

# Escala da Avaliação da Competência da Leitura pelo Professor (EACOL) – Evidências de Validade de Critério

Patrícia Silva Lúcio<sup>1</sup>

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

Ângela Maria Vieira Pinheiro

*Departamento de Psicologia da Universidade Federal de Minas Gerais,  
Belo Horizonte, Brasil*

## Resumo

A Escala da Avaliação da Competência da Leitura pelo Professor (EACOL) foi criada com o objetivo de fornecer uma maneira rápida e de baixo custo para a triagem das dificuldades de aprendizagem dos alunos pelos professores. O presente estudo investiga a concordância entre o desempenho de crianças na leitura em voz alta de palavras isoladas e a avaliação dos professores na EACOL. Foram avaliadas 137 crianças (52% do sexo feminino) cursando do 2º ao 5º ano de escolarização da cidade de Belo Horizonte. Em geral, as correlações entre o desempenho das crianças na leitura de palavras e a EACOL foram moderadas, sendo que a maior concordância de desempenho esteve entre as classificações de habilidade mais baixas. Os resultados são promissores, mas apontam para a necessidade de ajustes nos itens da EACOL, a fim de melhorar a concordância nos níveis mais elevados de habilidade.

**Palavras-chave:** Leitura, heteroavaliação, validade de critério.

## Scale of Evaluation of Reading Competence by the Teacher (EACOL) – Evidence of Criterion Validity

### Abstract

The Scale of Assessment of Reading Competence of Students by the Professor (EACOL) was created with the goal of providing a quick and inexpensive screening for learning difficulties of students by teachers. The present study investigates the correlation between the performance of children in reading single words aloud and their evaluation by the teachers in the EACOL. We evaluated 137 children (52% female) attending the 2nd to the 5th year of schooling in the city of Belo Horizonte. In general, the correlations between children's performance in reading words and EACOL were moderate, and the agreement in performance was better between lower ability classifications. The results are promising, but point to the need for adjustments to the items of EACOL in order to improve the agreement in the highest levels of skill.

**Keywords:** Reading, hetero-assessment, criterion validity.

---

<sup>1</sup> Endereço para correspondência: Departamento de Psicologia e Psicanálise, Universidade Estadual de Londrina, Rodovia Celso Garcia Cid, Pr. 445, Km 380, Campus Universitário, Londrina, PR, Brasil, 86057-970. Cx. Postal: 10.011. E-mail: pslucio@gmail.com

Agradecimentos: Este trabalho teve apoio da Fundação de Amparo à Pesquisa do Estado de Minas Gerais (Fapemig) e do Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq).

## Escala de Evaluación de las Competencias en Lectura por el Profesor (EACOL) – La Evidencia de la Validez de Criterio

### Resumen

La Escala de Evaluación de Competencia de Lectura de Estudiantes por el Profesor (EACOL) fue creada con el objetivo de proporcionar una manera rápida y económica de evaluar las dificultades de aprendizaje de los alumnos por los profesores. El presente estudio investiga la correlación entre el desempeño de niños al leer en voz alta palabras aisladas y la evaluación de los profesores en la EACOL. Fueron evaluados 137 niños (52% sexo femenino) cursando entre el 2° y 5° año de sistema escolarizado en la ciudad de Belo Horizonte. En general, la correlación entre el desempeño de lectura de los niños y la EACOL fue moderado, y la mayor concordancia en el desempeño estuvo entre las clasificaciones de habilidad más bajas. Los resultados son prometedores, pero apuntan a la necesidad de ajustes en los elementos de la EACOL para mejorar la concordancia en los niveles de habilidad más altos.

**Palabras clave:** Lectura, heteroevaluación, validez de criterio.

A dificuldade de aprendizagem da leitura atinge um grande número de crianças em todo o mundo. Dados internacionais apontam que cerca de 80% das crianças que apresentam dificuldades de aprendizagem escolar concentram seus problemas na área da leitura, sendo que a proporção de crianças disléxicas gira em torno de 15-20% (Kenyon, 2003). No Brasil, um estudo recente mostrou que a prevalência da dislexia entre crianças de Pelotas (RS) foi de 12,5% (Gutierrez & Tomasi, 2011). Além disso, o resultado geral da última avaliação dos alunos do 5° ano em português no Sistema de Avaliação da Educação Básica (Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira [INEP], 2011) mostrou que o desempenho de nossas crianças ainda se distancia da meta de 200 pontos estipulada pelo governo, ficando em 190,6 pontos. Paralelamente, a falta de profissionais especializados nas escolas impede que crianças portadoras de dificuldades de leitura sejam adequadamente diagnosticadas, retardando uma intervenção precoce de seus problemas. Essa questão nos faz refletir sobre a importância do professor em proporcionar uma primeira avaliação rápida e eficaz de seus alunos, de modo a identificar e encaminhar aqueles que possivelmente estão apresentando um atraso na aquisição da habilidade de leitura.

Estudos têm apontado que o professor é um bom avaliador do real desempenho acadêmico de seus alunos (para uma revisão, ver Hoge &

Caladarci, 1989). As correlações médias entre o desempenho do aluno e a avaliação do professor giram em torno de 0,66, mas há uma grande amplitude nesses valores, podendo variar de 0,28 a 0,92 (Hecht & Greenfield, 2002). Bennett, Gottesman, Rock e Gerullo (1993) argumentam que fatores externos ao desempenho acadêmico das crianças (como o gênero ou o comportamento em sala de aula) afetam a avaliação do professor, reduzindo a validade preditiva das escalas.

No Brasil, pesquisas têm avaliado a eficácia do julgamento do professor sobre as habilidades acadêmicas de seus alunos. É o caso, por exemplo, do estudo de Pinheiro (2001), que avaliou o desempenho de alunos considerados bons leitores pelos professores em tarefa de leitura. A autora descobriu que 35% dos estudantes classificados como bons leitores pelos professores apresentaram desempenho real aquém do esperado. Destes, três casos foram considerados críticos. Esses resultados sugerem que os professores podem ser capazes de julgar de maneira acurada o desempenho de seus alunos de uma forma mais geral, mas alguns aspectos específicos do modelo de processamento de informação devem ser avaliados por meio de tarefas construídas para esta finalidade.

