

Evidências de validade por processo de resposta no Cloze

Neide de Brito Cunha,^{Orcid, I, *} Acácia Aparecida Angeli dos Santos,^{Orcid, II} Katya Luciane de Oliveira^{Orcid, III}

^I Universidade do Vale do Sapucaí, Pouso Alegre, MG, Brasil / Centro Paula Souza - Fatec Bragança Paulista, Bragança Paulista, SP, Brasil

^{II} Universidade São Francisco, Campinas, SP, Brasil

^{III} Universidade Estadual de Londrina, Londrina, PR, Brasil

Resumo

O objetivo deste estudo foi explorar as diferenças qualitativas nos erros apresentados em um Teste de Cloze nas pontuações extremas de crianças e encontrar evidências de validade por processo de resposta. Participaram 277 alunos matriculados nos 3º, 4º e 5º anos do ensino fundamental público do Estado de São Paulo. Para a coleta de dados, que ocorreu de forma coletiva, foi empregado o teste de Cloze estruturado de forma tradicional, na qual se omitem todos os quintos vocábulos do texto e no local coloca-se um traço de tamanho proporcional ao da palavra omitida, bem como a correção verbatim. Os resultados evidenciaram evidências de validade baseadas no processo de resposta, indicando também divergências nas respostas entre os alunos que pontuaram menos e aqueles que obtiveram maiores pontuações. A discussão dos dados aponta para algumas implicações acerca dos processos cognitivos envolvidos na emissão das respostas das crianças.

Palavras-chave: processo de resposta; técnica de cloze; ensino fundamental.

Validity evidence based on response process in the Cloze test

Abstract

This study aims to explore the errors differences presented in a Cloze test of kids' extreme punctuations and find evidences of the validity for answer process. 277 students from third, fourth and fifth grades of public education of São Paulo participated on the survey. For data collection, which occurred collectively, it was held a traditional Cloze test that substitutes the fifth words for a dash of the same size of the word hidden, as well as verbatim correction. Results showed evidences of validity based on answer process, indicating answer divergences between students with high punctuation and low punctuation. Data discussion points some implications about cognitive process involved on the emission of kids' answers.

Keywords: error analysis; Cloze technique; elementary school.

A validade é o conceito central da psicometria e está relacionada à interpretabilidade dos escores de um teste, conforme indicam a *American Psychological Association*, *American Educational Research Association* and *National Council on Measurement in Education* (1999). E a qualidade de um teste psicológico está diretamente relacionada às suas evidências de validade, que são as confirmações de que o instrumento realmente mede aquilo a que se propõe (ANASTASI; URBINA, 2000).

A partir da publicação do documento *Standards for Educational and Psychological Testing* (APA; AERA, NMCE, 1999), o conceito de validade sofreu uma reformulação, proposta por Messick (1995). O pesquisador indicou que todos os tipos de validade se resumiam à validade de construto e que a validade de um instrumento estaria vinculada a um conjunto de evidências científicas requeridas, à interpretação proposta no instrumento e à sua interpretação e aplicabilidade para o contexto proposto. Esse novo documento definiu cinco fontes usadas ao evidenciar a validade de um teste: (a) evidência com base no conteúdo; (b) evidência com base no processo de resposta; (c) evidência com base na estrutura interna; (d) evidência com base na relação com variáveis externas; (e) evidências baseadas nas consequências da testagem.

Prieto e Delgado (1999), na busca de instrumentos de medida que se ajustassem a um marco que não fosse estritamente estatístico, salientaram a importâncias do estudo dos processos cognitivos envolvidos na resolução dos

itens, por serem uma importante fonte de informação. Essa metodologia inclui predominantemente a análise de protocolos de resposta, entrevistas ou outros procedimentos que permitam a análise individualizada do par sujeito/item.

Esses tipos de investigação são úteis para a referência de modelos cognitivos porque permitem mostrar possíveis discrepâncias entre grupos, relativas ao processo de resposta. Desse modo, são formulações que reúnem a representação formal e a psicológica, decompondo a dificuldade dos itens em parâmetros representativos de seus componentes (OLIDEN, 2003). Macías (2007) acrescenta que a evidência baseada no processo de resposta reúne argumentos que evidenciam as consistências entre as respostas ao instrumento (evidência empírica) e os processos para os quais se estabelece, desde a teoria, que assumam os avaliados nas tarefas propostas. Nessa concepção, os argumentos relacionados com esse aspecto se encontram no marco conceptual, e os processos seguidos pelos avaliados para se chegar às respostas que são formuladas devem ser explicitados.

