions: inferential, elaborative, and the ones that required the construction a mental representation of the text; (c) the younger children. Item analysis showed that most questions presented low discrimination indices and functioned to lower levels of ability. The results indicate that, in general, the items proved promising for the assessment of reading comprehension for the age range investigated, but they also point to the need for adjustments in some texts and questions.

**Keywords:** Reading comprehension, written language, reading, difficult, discrimination.

## Construcción de Instrumento para Evaluación de la Comprensión Lectora en la Primaria de la Educación Básica: Estudio Piloto

### Resumen

La comprensión lectora es el proceso por lo cual se extrae el significado y se construye una representación mental del contenido, lo que exige del lector adecuados conocimientos cognitivos y lingüísticos. En Brasil, el restringido número de testes para evaluar la comprensión lectora es conocido y a estos instrumentos les faltan estudios de validez, fiabilidad y presentación de la calidad psicométrica de sus ítems. Aquí se presenta el estudio piloto que reporta los procedimientos psicométricos de construcción de una prueba de comprensión lectora para niños. Estudiantes de tercero a quinto primaria de la Educación Básica de una escuela pública de San Pablo ( $N= 81$ ) leyeron siete textos y contestaron preguntas. El análisis de los datos consideró: (a) efectos del texto, de los procesos cognitivos implicados en la contestación a las cuestiones planteadas y de la escolaridad sobre la precisión de las respuestas (Análisis de varianza para medidas repetidas); (b) parámetros psicométricos de discriminación e dificultad de los ítems (Teoría de Respuestas a Ítems). El análisis de ítems mostró que la precisión fue menor para los textos complejos, para las cuestiones inferenciales, elaborativas o que envolvían la construcción de una representación textual, y además, para los niños más jóvenes. El análisis de los ítems mostró que las cuestiones, en general, presentaran bajos valores de discriminación y se han adaptado a los bajos niveles de habilidad. Los resultados indican que los ítems son prometedores para evaluar la comprensión lectora de niños, pero también sugieren la necesidad de modificaciones en algunos textos y cuestiones.

**Palabras clave:** Comprensión de lectura, lenguaje escrito, lectura, dificultad, discriminación.

Uma parte fundamental do processo de aquisição da leitura é certamente a compreensão do conteúdo lido. Isso implica que a leitura não

pode ser resumida apenas à tradução de grafemas em fonemas, mas deve atingir, como objetivo final, a apreensão do significado do texto pelo

leitor. De uma maneira ampla, Snow e Sweet (2003) definem a compreensão da leitura como o processo de extração e construção do significado, envolvendo três elementos essenciais: o leitor (munido de habilidades, conhecimentos prévios e expectativas no ato da leitura); o texto (ou aquilo que deve ser lido e transmite uma mensagem); e a atividade (que inclui o processo mental envolvido na leitura, assim como suas consequências para o leitor, na forma de experiências de aprendizagem). Snow e Sweet (2003) ressaltam, ainda, que esses três elementos ocorrem sempre dentro de um contexto sociocultural, o qual ao mesmo tempo modula e é modulado pelo leitor.

Dificuldades no desenvolvimento das habilidades de compreensão leitora são relativamente comuns em crianças em idade escolar (Catts & Weismer, 2006; Oakhill, Cain, & Bryant, 2003; Snowling & Stackhouse, 2004). A avaliação dessas dificuldades exige o uso de instrumentos confiáveis, contendo indicadores empíricos de validade, fidedignidade e normas. Sendo reconhecida a escassez de trabalhos brasileiros que apresentam esses indicadores, justifica-se a relevância de pesquisas que busquem proporcionar aos clínicos, profissionais da educação e pesquisadores de nosso país instrumentos com tais características psicométricas. Assim sendo, o presente trabalho relata os principais resultados das etapas iniciais da criação de um instrumento de avaliação da compreensão leitora para crianças dos anos iniciais do Ensino Fundamental, a Prova de Avaliação da Compreensão Leitora (PACL).

A compreensão da leitura pressupõe o funcionamento complexo e orquestrado de diferentes habilidades linguísticas e processamentos cognitivos (Cain, Oakhill, & Bryant, 2004) subjacentes a qualquer atividade de compreensão verbal. A compreensão leitora é simultaneamente processo e produto de um mesmo fenômeno geral (Cunningham & Cunningham, 1981). Enquanto processo, depende, inicialmente, do acurado – e automático – reconhecimento das palavras que compõem o texto. Com isso, ela congrega desde mecanismos cognitivos (ligados ao reconhecimento, organização, integração e associação de informações aos vários

tipos de memória, à percepção, ao conhecimento vocabular e de regras da língua); até o uso de estratégias metacognitivas e metalinguísticas, de elaboração de inferências e de monitoramento da representação mental criada a partir da informação recebida (Cain & Oakhill, 2006, 2008). Enquanto produto, a compreensão leitora resulta da extração de significados e da construção de uma representação mental do conteúdo lido a partir de elementos do texto e do conhecimento prévio do leitor (Yuill & Oakhill, 1996). O conhecimento de mundo dos leitores pertence à memória de longo prazo, a qual é ativada durante a compreensão, instanciando um conjunto de informações como inferências para o alcance do modelo situacional (Kintsch, 1988). Sendo assim, a compreensão da leitura é um processo ativo e estratégico, de busca consciente e voluntária pelo significado, que envolve o desenvolvimento de associações e expectativas elaboradas sobre os eventos que aparecem no texto (van den Broek, Rapp, & Kendeou, 2005).

Existem vários tipos de tarefas para avaliar a compreensão da leitura, que focalizam distintos aspectos do funcionamento cognitivo e linguístico envolvidos neste processo (Brandão & Spinillo, 1998) e apresentam vantagens e limitações. Por exemplo, a Técnica de Cloze, procedimento em que o leitor é instruído a completar palavras previamente suprimidas do texto, pode ser uma tarefa de fácil aplicação, mas se restringe a avaliar a compreensão do ponto de vista local, uma vez que pode ser resolvida compreendendo-se apenas a frase, não sendo necessária a compreensão global do conteúdo lido (Kida, Carvalho, & Avila, 2012). Sua restrição pelo enfoque em habilidades relacionadas ao processamento das informações locais e dos aspectos coesivos do texto, amplamente determinados pelo conhecimento e capacidade de operar com elementos sintáticos (Cain & Oakhill, 2006), bem como a interferência do vocabulário da criança (Snowling & Stackhouse, 2004) no desempenho medido pela tarefa, podem conduzir a falhas na detecção de dificuldades no processamento global do texto.

A tarefa de reconto após leitura possibilita o acesso à expressão direta da representação men-

tal construída pelo leitor a partir das informações do texto lido (Leslie & Caldwell, 2009; Reed & Vaughn, 2012). Permite identificar as informações julgadas essenciais pelo leitor, sua capacidade de integrar e de organizar logicamente estas informações, além de mostrar-se uma tarefa livre, sem indução ou interferência sobre a resposta do avaliado (Bernhardt, 1991). No entanto, a possibilidade de penalização de crianças com dificuldades de expressão, somada às dificuldades relacionadas à análise (transcrição e categorização das respostas) são suas principais desvantagens (Leslie & Caldwell, 2009).