Em outro estudo, cujo objetivo foi comparar o desempenho acadêmico de crianças de acordo com avaliações formais e a opinião dos professores, Capellini, Tonelotto e Ciasca (2004) utilizaram uma amostra de 164 alunos do 3° ao 5°

ano do Ensino Fundamental. Para a avaliação formal das habilidades de leitura, escrita e aritmética, foi utilizado o Teste de Desempenho Escolar (TDE). Os professores dividiram a amostra em dois grupos, de acordo com presença ou ausência de dificuldades de aprendizagem. Como resultado, o grupo de crianças sem dificuldades apresentou melhor desempenho em relação às crianças com dificuldades. A concordância entre a avaliação formal e a dos professores em todas as habilidades investigadas foi maior entre as crianças apontadas como sendo portadoras de dificuldades (76% na leitura a 95% no desempenho geral). Menores índices de concordância foram obtidos nas crianças classificadas como normais (51% em aritmética a 85% na leitura) e isto está de acordo com a conclusão de Pinheiro (2001) de que muitas das dificuldades específicas de aprendizagem podem ser camufladas durante o desenvolvimento dessas habilidades.

Salles e Parente (2007) relacionaram a opinião dos professores, medida pelo Protocolo de Caracterização do Desempenho Escolar do Aluno (PCDE), com o desempenho de crianças de 2ª série em tarefas de leitura, escrita e compreensão de texto. Como resultado, todas as questões presentes no PCDE (que avaliaram leitura, escrita e aritmética) se correlacionaram de forma moderada com o desempenho das crianças, exceto a velocidade, fluência e precisão na leitura, que apresentaram correlações não significativas ou baixas. Uma análise gráfica mostrou a concordância entre o desempenho das crianças em leitura e escrita e a avaliação dos professores. A maior concordância foi para as crianças classificadas com desempenho de muito bom a ótimo, enquanto uma maior variabilidade ocorreu para as crianças enquadradas com desempenho bom ou insuficiente. As autoras concluíram que a avaliação do professor foi útil na identificação de dificuldades acadêmicas dos alunos, mas não excluíram a interferência de fatores extrínsecos às habilidades acadêmicas das crianças nesta avaliação (como o comportamento em sala de aula, por exemplo).

Neste mesmo estudo, Salles e Parente (2007) relacionaram o desempenho das crianças com o resultado na Escala da Avaliação da Competência da Leitura pelo Professor (EACOL),

que classifica a leitura da criança pela perspectiva do professor (Pinheiro & Costa, 2005). De forma semelhante ao PCDE, a correlação da EACOL com o desempenho em leitura, escrita e compreensão de texto foi moderada. Além disso, a análise gráfica também mostrou uma variabilidade entre os alunos considerados leitores maus ou razoáveis pela escala, mas esta variabilidade também se estendeu para os bons leitores. Por fim, uma análise de variância (*One-Way*) mostrou a existência de diferenças de desempenho entre os leitores considerados bons, razoáveis e maus em todas as tarefas aplicadas, sendo que os três grupos diferiram entre si em relação à leitura de palavras, enquanto que em compreensão e na escrita os dois grupos menos proficientes não se diferiram.

Uma consideração deve ser feita sobre o método de análise dos dados no estudo de Salles e Parente (2007). Para resultados mais conclusivos, correlações também deveriam ter sido feitas entre o desempenho dos alunos e os subcomponentes da EACOL (leitura em voz alta e leitura silenciosa). É possível que as correlações não tenham sido mais altas porque um dos componentes da escala avalia melhor o desempenho dos alunos que o outro. Ademais, as Figuras contendo os gráficos de concordância consideraram o desempenho das crianças tanto em relação à leitura, quanto à escrita. Diferentemente do PCDE, a EACOL não apresenta questões relacionadas à escrita, o que pode ter reduzido a concordância.

O estudo de Salles e Parente (2007) sugere que o professor é um profissional capaz de avaliar o desempenho dos seus alunos com eficiência, apesar de possivelmente não estar preparado para avaliar aspectos mais específicos do desenvolvimento cognitivo, como foi demonstrado pela ausência de correlações entre as medidas de fluência, velocidade e precisão na leitura com as demais variáveis. Além disso, a maior variabilidade de desempenho entre os sujeitos classificados nos níveis mais baixos de desempenho, tanto no PCDE quanto na EACOL, indica que estes instrumentos precisam sofrer modificações de maneira a distinguir de forma mais eficaz aqueles que têm desempenho médio daqueles que possuem desempenho inferior.

Recentemente, Cogo-Moreira, Ploubidis, Ávila, Mari e Pinheiro (2012) apresentaram evidências de validade concorrente e discriminante para a EACOL em amostras de crianças de 2º ao 5º ano de Belo Horizonte ( $N = 335$ ) e de 2º ao 4º ano de São Paulo ( $N = 617$ ). Os autores mostraram que a acurácia (ou número de acertos por minuto) na leitura de palavras reais, de pseudopalavras e na leitura de textos foram preditoras do escore total na EACOL, demonstrando a validade concorrente da escala. Adicionalmente, modelos de regressão não foram significativos para a avaliação comportamental das crianças (feita pelo SDQ – *Strengths and Difficulties Questionnaire*) e para a inteligência (WISC-III), atestando para a sua validade discriminante.

O presente trabalho busca fornecer novas fontes de evidência de validade para a EACOL, as quais estão relacionadas ao critério. Segundo Pasquali (2009), a validade de critério constitui um conjunto de procedimentos usados para demonstrar a eficácia de um teste na predição do comportamento dos sujeitos. Este comportamento é o critério que será correlacionado com o desempenho no teste e deve ser uma medida independente deste. Na presente pesquisa, a validade de critério será investigada por meio de correlações de Pearson entre a classificação das crianças na EACOL e seu desempenho em uma tarefa de leitura em voz alta de palavras isoladas, o que constitui um tipo de validação concorrente, uma vez que as medidas de teste e de critério foram obtidas simultaneamente (Urbina, 2007). A concordância entre a avaliação dos professores e o desempenho das crianças também será investigada qualitativamente, por meio do uso de tabelas de contingência e análises gráficas, ou estatisticamente, por meio de análises de variância (ANOVA). Os resultados serão reportados separadamente para cada situação da escala (leitura em voz alta, leitura silenciosa e escore total). O esperado é que as crianças classificadas pelas professoras como boas leitoras apresentem escores mais altos na lista de palavras, enquanto que as classificadas como razoáveis ou más apresentem piores desempenhos. Consideraremos significativos os resultados com  $p < 0,05$ .