No Brasil, são poucas as pesquisas que buscaram esse tipo de evidência de validade. Entre elas, pode-se citar a de Primi (1995), que procurou explicar a complexidade dos problemas de raciocínio analógico geométrico presentes em testes de inteligência, bem como as diferenças individuais, analisando ocorrências particulares de interferências entre as fases do raciocínio (ou componentes) ligadas, por um lado, à percepção e, por outro, à abstração. As respostas e o tempo de reação foram gravados pelo programa informatizado a partir do qual a prova de racio-

*Endereço para correspondência: Universidade do Vale do Sapucaí – PROPPES - Avenida Prof. Tuany Toledo, 470 - CEP 37554-210 - Pouso Alegre, MG. E-mail: neidedebritocunha@gmail.com, acacia.santos@usf.edu.br, katyael@gmail.com



cínio foi aplicada. A conclusão da pesquisa indicou que a dinâmica entre as diferenças individuais e a complexidade da estimulação interfere na complexidade da tarefa.

Já em sua tese, Primi (1998) desenvolveu um modelo para indicar fatores de complexidade em problemas de raciocínio analítico a partir da psicologia cognitiva e de estudos sobre os componentes de processamento desse tipo de raciocínio. Utilizando a Teoria de Resposta ao Item (TRI), o autor criou uma escala que indica a proficiência em problemas de distintos níveis de exigência da memória de curto prazo e do gerenciamento metacognitivo, que permite também a comparação do desempenho relativo quanto à persistência, flexibilidade e gerenciamento eficaz do tempo.

Rueda e Sisto (2008) realizaram um estudo com o objetivo de buscar evidências de validade para o Teste Pictórico de Memória. Para levantar as evidências pelo processo de resposta, os autores partiram do pressuposto de que os ambientes presentes no teste (céu, terra e água) gerariam níveis de dificuldade significativamente diferenciados, afetando a recuperação da informação. A partir da confirmação de tal premissa, concluíram que o teste apresentava as evidências desejadas.

Também recorrendo à análise do processo de resposta, Cunha e Santos (2009) buscaram evidências de validade pelo processo de resposta para um teste de Cloze. Participaram da pesquisa 266 crianças, das quais 52,6% eram meninos. As idades variaram de 8 a 13 anos. Representaram 44,4% da amostra os alunos do quarto ano e 55,6% do quinto. O teste de Cloze, intitulado “A princesa e o fantasma”, apresentava evidências de validade de critério e estava adequado à escolaridade dos participantes. O texto tinha o quinto vocábulo de cada frase omitido, o qual deveria ser escrito numa lacuna com tamanho proporcional à palavra que deveria ser escrita. Tendo por base considerações sobre as evidências de validade por processo de resposta, foram analisados os erros cometidos num teste de Cloze. Participaram 266 crianças, de ambos os sexos, entre 8 e 13 anos, de terceiras e quartas séries do Ensino Fundamental de escolas públicas e particulares de São Paulo. Foram formados dois grupos de protocolos com as crianças que tiveram, respectivamente, as melhores e as piores pontuações, para a construção de uma escala dos tipos de erros cometidos, a saber: branco, fonológico, lexical, sintático e semântico. Os resultados mostraram que as crianças com médias mais altas cometeram mais erros lexicais, e as com médias mais baixas, erros semânticos. A evidência de validade por processo de resposta foi encontrada por meio da avaliação da homogeneidade na distribuição dos tipos de erros.

Esses estudos foram os únicos recuperados na literatura nacional, utilizando-se como palavra-chave o processo de resposta, o que confirma a afirmação de que estudos usando essa evidência não são comuns (ANASTASI; URBINA, 2000). Cabe salientar que Anastasi (1961/1977) advertiu que a variedade de técnicas disponíveis para a obtenção da validade de um instrumento poderia ser algo perigoso, pois um autor poderia usar diferentes métodos para encontrar evidências de validade de seu instrumento

e publicar apenas aquele que lhe tivesse sido mais positivo. Essa afirmação aponta a necessidade de que, cada vez mais, sejam levantadas evidências de validade dos instrumentos psicológicos, mesmo daqueles que já possuam alguns estudos desse tipo. Importante salientar que, para que seja considerada como uma evidência de validade, é essencial que as propostas explicativas dos processos mentais subjacentes forneçam uma previsão dos acertos em razão das características das demandas da tarefa (NUÑES; PRIMI, 2010; PADILLA; BENÍTEZ, 2014).