Os testes constituídos por perguntas (sejam estas abertas ou de múltipla escolha) apresentam-se como uma possibilidade de avaliar, mais controladamente, uma gama ampla de mecanismos de compreensão responsáveis tanto pelo processamento de informações literalmente expressas no texto, quanto de informações inferenciais. Estas últimas são tomadas como um importante preditor da compreensão leitora de crianças (Oakhill & Cain, 2012). As questões de múltipla escolha, no entanto, apresentam uma importante limitação, uma vez que permitem ao avaliado responder aleatoriamente a uma questão e acertá-la, sem que sequer tenha extraído o sentido das informações lidas no texto (Cain & Oakhill, 2006; Keenan & Betjemann, 2006). Modelos de Teoria de Resposta ao Item têm controlado esta limitação ao introduzir o parâmetro de “adivinhação” (*guessing*) em suas equações, também chamando de parâmetro *c* (Embretson & Reise, 2000). No entanto, esses modelos não resolvem o problema de entender qual processo levou o indivíduo à escolha da resposta.

Apesar da grande variedade de tarefas para a avaliação e o diagnóstico dos problemas de leitura e de compreensão leitora atualmente disponível no Brasil, em nosso contexto, a condução de pesquisas voltadas para a construção de instrumentos formais e objetivos, baseados em indicadores psicométricos, pode ser considerada recente, o que ainda dificulta o alcance de um diagnóstico e da intervenção precisos. Destacadas raras exceções (Cuetos, Rodrigues, & Ruano, 2012), a maioria das tarefas encontra-se disponível apenas em artigos científicos, dissertações ou

teses como parte de estudos experimentais sobre a leitura ou a escrita (p. ex., Brandão & Spinillo, 1998; Carvalho, Avila, & Chiari, 2009; Galvão & Mahon, 1998; Gomes & Boruchovitch, 2005; Salles & Parente, 2002). De modo geral, estes estudos fornecem apenas evidências de validação de construto para estas tarefas. Sabe-se que estas fontes de validação são preciosas para a delimitação das inferências pretendidas dos escores (Urbina, 2014), mas a coleta de medidas psicométricas diretas sobre os escores de testes, na forma de estudos de validação, fidedignidade e análise dos itens, mostra-se necessária.

As questões anteriormente levantadas a respeito da carência de instrumentos confiáveis e válidos para avaliação da compreensão da leitura em nosso contexto foram motivadoras para a idealização da Bateria de Avaliação da Compreensão Leitora. A Bateria de Avaliação da Compreensão Leitora é um instrumento voltado para a faixa etária de alunos dos anos iniciais do Ensino Fundamental, sendo composto por um conjunto de tarefas que avaliam a compreensão de textos, reconto de histórias lidas, nomeação rápida e leitura de palavras isoladas. O presente trabalho é um estudo preliminar que buscou investigar a adequação psicométrica dos itens (textos e questões) do subteste de compreensão da bateria, composto por textos narrativos e expositivos e por questões abertas a eles relacionadas. Estas questões variaram quanto ao tipo de processo cognitivo demandado para respondê-las, com base nos argumentos de (Yuill & Oakhill, 1996) de que os leitores pobres em compreensão diferem dos leitores típicos pela dificuldade em responder perguntas que demandam diversos tipos de inferência, ou a capacidade de extrair conclusões a partir de evidências ou pistas contextuais (Paul & Elder, 2012). Para a construção do instrumento, uma das hipóteses foi a de que, controlando-se o tipo de questão conferida à criança, poder-se-ia verificar o papel de diferentes processos cognitivos sobre o desempenho em compreensão leitora (Yuill & Oakhill, 1996). Desse modo, baseando-se nesses autores, as questões da PCL foram assim classificadas: questões de compreensão literal, de conexão entre parte do texto, de elaboração a partir das in-

formações lidas, de uso de conhecimento prévio associado a informações literalmente expressas e de generalização do conteúdo lido.

Neste trabalho, pretendeu-se relatar o estudo piloto da Prova de Compreensão Leitora, destinada a escolares do Ciclo I do Ensino Fundamental. Esta etapa teve como objetivo alcançar a adequação psicométrica de seus itens, por meio de três procedimentos distintos: (a) estudo da progressão da proporção de acerto nas questões em função do texto utilizado e do ano escolar dos participantes; (b) análise das diferenças de médias de proporção de acertos nos itens, classificados quanto a diferentes tipos de processos cognitivos; (c) identificação dos níveis de dificuldade e de discriminação dos itens, estabelecidos a partir dos parâmetros da Teoria de Resposta ao Item. Os resultados obtidos nesta etapa do estudo fundamentarão a análise exploratória dos textos e/ou das questões utilizadas, o que será útil na adequação desses elementos para elaborar a versão final do instrumento.

## Método

### Considerações Éticas

Considerando a necessidade de respeitar os aspectos éticos dos projetos de pesquisa envolvendo seres humanos, este estudo foi submetido e aprovado pelo Conselho de Ética em Pesquisa/ Universidade Federal de São Paulo - Hospital São Paulo (UNIFESP-HSP) sob o número de protocolo 38406/2012.

### Amostra

Participaram 81 escolares (59% meninas) do 3º ao 5º anos do Ensino Fundamental ( $N = 25, 32$  e  $24$ , respectivamente) de uma escola da rede pública de ensino de São Paulo. A média de idade foi de 9,26 anos ( $DP 1,04$ ). Como critério de inclusão no estudo, utilizou-se: ausência de queixas ou indicadores de problemas déficits sensoriais, e/ou cognitivos e distúrbios comportamentais e/ou neurológicos e de retenção escolar. Foram avaliados apenas aqueles que entregaram o Termo de Consentimento Livre e Esclarecido assinado pelos pais ou responsáveis.

### Instrumentos

Foram construídos, inicialmente sete textos (dois expositivos e quatro narrativos) e respectivas questões abertas. A escolha por utilizar textos narrativos e expositivos se deveu ao fato de as crianças serem frequentemente expostas a estes estilos textuais no início da escolarização. Tanto os textos quanto as questões são originais, elaborados pelos autores para compor a Prova de Compreensão Leitora. Sob a perspectiva do aumento progressivo da extensão e da complexidade do material linguístico ao longo do instrumento, os textos variaram quanto ao número de palavras, gênero (narrativo ou expositivo) e nível de complexidade (características sintáticas das sentenças e psicolinguísticas das palavras).