## Método

### Considerações Éticas

O presente trabalho foi aprovado pelo Conselho de Ética da Universidade Federal de Minas Gerais (processo nº ETIC 347/04).

### Amostra

Participaram do estudo 137 estudantes (52% do sexo feminino), cursando do 2º ao 5º ano do ensino fundamental e seus respectivos professores ( $N = 23$ ), de cinco escolas da cidade de Belo Horizonte (três estaduais, uma municipal e uma particular). A idade mínima da amostra de crianças foi de seis anos e a idade máxima foi de 11 anos (média 8,7,  $DP = 1,2$ ).

### Instrumento

A Escala da avaliação da Competência da Leitura pelo Professor (EACOL) é um instrumento criado com o objetivo de fornecer uma forma rápida e eficiente para avaliar a habilidade de leitura de crianças, no que diz respeito às situações de leitura em voz alta (LVA) e leitura silenciosa (LS; Pinheiro & Costa, 2005). A escala apresenta duas versões, sendo uma para crianças do 2º ano de escolarização e a outra para crianças do 3º ao 5º anos. Na situação de LVA, a escala apresenta 13 itens na versão para crianças do 2º ano e 17 itens para crianças mais velhas. Em ambas as versões, há 10 itens para a avaliação da LS. Os escores de ambas as situações são somados para compor o escore total em leitura. Para cada item, o professor deve responder “sim” se a frase proposta descreve a leitura de um determinado aluno e “não”, caso contrário. Por exemplo, no primeiro item da situação de LVA do 2º ano, o professor deve marcar se concorda ou não com a afirmativa “demonstra ter entendido o que leu quando indagado sobre o texto lido”, a respeito de um aluno específico.

Para ambas as versões, os itens da escala foram construídos de modo a descrever o desempenho dos alunos em três níveis distintos: se lê bem, se lê razoavelmente ou se lê mal. No caso do leitor razoável, há uma subdivisão entre aqueles que leem razoavelmente bem ou razoavelmente

mal. A pontuação da escala adotada neste estudo segue a versão modificada por Pinheiro e Costa (2011). Assim, para os itens cujos conteúdos se referem ao leitor “bom” e “mal”, a criança pode obter +2 ou -2 pontos, recebendo uma pontuação de +2 quando o professor responde “sim” a um item que descreve um bom leitor ou se responde “não” a um item que descreve um mau leitor. Caso contrário, a criança recebe um escore de -2. No exemplo dado no parágrafo anterior, a pergunta caracteriza o desempenho de um bom leitor. Desse modo, se a professora responder “sim” ao item, o escore da criança será de +2 e se responder “não” o escore será de -2.

Já para os itens que descrevem uma leitura razoável, a criança pode receber um escore de +1 ou -1. O escore de -1 ocorre nos casos em que a professora responde “sim” a um item que descreve um leitor razoável, e o escore de +1, se a resposta for “não”. Por exemplo, se a professora responder “sim” ao item 6 da situação de LVA da escala do 3º ao 5º ano (“Às vezes, lê e não sabe contar o que leu”), o escore será de -1, enquanto que se responder “não” o escore será de +1<sup>2</sup>.

Na situação de LVA, a pontuação varia de -21 a +22 pontos no 2º ano e de -29 a +29 pontos do 3º ao 5º ano. Na situação de LS, os escores variam de -17 a +17 para ambas as escalas. No escore total, a pontuação pode variar de -38 a +39 no 2º ano e de -46 a +46 do 3º ao 5º ano.

### *Procedimento*

As crianças leram individualmente uma lista contendo 323 palavras de baixa frequência, que variavam em características psicolinguísticas (Pinheiro, 2007). A aplicação ocorreu em ambiente fornecido pela escola e em horário de aula. As palavras isoladas apareciam aleatoriamente no centro da tela de um computador portátil e as crianças foram instruídas a ler o mais

rápido e acuradamente possível. Foram necessárias duas sessões para a aplicação da lista, sendo que em uma terceira sessão as crianças responderam a outros instrumentos de avaliação (assunto que não será tratado aqui). Foram consideradas corretas as respostas que correspondiam à pronúncia da palavra, emitidas sem hesitação ou correção espontânea (emissão da pronúncia incorreta da palavra, seguida da correção pela criança).

## **Resultados**

### *Estatística Descritiva*

A Tabela 1 resume os resultados do estudo. Por ela, vemos que a distribuição de todas as medidas se distancia da normalidade, sendo assimétricas à esquerda. Note que os valores médios apresentados no 2º ano na EACOL não são diretamente comparáveis aos dos demais anos, por estarem em escalas de medida diferentes.

A Tabela 2 apresenta os escores necessários para se atingir os percentis 25, 50, 75 e 99 nas situações da escala e na proporção de acertos na lista de palavras. Observa-se que um efeito de teto é obtido na situação de LS no 2º e 5º anos no P50, enquanto que nos demais grupos os valores de teto são atingidos no P75. Na situação de LVA, no 5º ano, desaparecem as diferenças entre os percentis 50 e 75. Em todas as outras medidas, os valores apresentados no P75 se aproximam dos escores máximos, com exceção do escore total na escala e na lista de palavras, em que valores de teto são atingidos no P99 em sua maioria.

### *Estatística Inferencial*

Para as análises aqui reportadas, os escores na EACOL foram convertidos em escores-z, a fim de se comparar o desempenho dos alunos do 2º ano com os demais grupos.

Foi conduzida uma ANOVA univariada para verificar se a quantidade de escalas respondidas pelos professores interferiu nos escores conferidos aos alunos. Em média, os professores corrigiram 7,4 escalas ( $DP = 2,6$ , mín. = 1; máx. = 11). Não houve diferenças significativas no desempenho das crianças em função da quantidade

<sup>2</sup> A exceção fica por conta do item 8 da escala do 2º ano (“às vezes comete erros ao ler palavras novas”). Neste item, a criança recebe um escore de 0 ou 1 ponto. Essa decisão foi tomada para não penalizar muito os alunos das séries iniciais, que enfrentam dificuldades na leitura de itens pouco frequentes de uma maneira geral.