Sob essa perspectiva, o presente estudo se propõe a buscar evidências de validade com base no processo de resposta para outro Teste de Cloze, cujo título é “Uma vingança infeliz”, na medida em que há pressupostos teóricos que fundamentam a existência de erros e acertos nas lacunas do teste usado. Cabe ressaltar que esse é um teste que já foi utilizado em vários estudos brasileiros. Dentre eles, serão citados os que buscaram evidências de validade.

Quanto às evidências de validade baseadas na relação com outras medidas que avaliam construtos relacionados, foram encontradas na literatura pesquisas como a de Santos, Sisto e Noronha (2010), que correlacionaram os escores do Teste de Cloze com o Teste TONI 3-Forma, que é um teste de inteligência não-verbal. Santos e Cunha (2012) compararam a pontuação do Cloze com a do Questionário de Avaliação da Consciência Metatextual, do Teste de Desempenho Escolar (TDE), por Mota e Santos (2014), do Questionário de Avaliação da Consciência Metatextual- versão preliminar, por Cunha e Santos (2014).

Suehiro e Magalhães (2014), por sua vez, utilizaram o teste de Cloze, a Escala de Avaliação da Escrita (EAVE) e a Escala de Reconhecimento de Palavras (EREP) e constataram validade de construto convergente-discriminante entre os instrumentos, sendo que o Cloze se mostrou sensível para captar a relação existente entre a aprendizagem da escrita e a compreensão de leitura. Os alunos com baixo desempenho na leitura foram os com pior desempenho na escrita, abarcando uma evidência de validade de critério concorrente para os instrumentos empregados.

Também Suehiro e Santos (2015) buscaram evidências de validade entre o Cloze, o Roteiro de Avaliação da Consciência Fonológica e a Prova de Consciência Fonológica por Produção Oral. Os resultados obtidos confirmaram a existência de validade de critério por grupos contrastantes, uma vez que as crianças que obtiveram escores altos na Prova de Consciência Fonológica por Produção Oral também apresentaram desempenhos significativamente superiores no Cloze e no Roteiro de Avaliação da Consciência Fonológica.

Quanto às evidências de validade de critério, elas foram encontradas em estudos, usando-se como critério o ano escolar. Com esse propósito foram localizadas as publicações de Santos e Cunha (2012), Mota e Santos (2014) e Cunha e Santos (2014).

De acordo com as pesquisas recuperadas, há vários estudos de validade para esse teste, no entanto não há evidências de validade por processo de resposta. Nesse sentido, o presente estudo se valerá dos conceitos de García

(1998), que descreveu o modelo da psicologia da leitura, com as contribuições da psicologia cognitiva e da teoria do processamento da informação. Esse modelo permitiu o desenho e a identificação de quatro grandes módulos com os submódulos respectivos, correspondentes a diversos processos participantes em toda tarefa da leitura.

O primeiro deles é o módulo perceptivo, que contém os processos de extração de informação, por meio da memória icônica e dos processos de memória de trabalho nos quais se efetuam tarefas de reconhecimento e análise linguística. Já o segundo, o módulo léxico, consiste na recuperação do conceito associado à unidade linguística, ou recuperação léxica. O terceiro módulo, denominado sintático, inclui os processos que contêm as estratégias de reconhecimento gramatical, tais como a consideração da ordem das palavras, do papel das palavras funcionais, do significado das palavras e do uso dos sinais de pontuação, entre outras. Tudo isso levará à proposta de diversos modelos de processamento sintático que se ajustem aos elementos léxicos e ao sentido do que for expresso na leitura. Finalmente, o módulo semântico contém os conhecimentos prévios referentes ao significado declarativo e procedimental, que exige a integração do léxico e das distintas palavras num todo coerente que permita a extração do significado da mensagem, além do que cada uma de suas partes componentes represente.

Além desse estudo da psicologia da leitura, há outros que auxiliam no entendimento do processamento da informação para a compreensão da leitura, especificamente nos testes de Cloze. Cohen (1975), Page (1975) e Santos (1981) indicaram em suas pesquisas a existência de diferenças de desempenho no teste de Cloze relacionadas com o conteúdo, como a familiaridade com o assunto abordado, que poderia aumentar a probabilidade de acerto das respostas.