Entre nove e 11 questões abertas foram elaboradas para avaliação da compreensão de cada texto lido. Estas buscaram eliciar, a partir da informação textual, diferentes tipos de processos cognitivos implicados na construção da representação mental do texto. Foram classificadas, segundo o mecanismo cognitivo demandado, em:

1. Questões de Compreensão Literal (LIT): elaboradas para avaliar a capacidade de entendimento de informações explícitas do texto;
2. Questões de Inferência por *Text-Connecting* (TC): processo pelo qual o leitor deve integrar informações implícitas para estabelecer a coesão entre diferentes frases ou períodos;
3. Questões de Compreensão por *Gap-Filling* (GAP): processo pelo qual o leitor utiliza seu conhecimento (prévio) para preencher lacunas de informação e integrar proposições do texto;
4. Questões de Inferência Elaborativa (EL): inferências que, apesar de pertinentes, não são essenciais para a compreensão da situação; e Questões de Modelo de Situação (MS): questões que propõem a aplicação pelo leitor de informações existentes no texto, em diferentes situações (dirigidas pelas questões), ou seja, o leitor evoca a situação do texto, liberto das palavras que o compõem (L. Baker & Stein, 1981; Cain & Oakhill, 1999; Sánchez, 2002).

A título de exemplo, reproduz-se abaixo o Texto 1, com as respectivas questões e classificações quanto ao processo cognitivo evocado:

*A menina esperava à janela. / – Olha, lá vem o Téo! Nós podemos tomar o sorvete que você comprou, mamãe? / – Vocês querem de chocolate, Ana? / – Eu sim! Mas ele gosta de morango. / Ana e Téo tomaram sorvete e brincaram no jardim.*

Questões do Texto 1: 1. *Onde a menina esperava?* (LIT); 2. *Quem chegou à casa de Ana?* (TC); 3. *Quem está conversando?* (TC); 4. *Quem gosta de sorvete de morango?* (TC); 5. *Por que Ana teve que pedir para tomar sorvete?* (EL); 6. *Qual o sabor do sorvete que Ana prefere?* (TC); 7. *Para que Ana esperava Téo?* (TC); 8. *Onde Ana e Téo brincaram?* (LIT); 9. *Como estava o tempo nesse dia?* (GAP); 10. *Do que Ana precisa para brincar? Por quê?* (MS).

### Procedimento de Coleta de Dados

As professoras foram treinadas para aplicação do procedimento de avaliação, que ocorreu em horários regulares de aula previamente estabelecidos com a direção e a equipe de educadores. Os textos foram aplicados da seguinte forma: 1º dia: Texto para iniciantes (35 palavras) e Texto 1 (43 palavras); 2º dia: Texto 2 (111 palavras) e Texto 3 (89 palavras); 3º dia: Texto 4 (172 palavras) e Texto 5 (158 palavras); 4º dia: Texto 6 (192 palavras).

A aplicação das tarefas foi feita de forma coletiva. Os textos foram entregues aos escolares pela professora da turma. Solicitou-se aos escolares que fizessem a leitura silenciosa do texto com atenção, de modo a entendê-lo, para responderem por escrito a algumas questões abertas, pertinentes à passagem lida, imediatamente após a leitura. A professora proferiu as perguntas oralmente e, caso necessário, repetiu a questão. Forneceram-se três minutos para a escrita autônoma de cada resposta.

Os escores foram atribuídos às respostas a partir de parâmetros pré-determinados por uma banca três de juízes, os quais criaram um crivo de correção. Às respostas corretas, atribuiu-se um ponto e às respostas incorretas, zero ponto. Os participantes receberam, portanto, um esco-

re por item (acerto ou erro), um escore total por texto (escore máximo variando de nove a 11, conforme o número de questões no texto) e um escore geral, correspondente à soma dos escores em todos os textos (escore máximo de 72 pontos). Estes valores constituíram a base para o estabelecimento das análises estatísticas.

### Análise dos Dados

Foram conduzidas estatísticas descritivas e inferenciais para a avaliação das diferenças entre os anos escolares e entre os textos (análises de variância para medidas repetidas – ANOVAS). Duas hipóteses são apresentadas:

1. A proporção de acertos nos textos será maior para os anos finais, em relação aos iniciais;
2. A proporção de acertos será maior para as questões LIT e TC em relação às demais. Adotou-se o nível de significância de  $\alpha = 0,05$ . A análise dos itens foi realizada para modelos de dois parâmetros da Teoria de Resposta ao Item (TRI), utilizando-se a Máxima Verossimilhança Robusta (MLR) como estimador. Índices de dificuldade e de discriminação foram obtidos para modelos unifatoriais (atestando-se, assim, para a unidimensionalidade dos mesmos).

## Resultados

### Efeitos na Precisão da Amostra

Como a quantidade de questões difere entre os textos, na Tabela 1 apresenta-se a proporção média de acertos, calculada a partir do escore total em cada texto dividido pelo número de questões. Pôde-se observar que, em geral, a precisão nas tarefas foi elevada, havendo níveis medianos de acertos apenas para os Textos 3, 5 e 6. Separando-se o desempenho por ano escolar, tem-se que, conforme esperado, no 3º ano a precisão foi elevada nos textos iniciais (até Texto 2), decaindo nos textos finais. No 4º ano, a proporção de acertos foi mais alta até o Texto 4, atingindo valores medianos nos textos finais (5 e 6). Para os escolares mais velhos, a precisão foi alta em todas as tarefas, com níveis médios de acerto apenas para os Textos 3 e 5.

**Tabela 1**  
**Proporção de Acertos Obtida pela Amostra Geral para Cada Texto e para os Diferentes Anos Escolares**

Texto	Ano Escolar			Amostra geral
	3º	4º	5º	
Iniciante	0,75	0,75	0,75	0,75
Texto 1	0,65	0,66	0,77	0,69
Texto 2	0,68	0,84	0,82	0,79
Texto 3	0,38	0,62	0,64	0,55
Texto 4	0,51	0,77	0,83	0,71
Texto 5	0,42	0,51	0,63	0,52
Texto 6	0,55	0,64	0,75	0,64

Uma ANOVA para medidas repetidas foi conduzida para verificar a existência de diferenças significativas entre o desempenho dos escolares para os diferentes textos. Foram utilizados como fator intrassujeitos os textos e como fator entre sujeitos o ano escolar. Houve um efeito geral do texto utilizado na precisão da amostra,  $F(6, 417) = 42,478, p < 0,001, \eta^2 = 0,362$ , e também do ano escolar cursado pela criança,  $F(2, 75) = 14365, p < 0,001, \eta^2 = 0,227$ . Isto indica que a precisão nos diferentes textos não foi a mesma, tampouco entre os anos escolares. Além disso, houve interação entre os fatores texto e o ano escolar,  $F(11, 417) = 5,855, p < 0,001, \eta^2 = 0,135$ , indicando que a precisão nos textos foi afetada pelo ano em curso. A Tabela 2 apresenta os resultados das análises de *post hoc* (Bonferroni) para os textos e para os anos escolares. Dentre as comparações significativas, estão fora da expectativa apenas as realizadas entre o Texto 1 e o Texto 2, entre o Texto 3 e os Textos 4 e 6 e entre o Texto 5 e o Texto 6, estando as demais comparações dentro do esperado (qual seja, os primeiros textos apresentando maior precisão do que os textos finais). Do ponto de vista da escolaridade, o 3º ano apresentou proporção de acertos significativamente menor que o 4º e o 5º anos.