**Tabela 1**

**Média e Desvio Padrão dos Escores dos Participantes nas Situações da EACOL e na Lista de Palavras (% acerto) e Valores Máximos, Mínimos, de Assimetria e Curtose para cada Medida, em Função do Ano Escolar.**

Ano	Avaliação	N	Média	DP	Mínimo	Máximo	Assimetria	Curtose
2º	Leitura em Voz Alta	40	11,9	8,8	-13	22	-1,3	1,5
	Leitura Silenciosa	40	14,0	5,1	-3	17	-2,4	5,1
	Escore Total Escala	40	25,9	11,3	-6	39	-1,5	1,9
	Lista de Palavras	40	71,0	19,5	5,6	95,4	-1,6	2,7
3º	Leitura em Voz Alta	36	13,6	15,2	-19	29	-0,9	-0,2
	Leitura Silenciosa	36	12,3	8,4	-13	17	-2,4	5,1
	Escore Total Escala	36	25,9	22,5	-32	46	-1,5	1,7
	Lista de Palavras	36	82,7	12,2	50,8	95,1	-1,3	0,8
4º	Leitura em Voz Alta	34	15,0	15,2	-29	29	-1,4	1,1
	Leitura Silenciosa	34	9,8	9,5	-17	17	-1,7	2,2
	Escore Total Escala	34	24,8	23,1	-34	46	-1,4	1,1
	Lista de Palavras	34	84,6	8,8	65,0	97,5	-0,6	0,3
5º	Leitura em Voz Alta	27	21,1	12,5	-21	29	-2,2	4,5
	Leitura Silenciosa	27	13,4	6,0	-5	17	-1,8	2,4
	Escore Total Escala	27	34,5	18,2	-26	46	-2,2	4,1
	Lista de Palavras	27	91,3	6,2	69,0	97,8	-1,9	5,3
Total	Leitura em Voz Alta	137	15,1	13,4	-29	29	-1,2	0,8
	Leitura Silenciosa	137	12,5	7,5	-17	17	-2,3	4,8
	Escore Total Escala	137	27,5	19,3	-34	46	-1,6	2,3
	Lista de Palavras	137	81,5	15,1	5,6	97,8	-2,0	5,9

de escalas respondidas pelos professores [LVA: (F(8, 128) = 1,656,  $p = 0,115$ ); LS: (F(8, 128) = 1,604,  $p = 0,130$ ); escore total (F(8, 128) = 1,677,  $p = 0,110$ )]. Tampouco se encontrou efeito da quantidade de escalas respondidas quando dividimos a amostra por tipo de escola [LVA: pública (F(7, 69) = 1,793,  $p = 0,102$ ); particular

(F(3, 56) = 0,873,  $p = 0,461$ ); LS: pública (F(7, 69) = 1,541,  $p = 0,168$ ); particular (F(3, 56) = 0,891,  $p = 0,452$ ); escore total: pública (F(7, 69) = 1,619,  $p = 0,101$ ); particular (F(3, 56) = 0,832,  $p = 0,482$ )].

ANOVAs *One-Way* foram conduzidas para verificar se há efeito das classificações descritti-

Tabela 2

**Percentis Obtidos a partir dos Escores dos Participantes nas Diferentes Situações de Leitura Propostas pela Escala e para a Precisão (%) na Leitura da Lista de Palavras, em Função do Ano Escolar**

Ano	Percentis	Leitura em Voz Alta	Leitura Silenciosa	Escore Total Escala	Lista de Palavras
2º	25	9,0	13,0	24,0	65,8
	50	13,0	17,0	29,0	76,5
	75	19,0	-	32,0	83,4
	99	22,0	-	39,0	95,4
3º	25	4,0	11,5	15,0	77,6
	50	16,0	15,0	31,0	86,5
	75	26,5	17,0	42,0	92,3
	99	29,0	-	46,0	95,0
4º	25	10,5	5,0	17,5	79,9
	50	21,0	15,0	34,0	83,9
	75	27,0	17,0	40,5	90,7
	99	29,0	-	46,0	97,5
5º	25	23,0	13,0	36,5	87,9
	50	27,0	17,0	42,0	92,9
	75	27,0	-	44,0	96,3
	99	29,0	-	46,0	97,8

vas das situações da escala no desempenho das crianças na leitura de palavras. O esperado é que os bons leitores apresentem o melhor desempenho e os maus leitores obtenham os menores escores. Houve efeito geral das classificações descritivas no desempenho em leitura [LVA:  $F(3, 133) = 11,779, p < 0,001$ ; LS:  $F(3, 133) = 6,778, p < 0,001$ ; escore total:  $F(3, 133) = 8,131, p < 0,001$ ]. Análises de *post hoc* (teste de Tukey) mostraram que, na situação de LVA, as comparações somente não foram significativas entre os bons leitores e os razoavelmente bons e entre os maus leitores com os razoavelmente maus. Na situação de LS, os leitores razoavelmente maus tiveram desempenho pior que os leitores bons e razoavelmente bons, não havendo diferenças

significativas entre as demais categorias. No escore total, somente houve diferenças significativas entre os bons leitores e as demais categorias. Todas as diferenças significativas ocorreram nas direções esperadas.

Para atestar em que medida a avaliação dos professores condizem com o desempenho real das crianças, foram feitas correlações entre os escores das crianças nas situações da escala com os escores na leitura da lista de palavras. Pela Tabela 3, pode-se verificar que a maioria das correlações foi significativa e apresentou valores moderados. As maiores correlações foram obtidas nas escolas particulares na situação de LVA e nas escolas públicas na situação de LS, em todos os anos investigados.