Foram também apontados fatores intrínsecos ao texto que interferem na sua compreensão, como local da pista do contexto, número de sílabas do período, recuperação de palavras de conteúdo, extensão da palavra a ser encontrada, número de respostas possíveis para a lacuna e número de formas possíveis a serem consideradas (ABRAHAM; CHAPELLE, 1992). Já Santos et al. (2002) argumentaram que o tipo de processamento requerido e a habilidade avaliada pelo teste de Cloze dependem basicamente do método adotado na criação das lacunas. Nessa perspectiva, o acerto em determinado tipo de lacuna pode depender do uso do contexto, enquanto o acerto de outro tipo de lacuna pode estar ligado à aplicação de conhecimento prévio (conhecimento lexical), favorecendo a posição dos pesquisadores que ressaltam o processamento local imediato ou externo ao texto (conhecimento vocabular prévio) como elementos centrais da resposta.

Diante do levantamento da literatura e com base no modelo de García (1988) e nas pesquisas levantadas sobre o Cloze, este estudo foi proposto. Assim, seu objetivo é o de explorar as diferenças qualitativas nos erros apresentados em um Teste de Cloze nas pontuações extremas de crianças e encontrar evidências de validade por processo de resposta.

Método

Participantes

Participaram 277 alunos de uma escola pública do interior de São Paulo. Do total, 148 eram meninos e 129 meninas, cujas idades variaram de 7 a 11 anos, com média 9,22 e desvio padrão 0,85. Os estudantes estavam cursando o terceiro ($n=70$, 25,3%), quarto ($n=104$, 37,5%) e quinto ($n=103$, 37,2) anos do Ensino Fundamental I.

Instrumento

Teste de Cloze (SANTOS, 2005) “Uma vingança infeliz” – trecho exemplificativo: “A mãe de Pedro _____ brava com ele e o castigou. _____ lhe disse que ao _____ uma fotografia ele também _____ destruindo uma lembrança.”

Esse teste foi estruturado na forma do Cloze tradicional, com os quintos vocábulos omitidos, substituídos por um traço proporcional ao tamanho da palavra omitida, que deveria ser recuperada pelo leitor para restituir o sentido completo à seleção. Especialmente montado para ser utilizado com crianças da faixa etária do ensino fundamental, também apresenta evidências de validade de critério. Para tanto, foi submetido a um estudo piloto com 314 alunos, de ambos os sexos, de segunda a quarta séries. Os resultados obtidos indicaram que o texto, com total de 103 palavras, das quais foram omitidas 15, mostrou-se adequado para ser utilizado na amostra pretendida. A análise da variância indicou haver diferença estatisticamente significativa entre as séries [$F(3,314)=55,75$; $p<0,001$]. Essa diferença foi justificada pelo teste de *post-hoc* de Tukey, que separou os escores das crianças nos grupos correspondentes às séries frequentadas. A análise da consistência interna apresentou índices de precisão satisfatórios, de acordo com as recomendações do Conselho Federal de Psicologia (2003), acima de 0,70, pois o alfa de Cronbach foi de 0,83 para as crianças estudadas. Procedeu-se, também, à análise da consistência interna por série e obteve-se o índice de 0,85 para a segunda série, 0,70 para a terceira série e 0,72 para a quarta série.

Crerios de Correção

Foi atribuído um ponto para cada acerto, sendo consideradas corretas as palavras que fossem exatamente as mesmas usadas pelo autor do texto. Essa forma de correção é denominada literal ou verbatim e é recomendada por evitar o subjetivismo na avaliação. Só pode haver 15 acertos, correspondentes ao número de omissões.

Procedimentos

A presente pesquisa foi analisada e aprovada pelo Comitê de Ética da Universidade São Francisco, sob o protocolo de Certificado de Apresentação para Apreciação Ética (CAAE) nº 0277.0.142.000-10. A coleta de dados ocorreu em instituições cujas diretoras autorizaram sua realização e com aqueles alunos cujos pais assinaram o Termo de Consentimento Esclarecido, permitindo a participação do(a) filho(a).

A coleta de dados ocorreu em situação de sala de aula, sendo realizada pela própria professora, num dia normal de atividades escolares. Importante salientar que elas seguiram um roteiro especialmente elaborado, com instruções detalhadas, por meio das quais se procurou garantir o mesmo padrão de aplicação dos testes. Aos alunos era solicitado que lessem a história até o seu final, antes de começarem a escrever as palavras que haviam sido omitidas nos traços que as substituíram.

Análise de dados

Foram utilizadas as estatísticas descritivas em protocolos separados pelo número de acertos. Além disso, recorreu-se à prova de qui-quadrado para avaliar a homogeneidade da distribuição dos pontos. Foram formados os

grupos (1988), considerando-se: 1 – branco; 2 – erro fonológico (erros ortográficos e de acentuação); 3 – erro lexical (uso de sinônimo); 4 – erro sintático; 5 – erro semântico.