**Tabela 2**  
**Resultado das Comparações entre os Textos e os Anos Escolares (Teste de Bonferroni)**

Comparações		Diferença de Média	p-valor	Resultado
Iniciantes	Texto 1	0,057	0,118	I = T1
	Texto 2	-0,031	1,000	I = T2
	Texto 3	0,202*	0,001	I > T3
	Texto 4	0,042	0,785	I = T4
	Texto 5	0,228*	0,001	I > T5
	Texto 6	0,102*	0,002	I > T6
Texto 1	Texto 2	-0,089*	0,001	T1 < T2
	Texto 3	0,145*	0,001	T1 > T3
	Texto 4	-0,015	1,000	T1 = T4
	Texto 5	0,171*	0,001	T1 > T5
	Texto 6	0,044	1,000	T1 = T6
	Texto 2	Texto 3	0,233*	0,001
Texto 4		0,074*	0,005	T2 > T4
Texto 5		0,259*	0,001	T2 > T5
Texto 6		0,133*	0,001	T2 > T6
Texto 3	Texto 4	-0,016*	0,001	T3 < T4
	Texto 5	0,026	1,000	T3 = T5
	Texto 6	-0,100*	0,001	T2 < T6
Texto 4	Texto 5	0,185*	0,001	T4 > T5
	Texto 6	0,059	0,181	T4 = T6
Texto 5	Texto 6	-0,126*	0,001	T5 < T6

Nota. I = iniciantes; T = Texto; \* = Comparações significativas.

A Tabela 3 apresenta as proporções médias de acertos de acordo com a classificação das perguntas utilizadas (LIT, GAP, TC, MS e EL), para cada texto e ano escolar, separadamente, e para a amostra geral. Os totais por ano escolar e para a amostra geral foram calculados a partir da proporção sobre a quantidade de itens presentes para cada tipo de questões, considerando-se o instrumento como um todo, com os sete textos. Estes totais foram utilizados para verificar a existência de efeito geral na proporção de acertos em cada texto, em função da classificação das respostas, por meio de ANOVAS para medidas repetidas.

Como fator intrassujeitos, utilizou-se a proporção de acertos de acordo com a classificação das

perguntas para cada texto e, como fator entre sujeitos, novamente o ano escolar.

**Tabela 3**  
**Proporção de Acertos de Acordo com a Classificação das Questões**

Ano	Texto	LIT	TC	GAP	EL	MS
3 <sup>o</sup>	Iniciantes	0,81	0,79	0,84	-	0,17
	Texto 1	0,92	0,76	0,48	0,04	0,32
	Texto 2	0,80	0,68	0,68	0,56	-
	Texto 3	0,40	0,68	0,19	-	0,24
	Texto 4	0,66	0,44	0,16	-	-
	Texto 5	0,70	0,32	0,36	-	-
	Texto 6	0,52	0,65	0,24	0,32	-
	Total	0,74	0,64	0,39	0,29	0,34
4 <sup>o</sup>	Iniciantes	0,88	0,75	0,91	-	0,25
	Texto 1	0,98	0,82	0,47	0,06	0,03
	Texto 2	0,91	0,82	0,84	0,94	-
	Texto 3	0,88	0,84	0,46	-	0,34
	Texto 4	0,94	0,78	0,03	-	-
	Texto 5	0,84	0,44	0,40	-	-
	Texto 6	0,75	0,75	0,06	0,47	-
	Total	0,90	0,75	0,48	0,49	0,36
5 <sup>o</sup>	Iniciantes	0,88	0,74	0,95	-	0,29
	Texto 1	0,94	0,89	0,75	0,01	0,58
	Texto 2	0,92	0,83	0,75	0,75	-
	Texto 3	0,96	0,81	0,41	-	0,71
	Texto 4	0,98	0,79	0,38	-	-
	Texto 5	0,92	0,63	0,48	-	-
	Texto 6	0,79	0,84	0,46	0,42	-
	Total	0,93	0,81	0,54	0,41	0,59
Total	Iniciantes	0,86	0,76	0,90	-	0,24
	Texto 1	0,95	0,82	0,56	0,04	0,28
	Texto 2	0,88	0,78	0,77	0,77	-
	Texto 3	0,75	0,78	0,36	-	0,42
	Texto 4	0,86	0,68	0,17	-	-
	Texto 5	0,82	0,46	0,41	-	-
	Texto 6	0,69	0,75	0,24	0,41	-
	Total	0,86	0,73	0,47	0,41	0,42

Nota. LIT = Literal; TC = *Text-connecting*; GAP = *Gap-filling*; EL = Elaborativa; MS = Modelo Situacional. Proporção de acertos (Pasquali, 2003): alta: ( $p \geq 0,65$ ); baixa: ( $p \leq 0,35$ ); mediana: (entre 0,36 e 0,64).



Os resultados apontaram um efeito geral dos diferentes tipos de perguntas,  $F(4, 210) = 182,384, p < 0,001, \eta^2 = 0,71$ , indicando que as médias de acertos não são as mesmas em função das diferentes classificações das perguntas. Também se observou efeito geral do ano escolar na precisão,  $F(2, 75) = 1835,422, p < 0,001, \eta^2 = 0,96$ , o que significa que as médias de acertos em função do tipo de questão foram afetadas pelo tempo de escolarização. Além disso, houve uma

interação do tipo de questão com o ano escolar, indicando que as proporções de acertos para os diferentes tipos de questão não foram as mesmas entre os textos, para cada ano escolar,  $F(8, 210) = 5,115, p < 0,001, \eta^2 = 0,12$ . A Figura 1 aponta que esta interação ocorreu em função da proporção de acertos nas questões de MS (ou seja, as crianças com maior idade apresentam um “salto” na proporção de acerto nesse tipo de questão, maior do que o que seria esperado pelo desempenho nos demais anos).