Tabela 3

Valores da Correlação de Pearson entre as Situações de Leitura em Voz Alta, Leitura Silenciosa e Escore Total na Escala com a Precisão na Leitura em Voz Alta de Palavras Isoladas

Amostra	Ano	Leitura em voz alta	Leitura silenciosa	Escore total
Escola Particular	2º ano	0,80**	NS	0,78**
	3º ano	0,82**	NS	0,79**
	4º ano	0,57**	0,54**	0,59**
	5º ano	0,61*	0,66*	0,64*
	Total	0,59**	0,30*	0,54*
Escola Pública	2º ano	NS	0,70**	0,42*
	3º ano	0,54**	0,57**	0,58**
	4º ano	0,67**	0,61**	0,69**
	5º ano	0,56*	0,54*	0,56*
	Total	0,29*	0,50**	0,41**
Amostra geral	2º ano	0,50**	0,53**	0,62**
	3º ano	0,59**	0,56**	0,61**
	4º ano	0,39*	0,32 <sup>1</sup>	0,39*
	5º ano	0,73**	0,76**	0,75**
	Total	0,43**	0,42**	0,48**

Nota. NS = não significativo; <sup>1</sup>efeito marginalmente significativo  $p < 0,06$ .

\*\* $p < 0,01$ ; \* $p < 0,05$ .

Tabela 4

Proporção de Crianças nas Diferentes Classificações Descritivas Propostas pelas Situações de Leitura da Escala, em Função do Percentil Obtido pelo Escore na Lista de Palavras (em negrito, a proporção de crianças classificadas de forma idêntica na lista de palavras e nas situações da escala)

Percentil na lista de palavras	Classificação na situação de leitura em voz alta			
	Lê mal	LR-	LR+	Lê bem
25	<b>63%</b>	85%	16%	16%
50	25%	<b>15%</b>	32%	25%
75	13%	0%	<b>26%</b>	29%
99	0%	0%	26%	<b>29%</b>
	Classificação na situação de leitura silenciosa			
25	<b>80%</b>	71%	50%	17%
50	0%	<b>14%</b>	29%	27%
75	20%	14%	<b>0%</b>	29%
99	0%	0%	21%	<b>27%</b>
	Classificação no escore total			
25	<b>71%</b>	88%	37%	16%
50	14%	<b>13%</b>	37%	25%
75	14%	0%	<b>11%</b>	30%
99	0%	0%	16%	<b>29%</b>

Nota. LR- (Lê razoavelmente mal); LR+ (Lê razoavelmente bem).



Uma análise qualitativa da concordância entre a classificação dos professores e os percentis na lista de palavras é mostrada na Tabela 4. Observa-se uma maior concordância entre as crianças classificadas como maus leitores com o percentil 25 da lista de palavras. As crianças classificadas como leitores razoavelmente maus oscilam entre os percentis 25 e 50, mas estão em uma maior proporção no percentil 25. Uma maior dispersão ocorre nas demais categorias. Assim, entre 41% e 79% das crianças que apresentaram percentil 25 ou 50 na lista de palavras foram classificadas pelos professores como leitores bons ou razoavelmente bons.

Gráficos de dispersão confirmam o que foi exposto. Nas Figuras 1 e 2, no sentido anti-horário, temos no 1º quadrante altos escores tanto na lista de palavras quanto nas situações da escala; no 2º quadrante, baixos escores na escala e altos escores na lista de palavras; no 3º quadrante, baixos escores em ambas as medidas; e no

4º quadrante, altos escores na escala, mas baixos na lista de palavras. A linha vertical dos gráficos representa o escore médio na EACOL, enquanto que a linha horizontal incide sobre a precisão média na lista de palavras. O esperado é que os escores dos maus leitores e dos razoavelmente maus tendam a se concentrar no 3º quadrante, enquanto que os mais proficientes se concentrem no 1º quadrante. Como previsto, nota-se uma tendência, em ambas as situações, de que leitores considerados maus ou razoavelmente maus fiquem alocados no 3º quadrante. Entretanto, contrariando as expectativas, os escores dos bons leitores tendem a se dispersar também para o 4º quadrante em ambas as Figuras, apesar de haver uma grande proporção que, como esperado, se agrupa no 1º quadrante. Na situação de LVA, os leitores razoavelmente bons tendem a apresentar escores medianos na EACOL e altos escores na lista de palavras, enquanto que na situação de LS há uma maior dispersão desses leitores no 2º e 3º quadrantes.