Resultados

Recorreu-se a provas de estatística descritiva para a análise dos dados obtidos. A média de acertos do total de alunos ($N=277$) foi de 5,13, com um desvio padrão de 3,24. Houve protocolos de sujeitos que fizeram zero ponto e outros com 14 pontos, dos 15 pontos permitidos pelo instrumento.

A Tabela 1 mostra as estatísticas descritivas dos tipos de erros cometidos no Cloze pelas crianças com médias altas, como descrito no método, no item da análise dos dados.

Tabela 1 – Estatísticas descritivas dos tipos de erros cometidos no Cloze pelas crianças que tiveram médias altas

Palavra correta	Tipos de Erro – Alunos com médias altas					Total
	Branco	Fonológico	Lexical	Sintático	Semântico	
uma				11		11
estavam				5		5
casa	1			1		2
ficou				1		1
ela				15	3	18
rasgar				1		1
estava	2	1		10	2	15
para	1		6	5	1	13
as				1		1
nossas	1			6		7
muito			2	14		16
seu				5		5
o	1					1
outro	2			20	1	23

seguintes grupos de protocolos: com até quatro pontos, que somaram 125; de cinco a sete pontos, que somaram 81; com oito pontos ou mais, que somaram 71 protocolos. Em seguida, os testes de Cloze das crianças que tiveram as melhores e as piores pontuações foram utilizados para a construção de uma escala dos tipos de erros cometidos. Para tanto, foram sorteados pelo programa SPSS 40% de protocolos das crianças que obtiveram as maiores notas ($n=71$). Pelo sorteio foram selecionados 28 protocolos, no entanto dois foram eliminados, o do participante que obteve 14 pontos e um que continha sete lacunas em branco. Assim, chegou-se ao número de 26 protocolos com pontuações de oito a 13 pontos. Para o sorteio dos 26 protocolos com as pontuações mais baixas (30% dos 86 restantes), primeiramente foram eliminados os protocolos das crianças que não tiveram qualquer acerto ou apenas um acerto ($n=39$). A classificação dos tipos de erros, dos menos aos mais graves, foi baseada nas divisões de Gar-

Os erros sintáticos foram os mais cometidos ($n=99$) para as crianças que tiveram as maiores médias, seguidos pelos brancos ($n=14$) e pelos semânticos ($n=11$). O erro fonológico foi cometido somente uma vez, seguido dos erros lexicais ($n=8$). Todas as crianças tiveram algum tipo de erro, sendo que houve apenas um erro para as palavras: “ficou”, “rasgar”, “as” e “o”. A palavra com maior quantidade de erros foi “outro”.

Para avaliar se havia homogeneidade na distribuição dos tipos de erros cometidos por essas crianças, foi realizada a Prova de Qui-quadrado. Essas análises são apresentadas na Tabela 2.

Tabela 2 – Prova de Quiquadrado para avaliar a homogeneidade na distribuição dos tipos de erros cometidos no Cloze pelas crianças com médias altas

Tipos de erros	F	fo(%)	fe	(F-fe) ² /fe
Branco	14	10,53	26,6	5,97
Fonológico	1	0,75	26,6	24,64
Lexical	8	6,01	26,6	13,00
Sintático	99	74,43	26,6	197,05
Semântico	11	8,27	26,6	9,14
Total	133	100,00	133,00	249,82

O índice χ^2 foi de 249,82 e o grau de liberdade para a prova foi de 4. Já o índice de significância foi estatisticamente significativo ($p < 0,001$), em decorrência de os resultados não evidenciarem homogeneidade na distribuição. Os valores obtidos para as categorias foram de 10,53 para o branco; 0,75 para o erro fonológico; 6,01 para o lexical; 74,43 para o sintático e 8,27 para o semântico. Na Tabela 3, estão demonstradas as estatísticas descritivas dos tipos de erros cometidos no Cloze pelas crianças com médias baixas.

Finalmente, verificou-se que houve erros para todas as palavras. Prosseguindo essa análise, foi efetuada a Prova de Quiquadrado para avaliar a homogeneidade na distribuição dos tipos de erros por categoria, que é mostrada na Tabela 4.

O índice χ^2 foi de 257,38, sendo que o grau de liberdade foi 4. O índice de significância foi de ($p < 0,001$), indicando que esse resultado é estatisticamente significativo e demonstrando que a distribuição não é equitativa.