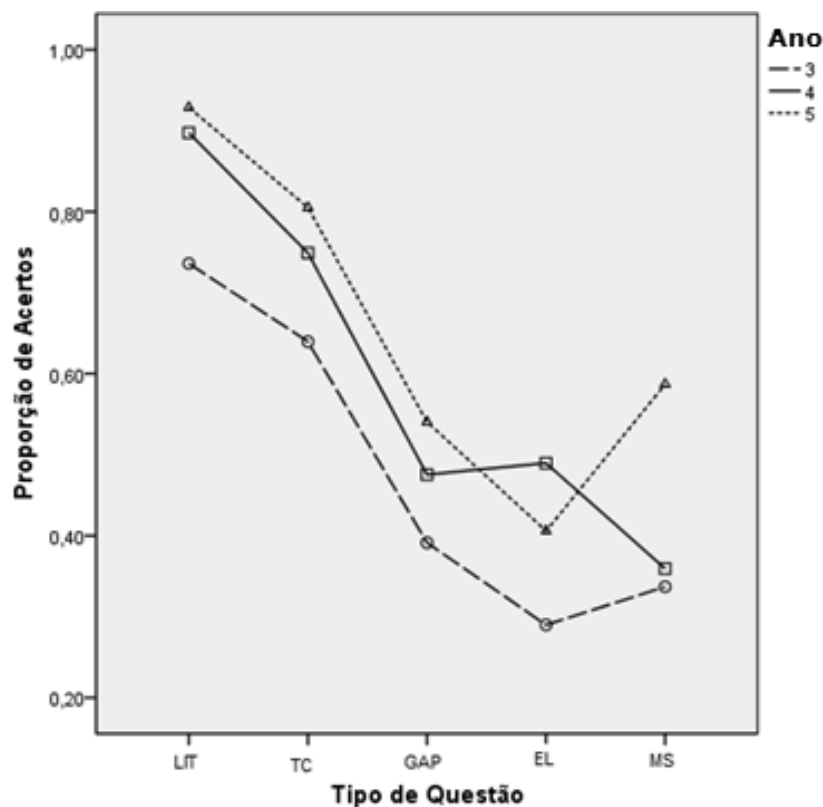


Figura 1. Proporção de acertos em função do tipo de questão para cada ano escolar.

Testes de *post hoc* são apresentados na Tabela 4. As análises mostram que, conforme esperado, as respostas LIT apresentaram maiores proporções de acerto do que todos os outros tipos de questões. Em seguida, ficaram as ques-

tões TC, com precisão significativamente maior do que as questões GAP, EL e MS. As questões GAP apresentaram maior precisão apenas que as EL, não se diferenciando das questões de MS. Estas últimas equivaleram às questões EL em termos de precisão.

**Tabela 4**  
**Teste de Post Hoc (Bonferroni) para Contraste da Precisão em Função do Tipo de Questões**

Tipo		Diferença entre médias	<i>p</i> -valor	Resultado
LIT	TC	0,123*	<0,001	LIT > TC
	GAP	0,385*	<0,001	LIT > GAP
	EL	0,459*	<0,001	LIT > EL
	MS	0,426*	<0,001	LIT > MS
TC	GAP	0,262*	<0,001	TC > GAP
	EL	0,336*	<0,001	TC > EL
	MS	0,303*	<0,001	TC > MS
GAP	EL	0,074*	0,039	GAP > EL
	MS	0,041	1,000	NS
EL	MS	-0,033	1,000	NS

Nota. LIT = Literal; TC = *Text-connecting*; GAP = *Gap-filling*; EL = Elaborativa; MS = Modelo Situacional. \* = comparações significativas; NS = Não significativo.

### Análise dos Itens

A Tabela 5 contém os índices de discriminação e de dificuldade dos itens, ambos os índices obtidos para modelos unifatoriais de dois parâmetros criados para os sete textos da PCL. Pelos índices de discriminação apresentados na Tabela 5, os quais foram resumidos na parte superior da Tabela 6, vê-se que, pelo critério de F. B. Baker (2001), a maioria dos itens (56,9%) apresentou discriminação insatisfatória, sendo 44% do total de itens com classificação muito baixa ( $a \leq 0,34$ ) e 12,5% com discriminação baixa ( $0,34 < a \leq 0,64$ ). Dos itens com discriminação adequada, a maioria (22,2%) apresentou índices moderados ( $0,64 < a \leq 1,34$ ), seguidos de 15,3% de itens com discriminação muito alta ( $a > 1,69$ ) e uma minoria (5,6%) com classificação alta ( $1,34 < a \leq 1,69$ ).

Na Tabela 6 (superior), observa-se que as questões LIT foram as que apresentaram a maior proporção de itens com classificação adequada, sendo 71,4% com discriminação moderada ou acima (10 de 14 itens). Em seguida, os melhores resultados ocorreram para as questões TC, com 45,1% de itens com discriminação adequada (14 de 31 itens). Os demais tipos de questão obtiveram índices semelhantes, com uma minoria de

itens adequados (28,6% para as questões MS (2 de 7 itens) e 25% para as GAP e EL (respectivamente, 4 de 16 itens e 1 de 4 itens).

Em relação à dificuldade (Tabelas 5 e parte inferior da Tabela 6), a maioria dos itens (31,9%) exigiu níveis medianos de habilidade (ou Theta,  $\theta$ ) para que uma probabilidade de acerto de 50% fosse atingida ( $-1 \leq b \leq 1$ ). Em segundo lugar, estiveram os itens com probabilidade de sucesso para indivíduos com habilidade baixa ( $-2 \leq b < -1$ ), seguidos dos itens para indivíduos com habilidade muito baixa ( $b < -2$ ), constituindo, respectivamente, 26,4% e 22,2% dos casos. Os demais itens funcionaram para sujeitos com níveis mais altos de habilidade, sendo 16,7% para nível muito superior ( $b > 2$ ) e uma pequena quantidade de itens (2,8%) para aqueles com habilidade superior ( $1 < b \leq 2$ ).

A maioria dos itens das questões EL, MS, LIT e GAP funcionaram em maior magnitude para os níveis medianos de habilidade (37,5% a 50% dos itens). Alguns itens funcionaram mais para os níveis altos de habilidade, com destaque para as questões EL (25%) e LIT (21,4%). As questões TC e GAP apresentaram itens mais informativos para os níveis baixos ou muito baixos de habilidade (respectivamente, 64,5% e 43,5%).