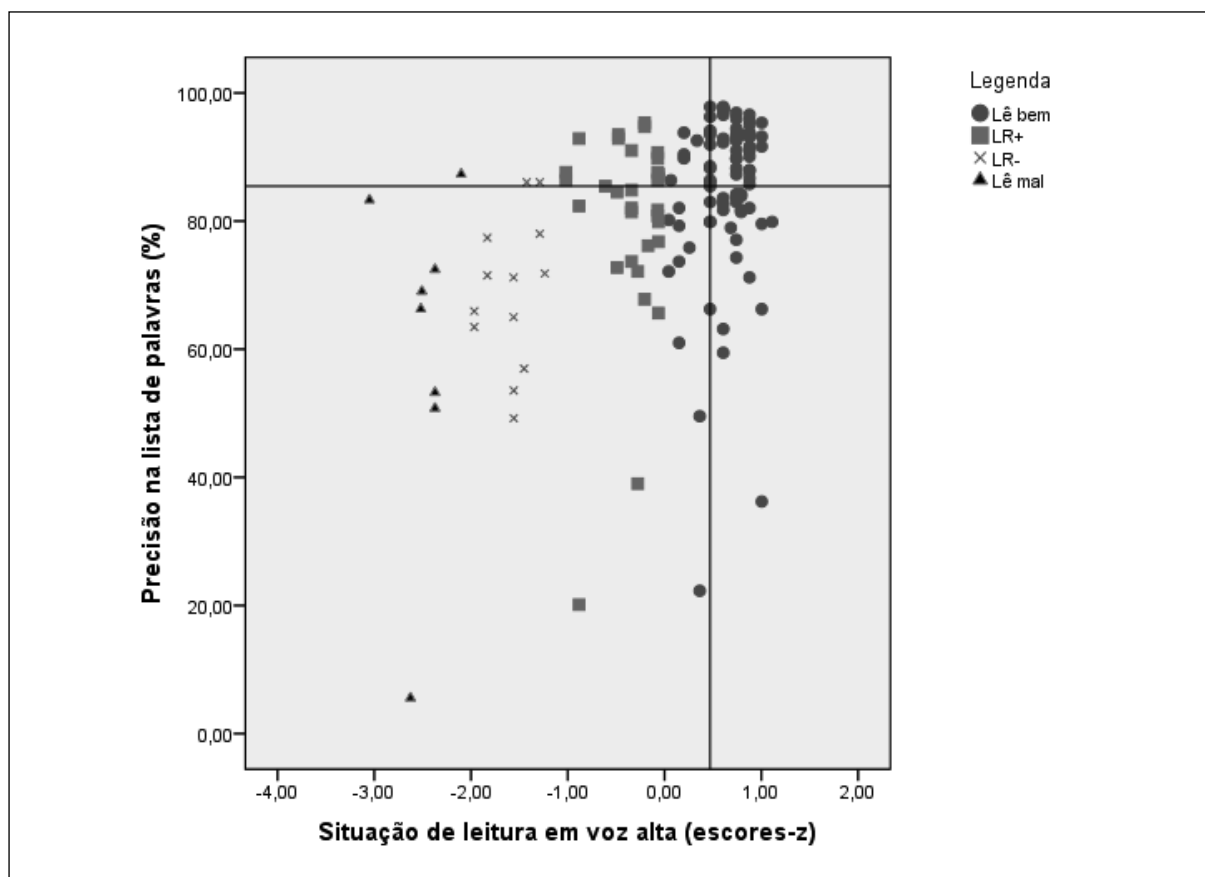
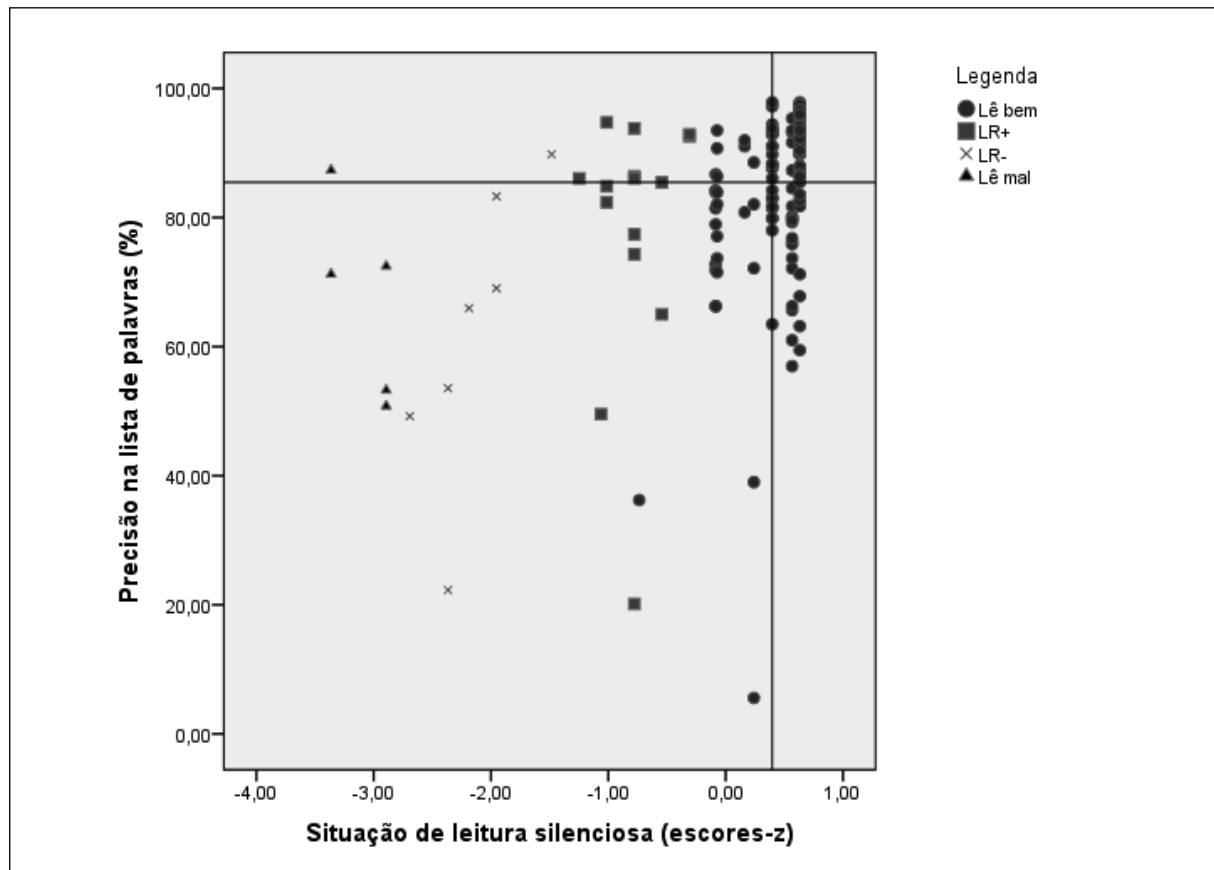


Figura 1. Escores das crianças (escores-z) na situação de leitura em voz alta, em função do desempenho na lista de palavras (% acerto) e da classificação descritiva proposta pela escala (amostra geral).



**Figura 2.** Escores das crianças (escores-z) na situação de leitura silenciosa, em função do desempenho na lista de palavras (% acerto) e da classificação descritiva proposta pela escala (amostra geral).

## Discussão

O estudo buscou coletar evidências de validade para a Escala da Avaliação da Competência da Leitura pelo Professor (EACOL), que está em fase de desenvolvimento. Para esta finalidade, o principal procedimento adotado foi investigar a relação entre o desempenho das crianças na EACOL e seus escores na leitura de uma lista de palavras de baixa frequência. Os resultados obtidos estão de acordo com a literatura nacional (p. ex., Salles & Parente, 2007) e internacional (p. ex., Hecht & Greenfield, 2002) de que, em geral, correlações moderadas são encontradas entre as medidas. Entretanto, os dados também apontam para necessidades de ajustes nos itens da escala, de modo a aumentar o grau de concordância entre a avaliação dos professores e o desempenho real das crianças.