Tabela 3 – Estatísticas descritivas dos tipos de erros cometidos no Cloze pelas crianças com médias baixas

Palavra correta	Tipos de Erro – Alunos com médias baixas					Total
	Branco	Fonológico	Lexical	Sintático	Semântico	
uma				18	2	20
estavam				10	6	16
casa	2			5	5	12
ficou		3		3	3	9
ela	1			15	6	22
rasgar	1	1		3	15	20
estava	2			11	10	23
para	1		3	4	11	19
as				11	11	22
nossas				17	4	21
muito	2	4		1	14	21
seu	1	1		15	1	18
o	4			3	2	9
outro	3			6	12	21
porta	6				14	20

Verifica-se, na Tabela 3, que os erros sintáticos foram os mais cometidos (122), seguidos dos erros semânticos ($n=116$). As crianças erraram mais a palavra “estava”, no entanto outras palavras obtiveram total de erros muito próximos, como 22 para “ela”, 21 para “nossa”, “muito” e “outro” e 20 para “uma” e “porta”. As palavras que tiveram menor número de erros foram “ficou” ($n=9$) e “o” ($n=9$).

Discussão

O objetivo desta pesquisa foi o de explorar as diferenças qualitativas nos erros apresentados em um Teste de Cloze nas pontuações extremas de crianças e encontrar evidências de validade por processo de resposta. Cabe lembrar que o teste utilizado já apresenta vários estudos com evidências de validade, a saber: entre construtos relacionados; de critério, quanto ao ano escolar; pelos gru-

Tabela 4 – Prova de Quiquadrado para avaliar a homogeneidade na distribuição dos tipos de erros cometidos no Cloze pelas crianças com médias baixas

Tipos de erros	<i>F</i>	<i>f</i> ₀ (%)	<i>f</i> _e	$(F-f_e)^2/f_e$
Branco	23	8,42	54,6	18,29
Fonológico	9	3,30	54,6	38,08
Lexical	3	1,10	54,6	48,76
Sintático	122	44,69	54,6	83,20
Semântico	116	42,49	54,6	69,05
Total	273	100	54,6	257,38

pos contrastantes em relação aos escores de instrumentos de medida de consciência fonológica, como demonstrado em vários estudos levantados neste trabalho.

No entanto, este estudo buscou acentuar o estudo dos processos cognitivos envolvidos na resolução dos itens, decompondo a dificuldade dos itens em parâmetros representativos de seus componentes (OLIDEN, 2003; PRIETO; DELGADO, 1999). Além disso, quis acrescentar outra técnica disponível e pouco utilizada às evidências já encontradas anteriormente (ANASTASI, 1961/1977; ANASTASI; URBINA, 2000).

De acordo com os resultados, pode-se observar que não houve homogeneidade na distribuição dos erros, tanto para os estudantes com pontuações mais altas quanto para aqueles com as mais baixas, pois os índices de significância foram estatisticamente significativos na prova de quiquadrado. Assim, pode-se observar que, para os alunos com médias mais altas, os erros sintáticos ($n=99$) foram os mais cometidos, seguidos pelos brancos e pelos semânticos ($n=11$). Este resultado indica que essas crianças tiveram mais problemas com a recuperação da palavra exata omitida, visto que o módulo sintático inclui os processos que contêm as estratégias de reconhecimento gramatical, como a regência, a colocação e a concordância nominal e verbal. Os outros tipos de erros foram pouco cometidos. Tal forma de resposta está congruente com as demandas do processo cognitivo envolvido (NUNES; PRIMI, 2010; PADILLA; BENÍTEZ, 2014).

Na pesquisa de Cunha e Santos (2009) com um teste de Cloze com características semelhantes, as crianças com médias mais altas cometeram mais erros lexicais. Considerando que os erros foram classificados por ordem crescente de gravidade, de acordo com as divisões de García (1988), na sequência branco, fonológico, lexical, sintático e semântico, seria esperado que os alunos com pontuações mais altas tivessem cometido também mais erros lexicais, que incidem no uso do sinônimo.

No entanto, os erros sintáticos foram os mais cometidos, indicando alguma compreensão, mas apresentando problemas em relação às estratégias de reconhecimento gramatical, como a consideração da ordem das palavras, do papel das palavras funcionais, do significado das palavras e do uso dos sinais de pontuação. Esse fato pode ser atribuído a fatores intrínsecos ao texto que interferem na sua compreensão, como a extensão da palavra a ser encontrada, o número de respostas possíveis para a lacuna e o número de formas possíveis a serem consideradas (ABRAHAM; CHAPELLE, 1992). Também podem ser

considerados: o método adotado na criação das lacunas, que pode ter dificultado o uso do contexto; a aplicação de conhecimento prévio ou o processamento local imediato ou externo ao texto (SANTOS et al., 2002).