**Tabela 5**  
Valores de Discriminação (a) e de Dificuldade (b) dos Itens do PACL

Item	Tipo	a	b	Item	Tipo	a	b
I_1	LIT	0,990	-3,467	T3_6	GAP	0,903	-0,519
I_2	TC	-0,095	-1,256	T3_7	GAP	0,081	-4,301
I_3	LIT	3,000	-0,969	T3_8	GAP	0,070	4,607
I_4	TC	-0,068	4,052	T3_9	MS	0,436	-0,482
I_5	GAP	-1,943	1,934	T4_1	LIT	1,631	-0,580
I_6	TC	-0,525	3,554	T4_2	LIT	0,654	-2,507
I_7	TC	0,756	-2,436	T4_3	LIT	0,814	-0,879
I_8	TC	3,551	-2,125	T4_4	TC	0,268	0,308
I_9	LIT	-0,302	4,486	T4_5	TC	0,280	-3,377
I_10	GAP	-0,190	-1,509	T4_6	LIT	1,496	0,200
I_11	MS	0,058	2,455	T4_7	TC	0,174	-2,391
T1_1	LIT	-0,424	5,064	T4_8	GAP	0,733	0,424
T1_2	TC	0,007	-3,632	T4_9	TC	0,396	0,292
T1_3	TC	0,295	-1,783	T4_10	GAP	0,202	-3,560
T1_4	TC	1,471	-1,649	T4_11	MS	0,029	-7,349
T1_5	EL	0,068	3,061	T5_1	GAP	0,305	-1,050
T1_6	TC	3,768	-1,659	T5_2	TC	0,914	-2,578
T1_7	TC	-0,397	4,563	T5_3	LIT	0,278	-3,013
T1_8	LIT	0,750	-1,950	T5_4	GAP	0,484	0,060
T1_9	GAP	-0,515	4,798	T5_5	GAP	0,618	-0,488
T1_10	MS	3,150	0,999	T5_6	GAP	0,416	-1,912
T2_1	TC	1,360	-1,007	T5_7	LIT	0,941	-1,401
T2_2	LIT	-0,079	4,451	T5_8	TC	2,337	0,846
T2_3	TC	0,118	-3,444	T5_9	TC	5,230	0,713
T2_4	TC	0,177	-2,414	T5_10	GAP	5,405	0,661
T2_5	EL	-0,073	-6,584	T5_11	MS	1,286	1,048
T2_6	TC	2,317	-1,408	T6_1	EL	0,740	-0,346
T2_7	GAP	1,178	-1,286	T6_2	TC	0,297	-0,683
T2_8	TC	0,804	-1,561	T6_3	TC	2,214	-1,675
T2_9	MS	0,441	-1,366	T6_4	TC	2,158	-1,684
T2_10	EL	0,440	-0,700	T6_5	TC	0,503	-2,086
T3_1	TC	-0,221	7,693	T6_6	TC	0,999	-0,482
T3_2	TC	0,274	-1,239	T6_7	LIT	0,698	-0,413
T3_3	TC	7,914	-1,598	T6_8	TC	0,074	3,319
T3_4	GAP	0,278	-1,013	T6_9	GAP	0,472	0,903
T3_5	LIT	0,878	-0,904	T6_10	MS	-0,264	-0,175

Nota. T = texto; LIT = Literal; TC = Text-connecting; GAP = Gap-filling; EL = Elaborativa; MS = Modelo Situacional.

**Tabela 6**  
**Proporção da Classificação dos Índices de Dificuldade e de Discriminação em Função do Tipo de Questão**

Tipo	Discriminação				
	$a \leq 0,34$	$0,34 \leq a < 0,64$	$0,64 \leq a < 1,34$	$1,34 \leq a < 1,69$	$a \geq 1,69$
LIT	28,6%	0,0%	50,0%	14,3%	7,1%
TC	48,4%	6,5%	12,9%	6,5%	25,8%
GAP	50,0%	25,0%	18,8%	0,0%	6,3%
EL	50,0%	25,0%	25,0%	0,0%	0,0%
MS	42,9%	28,6%	14,3%	0,0%	14,3%
Total	44,4%	12,5%	22,2%	5,6%	15,3%

	Dificuldade				
	$b \leq -2$	$-2 \leq b < -1$	$-1 \leq b < 1$	$1 \leq b < 2$	$b \geq 2$
LIT	21,4%	14,3%	42,9%	0,0%	21,4%
TC	29,0%	35,5%	19,4%	0,0%	16,1%
GAP	12,5%	31,3%	37,5%	6,3%	12,5%
EL	25,0%	0,0%	50,0%	0,0%	25,0%
MS	14,3%	14,3%	42,9%	14,3%	14,3%
Total	22,2%	26,4%	31,9%	2,8%	16,7%

Nota. LIT = Literal; TC = *Text-connecting*; GAP = *Gap-filling*; EL = Elaborativa; MS = Modelo Situacional;  $a$  = discriminação;  $b$  = dificuldade.

## Discussão

A Bateria de Avaliação da Compreensão Leitora, que está em fase de desenvolvimento, tem como objetivo avaliar de forma fidedigna e válida as habilidades básicas de escolares dos primeiros anos do Ensino Fundamental em tarefas de compreensão textual e outras capacidades cognitivas/linguísticas correlatas, as quais constituem: decodificação leitora, nomeação rápida e memória de trabalho verbal. Relatou-se nesta pesquisa o primeiro estudo piloto da Prova de Compreensão da referida bateria. A prova é composta por um texto de adaptação (Texto Iniciantes) e seis textos de prova (quatro narrativos e dois expositivos, dispostos sequencialmente segundo a extensão e complexidade sintática e vocabular, a priori pretendidas, mas não analisadas no presente estudo).

Pela análise da precisão nas tarefas, foi possível mostrar que, de maneira geral, há uma

progressão nas médias de acerto dos textos, sendo que os escolares do 3º ano conquistaram precisão acentuada apenas nos primeiros textos, enquanto alunos do 5º ano mantiveram médias elevadas em todos os textos, com exceção dos Textos 3 e 5, em que os escores foram medianos. A progressão observada em termos do aumento dos escores dos participantes está de acordo com as expectativas e constitui uma medida do desenvolvimento da compreensão (Cadime et al., 2013; Cain & Oakhill, 2012; Cunha, Silva, & Capellini, 2012). Análises de variância para medidas repetidas confirmaram esta observação, ao demonstrarem a presença de diferenças significativas entre a proporção de acertos nos textos entre os anos escolares, além de uma interação entre esses fatores. O efeito geral do ano indica que as crianças mais jovens obtiveram escores consistentemente mais baixos do que seus pares de idade mais avançada, enquanto que a interação do texto com o ano escolar mostra que a pre-

cisão em certos textos foi afetada de modo diferenciado de acordo com a faixa de escolaridade das crianças. Estes resultados apontam que os textos iniciais são mais indicados para avaliação da compreensão das crianças menores, enquanto que os textos finais são mais apropriados para as mais velhas, o que está de acordo com o esperado (Cadime et al., 2013; Cunha & Capellini, 2013; Cunha et al., 2012; Joly & Istome, 2008).

A progressão da proporção de acertos não foi mantida por apenas alguns textos, tendo o Texto 1 menor precisão que o Texto 2; os Textos 4 e 6 menos precisos que o Texto 3; e o Texto 5 com menor precisão do que o Texto 6. É possível que um desequilíbrio nos tipos de questões em cada texto esteja relacionado a estes resultados. Conforme foi demonstrado, houve um efeito geral do tipo de questão na precisão dos escores, sendo que as questões de compreensão LIT e TC obtiveram, em média, maior precisão para a amostra, enquanto que as questões GAP, MS e EL produziram mais erros, o que está de acordo com Cain e Oakhill (1999). Por exemplo, o Texto 4 é composto, em sua maior parte, por questões LIT e TC (88,8% no total), sendo apenas uma pequena proporção de questões GAP (10%). Já o Texto 6 apresenta apenas 11% de questões do tipo GAP. Por outro lado, o Texto 3 é composto principalmente por questões GAP (44%) e ainda apresenta uma questão de MS. Esta configuração de questões está provavelmente na raiz de uma maior precisão da amostra nos Textos 4 e 6 em relação ao Texto 3. Os resultados obtidos para o tipo de questões apontam a pertinência da classificação das inferências adotada no estudo e, portanto, de se avaliar os diferentes níveis de integração de informações durante a leitura compreensiva de texto (Cain & Oakhill, 2012; Díaz & Vega, 2003; Martínez & Gámez, 1998). Por sua vez, reiteram a necessidade de se utilizar instrumentos de avaliação que investiguem os diferentes processos cognitivos implicados nos diversos níveis de processamento do texto – retenção e integração local e global das ideias veiculadas – de modo a caracterizar os possíveis déficits capazes de influenciar o desempenho em compreensão de leitura (Cain & Oakhill, 1999; Martínez & Gámez, 1998).