As ANOVAS mostraram que a classificação descritiva da escala produziu diferenças significativas no desempenho em leitura de palavras da

amostra. A direção das diferenças entre as médias mostrou que os leitores das categorias superiores (lê bem ou razoavelmente bem) obtiveram maiores escores na leitura de palavras do que as categorias inferiores (lê mal ou razoavelmente mal), indicando que houve uma correspondência entre a avaliação feita pelos professores e diferenças reais no desempenho das crianças. Entretanto, as categorias da escala nem sempre produziram diferenças significativas entre si. Desse modo, na situação de LVA, modificações devem ser feitas de modo a produzir uma maior diferenciação entre os leitores dos níveis extremos de habilidade em relação aos intermediários (bons leitores vs razoavelmente bons e maus leitores vs razoavelmente maus). Na situação de LS, o problema maior está na produção de diferenças com os maus leitores. Já no escore total, apenas os bons leitores se diferenciam das demais classificações.

A concordância entre as avaliações também foi demonstrada por meio de correlações

com a lista de palavras que, em sua maioria, foram significativas, embora moderadas. Na situação de LVA não houve correlações significativas apenas no 2º ano das escolas públicas, enquanto que na situação de LS as correlações significativas não ocorreram somente para o 2º e 3º anos da escola particular. O escore total também apresentou bons resultados, havendo em todos os anos correlações significativas, tanto nas escolas públicas quanto na particular e na amostra geral.

A ausência de correlações mostrou-se mais frequente na situação de LS, o que é previsível, uma vez que a medida de comparação que está sendo considerada é a de leitura em voz alta de palavras, que favorece as correlações com a situação de LVA da escala (Cogo-Moreira et al., 2012). Mesmo assim, em geral, as duas medidas produziram correlações moderadas. Isso implica que, embora as situações de LVA e LS sejam de naturezas distintas, pode-se considerar que o instrumento mostrou-se útil na detecção de alunos que apresentam diferentes níveis de desempenho na leitura silenciosa, baseando-se em seu resultado na leitura em voz alta. Apesar disso, reconhece-se a necessidade de se utilizar uma medida específica de leitura silenciosa para servir como critério de comparação com a situação de LS da escala.

Outro fato que corrobora o que foi acima exposto é a presença de escores muito elevados na situação de LS. Enquanto que na LVA os escores tendiam a se dispersar mais em torno da média, na LS havia uma tendência maior para o efeito de teto (ver Tabelas 1 e 2), o que repercutiu na ausência do percentil 99 em todos os anos escolares e do percentil 75 no 2º ano. Para essa divergência, há pelo menos duas hipóteses: ou os professores têm dificuldade de julgar a leitura silenciosa dos alunos (por não ser diretamente observável, como é o caso da leitura em voz alta), ou as crianças tendem a obter melhores resultados na compreensão de materiais lidos silenciosamente devido às próprias características da avaliação. Desta forma, é possível que durante a leitura silenciosa, a criança possa estar em contato com o texto por mais tempo, tornando mais provável o uso de estratégias compensató-

rias de compreensão, como a releitura de trechos obscuros (Afflerbach & Cho, 2009).

A análise do grau de concordância entre a precisão da leitura e a avaliação dos professores demonstrou a necessidade de ajustes nos itens da EACOL, pois a classificação descritiva proposta pela escala produziu resultados mais consistentes apenas para as crianças com desempenho inferior (Tabela 4). Assim, observou-se uma concordância maior entre os leitores com os piores níveis de habilidade (lê mal/percentil 25) em todas as situações da escala, variando de 63% a 80%. Este resultado está em consonância com a literatura nacional (p. ex., Capellini et al., 2004). Nesse sentido, as classificações propostas pela escala mostram-se mais úteis quando o interesse é, por exemplo, a triagem. Ou seja, há uma maior margem de segurança de que uma criança classificada como má leitora de fato apresente problemas na leitura. No entanto, houve menor concordância nos níveis mais elevados (leitores bons e razoavelmente bons), sugerindo que os resultados de alunos com alto desempenho na escola foram superestimados. Resultado semelhante foi encontrado por Salles e Parente (2007), utilizando a mesma escala. Essa configuração de resultados indica que melhorias devem ser feitas nos itens de modo aumentar a concordância nas classificações e percentis mais elevados – e, portanto, aumentar a precisão da medida.

Por fim, a consideração feita anteriormente (ver Introdução) de que o método utilizado no trabalho Salles e Parente (2007) poderia ter contribuído para a redução da concordância na EACOL se confirmou. As autoras reportaram análises apenas para o escore total da escala, assim como utilizaram a escrita como medida de comparação do desempenho das crianças com a avaliação das professoras. Como resultado, encontraram que as classificações descritivas propostas pela EACOL produziram dissociações tanto para os leitores maus ou razoavelmente maus quanto para os bons ou razoavelmente bons. Neste estudo, apesar de em geral as dissociações terem ocorrido com maior frequência apenas entre os leitores bons e razoavelmente bons, ficou evidente que os resultados para as

situações de LVA e LS não foram idênticos: a leitura silenciosa foi mais problemática em termos de variabilidade dos escores<sup>3</sup>, apresentando um maior efeito de teto. Por outro lado, a concordância entre os maus leitores foi maior para a situação de LS (Figuras 1 e 2). Isto torna evidente a importância de conduzir análises separadas para cada situação na escala. Finalmente, o uso da tarefa de leitura em voz alta como medida de critério produziu mais correlações significativas para a situação de LVA, o que reforça a noção de que medidas do mesmo construto devem ser usadas como critério durante a investigação da qualidade dos instrumentos.

### Conclusão

Os resultados do estudo apontam que a EACOL é uma ferramenta útil para a triagem das dificuldades de leitura em crianças cursando os anos iniciais. Isso foi demonstrado quando observamos que o desempenho das crianças na tarefa de leitura variou conforme a atribuição de desempenho feita pelos professores, sendo que os leitores considerados bons tiveram desempenho melhor nesta tarefa dos que os considerados “maus” ou razoáveis. Apesar disso, reconhecemos a necessidade de ajustes nos itens, de modo a controlar os fatores que reduzem a concordância entre a avaliação dos professores e do desempenho efetivo das crianças.