Para os estudantes com médias mais baixas, os erros sintáticos ($n=122$) foram seguidos de perto pelos erros semânticos ($n=116$). Na pesquisa de Cunha e Santos (2009), os resultados mostraram que os participantes com médias mais baixas cometeram os erros semânticos em maior quantidade. Essa diferença entre os resultados das pesquisas pode ser explicada por características do texto que podem ter dificultado a localização da pista do contexto, a recuperação de palavras de conteúdo, ou pode ter havido um número grande de respostas possíveis para a lacuna (ABRAHAM; CHAPELLE, 1992). Outra explicação possível diz respeito ao método adotado na criação das lacunas do texto, que pode ter levado os respondentes a depender mais do uso do contexto ou da aplicação de conhecimento prévio (SANTOS et al., 2002).

As crianças com pontuações mais baixas erraram mais a palavra “estava”. No entanto, outras palavras obtiveram total de erros muito próximo. As palavras que tiveram menor número de erros para esses alunos foram “ficou” ($n=9$) e “o” ($n=9$). Provavelmente, as pistas do contexto ajudaram a recuperar essas palavras, pois os alunos com pontuação alta tiveram apenas um erro em cada uma delas.

Finalmente, verifica-se que houve erros para todas as palavras.

Considerações finais

As evidências de validade baseadas no processo de resposta levaram a uma análise qualitativa do teste de Cloze, que favoreceu a retomada do construto, dos efeitos do teste e principalmente dos processos cognitivos envolvidos na recuperação das palavras pelas crianças. Ficaram também evidentes as divergências entre alunos que pontuaram menos – em princípio aqueles que têm pior nível de compreensão – e os com maiores pontuações, ou seja, aqueles com bom nível de compreensão leitora.

Neste estudo, de acordo com o modelo cognitivo sugerido para a leitura, puderam ser evidenciadas as discrepâncias entre grupos extremos e, com a comparação com outro teste similar, a diferença nos resultados de acordo com o texto escolhido para a elaboração do teste. Nesse sentido, foram alcançados os parâmetros estabelecidos de que as evidências empíricas, embasadas na teoria, se estabeleceram nos avaliados, nas tarefas propostas.

Um resultado inesperado foi encontrado com a comparação do processo de resposta envolvido em teste de Cloze similar, também amplamente usado para mensuração do nível de compreensão de leitura. Esse fato reafirma a importância do estudo dos processos cognitivos envolvidos na resolução dos itens de um teste, com uma análise qualitativa dos protocolos de resposta.

Estudos futuros poderiam investigar se haveria diferença nos resultados se fossem incluídas crianças de escolas públicas e particulares. Outra possibilidade seria pesquisar os padrões de erros entre os anos escolares para observar se eles iriam na direção dos processos cognitivos determinados neste estudo.

Referências

ABRAHAM, R. G.; CHAPPELLE, C. A. The meaning of Cloze test scores: an item difficulty perspective. *The Modern Language Journal*, Iowa City, v. 76, n. 4, p. 468-479, winter, 1992.

AMERICAN PSYCHOLOGICAL ASSOCIATION; AMERICAN EDUCATIONAL RESEARCH ASSOCIATION; NATIONAL COUNCIL ON MEASUREMENT IN EDUCATION. *Standards for educational and psychological testing*. Washington, DC: APA, 1999.

ANASTASI, A. *Testes psicológicos* (1961). Tradução de Dante Moreira Leite. 2. ed. São Paulo: EPU, 1977.

ANASTASI, A.; URBINA, S. *Testagem Psicológica*. Tradução de Maria Adriana Verissimo Veronese. 7. ed. Porto Alegre: Artes Médicas, 2000.

COHEN, J. H. The effect of content area material on Cloze test performance. *Journal of Reading*, Newark, v. 19, n. 3, p. 247-250, 1975.

CONSELHO FEDERAL DE PSICOLOGIA. *Resolução nº 002/2003*. Brasília: CFP, 2003. Disponível em: <https://site.cfp.org.br/wp-content/uploads/2012/05/resoluxo022003.pdf>. Acesso em: 23 jul. 2015.

CUNHA, N. B.; SANTOS, A. A. A. Validade por processo de resposta no teste de Cloze. *Fractal: Revista de Psicologia*, Niterói, v. 21, n. 3, p. 549-562, 2009. [CrossRef](#)

CUNHA, N. B.; SANTOS, A. A. A. Estudo de validade do questionário de avaliação de consciência metatextual. *Psicologia: teoria e prática* [online]. São Paulo, v. 16, n. 1, p. 141-154, 2014. Disponível em: http://pepsic.bvsalud.org/scielo.php?script=sci_arttext&pid=S1516-36872014000100012&lng=pt&nrm=iso. Acesso em: 23 jul. 2015.