A análise dos itens mostrou que uma grande proporção de itens não atingiu níveis adequados de discriminação. Quase 57% dos itens apresentou discriminação inadequada (baixa, muito baixa ou negativa). As questões LIT e TC, mesmo com as maiores proporções de acertos, foram as que conseguiram preservar o maior número de itens com níveis adequados de discriminação. Enquanto isso, as questões elaborativas apresentaram os piores índices, o que sugere que ajustes devem ser conduzidos principalmente entre essas questões, de modo a produzir resultados mais satisfatórios para o conjunto final de itens no que diz respeito à discriminação.

Em relação à dificuldade, observou-se que os itens parecem funcionar mais para os níveis medianos de habilidade ou Theta, apesar da grande proporção de itens que funcionaram para níveis baixos ou muito baixos. Desse modo, o desafio futuro é conseguir, simultaneamente, contemplar de uma maneira mais ampla os níveis mais elevados de Theta (os quais representaram apenas 2,8% dos itens) e reduzir a quantidade de itens que contemplam os níveis muito baixos (que compõem 22,2% do total). Um fato interessante, e aparentemente contraditório, é que as questões LIT, juntamente com as questões EL, funcionaram em grande proporção para os níveis mais altos de habilidade (o que não seria esperado, em virtude de terem sido respondidas com um nível de precisão significativamente maior do que as demais, conforme demonstrado pelas ANOVAS). Assim, pôde-se observar que todos os três itens LIT que funcionaram para um valor de  $b > 1,5$  obtiveram discriminação negativa (sendo, portanto, itens que devem ser revisados ou removidos). De um modo geral, todos os itens que funcionaram para os valores altos de Theta apresentaram índices de discriminação inadequados, o que reforça o desafio de se criar itens para esta faixa de habilidade.

### **Considerações Finais**

Os resultados obtidos neste estudo piloto fundamentarão a análise exploratória dos textos e/ou das questões utilizadas na Prova de Compreensão Leitora, o que será útil para a condução

de um estudo mais amplo, destinado a elaborar a versão final do instrumento. Em geral, os resultados apontam que, apesar da necessidade de ajustes, as tarefas criadas parecem ser adequadas para a avaliação da compreensão leitora de crianças na fase inicial de escolarização, uma vez que os modelos de TRI mostraram-se ajustados e funcionaram para a amostra. A progressão da dificuldade dos textos, a diferenciação pela escolaridade, assim como as diferenças nos escores em função do tipo de questão utilizada fornecem fontes valiosas de validação do construto que se pretende investigar (Urbina, 2014). Isto aumenta a confiança de que o PACL será uma ferramenta útil para a avaliação da compreensão leitora da população de interesse.

Dentre os principais desafios para as pesquisas futuras com o instrumento apontadas neste piloto, destacam-se: (a) a necessidade de modificações das questões para ampliação da oferta de itens que atendam níveis mais elevados de Theta, na busca de aumentar, progressivamente, a demanda cognitiva imposta na avaliação da compreensão leitora ao longo do instrumento; (b) ajustes nas questões de modo a melhorar os índices de discriminação dos itens; (c) verificar a validade de construto dos tipos de questões a fim de constatar se elas realmente avaliam os processos cognitivos que se pretende; (d) e, por fim, a necessidade de revisão de alguns textos que romperam a expectativa de aumento progressivo da dificuldade do instrumento, a saber, os Textos 1, 4 e 6. A revisão destes aspectos propiciará a obtenção de um instrumento psicometricamente ajustado, com a adequada validação de construto e que apresente, na delimitação das inferências pretendidas pelos seus escores, um maior apuro para o diagnóstico dos transtornos da compreensão leitora, assim como traçar o percurso esperado de seu desenvolvimento entre as crianças típicas.

Cabe ressaltar que o trabalho seguiu todas as etapas propostas pela psicometria para a construção de instrumentos de avaliação (Urbina, 20104), as quais são muitas vezes desconsideradas pelos pesquisadores. Nesse sentido, demonstra-se a importância desse estudo para aqueles que estão interessados na construção de tarefas de avaliação da linguagem.

## Referências

- Baker, F. B. (2001). *The basics of item response theory*. Washington, DC: ERIC Clearinghouse on Assessment and Evaluation.
- Baker, L., & Stein, N. (1981). The development of prose comprehension skills. In C. M. Santa & B. L. Hayes (Eds.), *Children's prose comprehension: Research and practice* (pp. 7-43). Newark, NJ: International Reading.
- Bernhardt, E. B. (1991). *Reading development in a second-language*. Norwood, NJ: Ablex.
- Brandão, A. C. P., & Spinillo, A. G. (1998). Aspectos gerais e específicos na compreensão de textos. *Psicologia: Reflexão e Crítica*, 11(2), 253-272. doi:10.1590/S0102-79721998000200006
- Cadime, I., Ribeiro, I., Viana, F. L., Santos, S., Prieto, G., & Maia, J. (2013). Validity of a reading comprehension test for Portuguese students. *Psicothema*, 25(3), 384-389. doi:10.7334/psicothema2012.288
- Cain, K., & Oakhill, J. V. (1999). Inference making ability and its relation to comprehension failure in young children. *Reading and Writing: An Interdisciplinary Journal*, 11, 489-503. doi:10.1023/A:1008084120205
- Cain, K., & Oakhill, J. V. (2006). Profiles of children with specific reading comprehension difficulties. *British Journal of Educational Psychology*, 76, 683-696. doi:10.1348/000709905X67610
- Cain, K., & Oakhill, J. V. (2008). *Children's comprehension problems in oral and written language: A cognitive perspective*. New York: The Guilford Press.
- Cain, K., & Oakhill, J. V. (2012). Reading comprehension development from seven to fourteen years: Implication for assessment. In J. Sabatini, E. Albro, & T. O'Reilly (Eds.), *Measuring up: Advances in how we assess reading ability* (pp. 59-76). Lanham, MD: Rowman & Littlefield Education.
- Cain, K., Oakhill, J., & Bryant, P. (2004). Children's reading comprehension ability: Concurrent prediction by working memory, verbal ability, and component skills. *Journal of Educational Psychology*, 96(1), 31-42. doi:10.1037/0022-0663.96.1.31
- Carvalho, C. A. F., Avila, C. R. B. de, & Chiari, B. M. (2009). Níveis de compreensão de leitura