O formato das respostas aos itens pode ser um desses fatores. Ao ter que responder “sim” ou “não” para determinado item, o professor pode se isentar da resposta (como ocorreu em alguns casos na amostra) ou marcar aleatoriamente em uma das opções. Uma alternativa seria a criação de uma escala likert de três a cinco níveis. Outra dificuldade parece estar na inteligibilidade dos itens da categoria “lê razoavelmente”. Muitos desses itens apresentam o termo “às vezes” e isso pode confundir e reduzir a acurácia da ava-

liação dos professores. A solução pode ser a exclusão deste termo nas perguntas e sua inclusão apenas nas respostas, que passaria a contar com quatro opções: sim, não e às vezes. Finalmente, outra possibilidade seria adicionar a opção “não sei avaliar” (Pinheiro, no prelo). Estas são medidas a serem adotadas nas próximas pesquisas com a finalidade de melhorar o poder preditivo da escala.

Além de tarefas de leitura silenciosa, os próximos estudos devem incluir medidas de compreensão de texto, pois, como lembram Salles e Parente (2007), os professores podem basear seu julgamento na compreensão dos materiais lidos pela criança. Conforme vimos, as autoras encontram correlações moderadas entre o escore total na EACOL com compreensão de texto, mas não conduziram análises separadas para as situações na escala. De forma semelhante, Cogo-Moreira et al. (2012) correlacionaram o escore total na EACOL com a acurácia na leitura de textos, que é um construto relacionado, mas não uma medida direta da compreensão textual. Estas resoluções podem aumentar a confiança no poder da escala em classificar corretamente os alunos nas situações de LVA e LS, o que pode gerar resultados promissores.

### Referências

- Afflerbach, P., & Cho, B. Y. (2009). Identifying and describing constructively responsive comprehension strategies in new and traditional forms of reading. In S. E. Israel & G. G. Duffy (Eds.), *Handbook of research on reading comprehension* (pp. 69-90). New York: Taylor & Francis Group.
- Bennett, R. E., Gottesman, R. L., Rock, D. A., & Cerullo, F. (1993). Influence of behavior perceptions and gender on teachers' judgments of students' academic skill. *Journal of Educational Psychology, 85*, 347-356.
- Capellini, S. A., Tonelotto, J. M. F., & Ciasca, S. M. (2004). Medidas de desempenho escolar: Avaliação formal e opinião de professores. *Estudos de Psicologia (Campinas), 21*(2), 79-90.
- Cogo-Moreira, H., Ploubidis, G. B., De Ávila, C. R. B., Mari, J. J., & Pinheiro, A. M. V. (2012). EACOL (scale of evaluation of reading competency

<sup>3</sup> Este resultado corrobora com o obtido por Cogo-Moreira et al. (2012), que demonstrou que a estrutura latente da situação de leitura em voz alta foi mais adequada que a situação de silenciosa na amostra de Belo Horizonte.

- by the teacher): Evidence of concurrent and discriminant validity. *Neuropsychiatric Diseases and Treatment*, 8, 443-454.
- Gutierrez, L., & Tomasi, E. (2011). *Prevalência de dislexia e fatores associados em Escolares do 1º ao 4º anos*. Trabalho apresentado no XVI Congresso Internacional de la Asociación de Lingüística y Filología de la América Latina – ALFAL, Madrid, Espanha.
- Hecht, S. A., & Greenfield, D. B. (2002). Explaining the predictive accuracy of teacher judgments of their students' reading achievement: The role of gender, classroom behavior, and emergent literacy skills in a longitudinal sample of children exposed to poverty. *Reading and Writing: An Interdisciplinary Journal*, 15, 789-809.
- Hoge, R. D., & Caladarci, T. (1989). Teacher-based judgments of academic achievement: A review of literature. *Review of Educational Research*, 59, 297-313.
- Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira. (2011). *Sistema Nacional de Avaliação da Educação Básica – SAEB. Resultados Saeb/Prova Brasil 2011 – Língua Portuguesa*. Recuperado em <http://portal.inep.gov.br/web/prova-brasil-e-saeb/edicao-2011>
- Kenyon, R. (2003, September). Facts and statistics on learning disabilities and literacy. In *Project Director – Florida's Bridge to Practice: A Project of Florida Human Resources Development*. Hollywood, FL: Florida Human Resources Development.
- Pasquali, L. (2009). Psicometria [Edição especial]. *Revista da Escola de Enfermagem da USP*, 43, 992-999.
- Pinheiro, A. M. V. (2001). Heterogeneidade entre leitores julgados competentes pelos professores. *Psicologia: Reflexão e Crítica*, 14(3), 537-551.
- Pinheiro, A. M. V. (2007). Banco de palavras de baixa frequência de ocorrência, para crianças brasileiras da 1ª à 4ª série do Ensino Fundamental, classificadas em termos de estrutura silábica, número de letras e regularidade para leitura e para escrita. In I. Sim-Sim & F. L. Vianna, *Para avaliação do desempenho da leitura* (pp. 119-130). Lisboa, Portugal: Ministério da Educação.
- Pinheiro, A. M. V. (no prelo). EACOL: Uma escala da avaliação da competência de leitura de alunos pelo Professor: os primeiros estudos de validação. In F. L. Viana, R. Ramos, E. Coquet, & M. Martins (Eds.), *Leitura, literatura infantil e ilustração: Investigação e prática docente 9*. Braga, Portugal: Centro de Investigação em Estudos da Criança, Universidade do Minho.
- Pinheiro, A. M. V., & Costa, A. E. B. (2005, mar.). Escala da avaliação de competência em leitura pelo professor. In Instituto Mineiro de Avaliação Psicológica (Org.), *VII Encontro Mineiro de Avaliação Psicológica. Resumos* (p. 27). Belo Horizonte, MG: Instituto Mineiro de Avaliação Psicológica.
- Pinheiro, A. M. V., & Costa, A. E. B. (2011). *Construção e Validação da Escala de Avaliação de Competência de Leitura de Alunos pelo Professor – EACOL*. Trabalho apresentado no **Seminário Internacional de Alfabetização na Perspectiva da Psicologia Cognitiva da Leitura**, São Paulo, SP, Brasil.
- Salles, J. F., & Parente, M. A. M. P. (2007). Relação entre desempenho infantil em linguagem escrita e percepção do professor. *Cadernos de Pesquisa*, 37(132), 687-709.
- Urbina, S. (2007). *Fundamentos da testagem psicológica* (C. Dornelles, Trad.). Porto Alegre, RS: Artmed.

Recebido: 08/03/2013  
1ª revisão: 20/05/2013  
Aceite final: 31/05/2013