- em escolares. *Pró-Fono Revista de Atualização Científica*, 21(3), 207-212. doi:10.1590/S0104-56872009000300005
- Catts, H. W., & Weismer, S. E. (2006). Language deficits in poor comprehenders: A case for the simple view of reading. *Journal of Speech, Language, and Hearing Research*, 49, 278-293. doi:10.1044/1092-4388(2006/023)
- Cuetos, F., Rodrigues, B., & Ruano, E. (2012). *Provas de avaliação dos processos de leitura – PROLEC [Manual]* (S. A. Capellini, A. M. de Oliveira, & F. Cuetos, Adaptação brasileira, 2. ed.). São Paulo, SP: Casa do Psicólogo.
- Cunha, V. L. O., & Capellini, S. A. (2013). Construção e validação de instrumento de avaliação da compreensão de leitura para escolares do terceiro ao quinto ano do ensino fundamental. *CoDAS*, 26(1), 28-37. doi:10.1590/S2317-17822014000100005
- Cunha, V. L. O., Silva, C. da, & Capellini, S. A. (2012). Correlação entre habilidades básicas de leitura e compreensão de leitura. *Estudos de Psicologia (Campinas)*, 29(Supl. 1), 799-807. doi:10.1590/S0103-166X2004000200003
- Cunningham, J. W., & Cunningham, P. M. (1981). *Middle and secondary school reading*. New York: Longman.
- Díaz, J. M., & Vega, M. (2003). Modelos de la situación e inferencias en la comprensión de textos. In J. A. León (Ed.), *Conocimiento y discurso. Claves para inferior y comprender* (pp. 132-152). Madrid, España: Pirâmide.
- Embretson, S. E., & Reise, S. P. (2000). *Item response theory for psychologists*. Mahwah, NJ: Lawrence Erlbaum.
- Galvão, A., & Mahon, R. (1998). Compreensão de texto em crianças: Comparações entre diferentes classes de inferência a partir de uma metodologia on-line. *Psicologia: Reflexão e Crítica*, 20(3), 463-471. doi:10.1590/S0102-79722007000300014
- Joly, M. C. R. A., & Istome, A. C. (2008). Compreensão em leitura e capacidade cognitiva: Estudo de validade do teste Cloze Básico - MAR. *Psicologia: Revista da Vetor Editora*, 9(2), 219-228. Recuperado em <http://pepsic.bvsalud.org/pdf/psic/v9n2/v9n2a10.pdf>
- Keenan, J. M., & Betjemann, R. S. (2006). Comprehending the gray oral reading test without reading it: Why comprehension tests should not include passage-independent items. *Scientific Studies of Reading*, 10(4), 363-380. doi:10.1207/s1532799xssr1004\_2
- Kida, A. de S. B., Carvalho, C. A. F., & Avila, C. R. B. de. (2012). Avaliação da compreensão nos diversos transtornos de leitura. In S. A. Capellini, M. N. Sampaio, & A. M. De Oliveira (Eds.), *Tópicos em transtornos de aprendizagem* (pp. 95-107). São José dos Campos, SP: Pulso Editorial.
- Kintsch, W. (1988). The role of knowledge in discourse comprehension: A construction-integration model. *Psychological Review*, 95(2), 163-182. doi:10.1037/0033-295X.95.2.163
- Leslie, L., & Caldwell, J. (2009). Formal and informal measures of reading comprehension. In S. E. Israel & G. G. Duffy (Eds.), *Handbook of research on reading comprehension* (pp. 403-427). New York: Taylor y Francis.
- Gomes, M. A. M., & Boruchovitch, E. (2005). Desempenho no jogo, estratégias de aprendizagem e compreensão na leitura. *Psicologia: Teoria e Pesquisa*, 21(3), 319-326. doi:10.1590/S0102-37722005000300008
- Martínez, G., & Gámez, E. V.-A. (1998). ¿Por qué los textos son difíciles de comprender? Las interferencias son la respuesta. *Textos de Didáctica de la Lengua y la Literatura*, 16, 85-97.
- Oakhill, J. V., & Cain, K. (2012). The precursors of reading ability in young readers: Evidence from a four-year longitudinal study. *Scientific Studies of Reading*, 16(2), 91-121. doi:10.1080/10888438.2010.529219
- Oakhill, J. V., Cain, K., & Bryant, P. E. (2003). The dissociation of word reading and text comprehension: Evidence from component skills. *Language and Cognitive Processes*, 18(4), 443-468. doi:10.1080/01690960344000008
- Pasquali, L. (2003). *Psicometria: Teoria dos testes na psicologia e na educação*. Rio de Janeiro, RJ: Vozes.
- Paul, R., & Elder, L. (2012). *Critical thinking: Tools for taking charge of your learning & your life* (3<sup>rd</sup> ed.). Mahwah, NJ: Prentice Hall.
- Reed, D. K., & Vaughn, S. (2012). Retell as an indicator of reading comprehension. *Scientific Studies of Reading*, 16(3), 187-217. doi:10.1080/10888438.2010.538780

- Salles, J. F. de, & Parente, M. A. de M. (2002). Processos cognitivos na leitura de palavras em crianças: Relações com compreensão e tempo de leitura. *Psicologia: Reflexão e Crítica*, 15(2), 321-331. doi:10.1590/S0102-79722002000200010
- Sánchez, E. (2002). *Compreensão e redação de textos: Dificuldades e ajudas*. Porto Alegre, RS: Artmed.
- Snow, C. E., & Sweet, A. P. (2003). Reading for comprehension. In A. P. Sweet & C. E. Snow (Eds.), *Rethinking reading comprehension* (pp. 1-11). New York: Guilford Press.
- Snowling, M. J., & Stackhouse, J. (2004). *Dislexia, fala e linguagem: Um manual profissional*. Porto Alegre, RS: Artmed.
- Urbina, S. (2014). *Essentials of psychological testing* (2<sup>a</sup> ed.). Hoboken, New Jersey: John Wiley & Sons.
- Van den Broek, P., Rapp, D., & Kendeou, P. (2005). Integrating memory-based and constructionist processes in accounts of reading comprehension. *Discourse Processes*, 39(2), 299-316. doi:10.1207/s15326950dp3902&3\_11
- Yuill, N. M., & Oakhill, J. V. (1996). Higher order factors in comprehension disability: Processes and remediation. In C. Cornoldi & J. V. Oakhill (Eds.), *Reading difficulties: Processes and intervention* (pp. 69-92). Mahwah, NJ: Lawrence Erlbaum.

Recebido: 01/09/2014  
1<sup>a</sup> revisão: 28/10/2014  
Aceite final: 11/12/2014