GARCÍA, J. N. *Manual de dificuldades de aprendizagem: linguagem, leitura, escrita e matemática*. Tradução de Jussara Haubert Rodrigues. Porto Alegre: Artmed, 1998.

MACÍAS, A. B. La consulta a expertos como estrategia para la recolección de evidencias de validez basadas en el contenido. *Investigación Educativa*, Durango, v. 7, n. 2, p. 5-14, jul./dez. 2007.

MESSICK, S. Validity of psychological assessment. Validation of inferences from person's responses and performances as scientific inquiry into score meaning. *American Psychologist*, v. 50, n. 9, p. 741-749, 1995. [CrossRef](#)

MOTA, M. M. P. E.; SANTOS, A. A. A. O Cloze como instrumento de avaliação de leitura nas séries iniciais. *Psicologia Escolar e Educacional* [online], Maringá, v. 18, n. 1, p. 135-142, 2014. [CrossRef](#)

NUNES, C. H. S. S.; PRIMI, R. Avaliação Psicológica: diretrizes na regulamentação da profissão. In: SANTOS, A. A. A. et al. (Org.). *Aspectos técnicos e conceituais da ficha de avaliação dos testes psicológicos*. Brasília: Conselho Federal de Psicologia, 2010. p. 101-128.

OLIDEN, P. E. Sobre la validez de los tests. *Psicothema*, Oviedo, v. 15, n. 2, p. 315-321, maio/jul. 2003. Disponível em: <http://www.psicothema.com/psicothema.asp?id=1063>. Acesso em: 12 out. 2015.

PADILLA, J. L.; BENÍTEZ, I. Validity evidence based on response process. *Psicothema*, Oviedo, v. 26, n. 1, p. 136-144, 2014. Disponível em: <http://www.psicothema.com/pdf/4171.pdf>. Acesso em: 27 Sept. 2016.

PAGE, W. D. The post-oral reading Cloze test: new link between oral reading and comprehension. *Journal of Reading Behavior*, Chicago, v. 7, n. 4, p. 383-389, 1975. Disponível em: <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.860.7161&rep=rep1&type=pdf>. Acesso em: 11 jan. 2013.

PRIETO, G.; DELGADO, A. R. Medición cognitiva de las aptitudes. In: OLEA, J.; PONSODA, V.; PRIETO, G. (Ed.). *Tests informatizados: fundamentos y aplicaciones*. Madrid: Pirâmide, 1999. p. 207-226.

PRIMI, R. *Inteligência, processamento de informação e teoria gestalt: um estudo experimental*. 1995. Dissertação (Mestrado em Psicologia)-Departamento de Pós-graduação em Psicologia da Pontifícia Universidade Católica de Campinas, Campinas. 1995.

PRIMI, R. *Desenvolvimento de um instrumento informatizado para a avaliação do raciocínio analítico*. 1998. Tese (Doutorado em Psicologia)-Instituto de Psicologia da Universidade de São Paulo, São Paulo, 1998.

RUEDA, F. J. M.; SISTO, F. Versão preliminar do teste pictórico de memória: estudo de validade. *Estudos de psicologia*, Campinas, v. 25, n. 2, p. 223-231, 2008. [CrossRef](#)

SANTOS, A. A. A. *Desenvolvimento do hábito de leitura e compreensão de textos através da aplicação de fichas: um estudo com adolescentes carentes*. 1981. Dissertação (Mestrado)-Pontifícia Universidade de Campinas, Campinas, 1981.

SANTOS, A. A. A. *Evidências de validade de critério para o Teste de Cloze*. 2005. Trabalho de Conclusão de Curso (Especialização). Programa de Pós-graduação Stricto-Sensu em Psicologia, Universidade São Francisco, São Paulo, 2005.

SANTOS, A. A. A. et al. O teste de Cloze na avaliação da compreensão em leitura. *Psicologia: Reflexão e Crítica*, Porto Alegre, v. 15, n. 3, p. 549-560, set./dez. 2002. [CrossRef](#)

SANTOS, A. A. A.; CUNHA, N. B. Consciência metatextual: evidências de validade para instrumento de medida. *Psico-USF* [online], v. 17, n. 2, p. 233-241, 2012. [CrossRef](#)

SANTOS, A. A. A.; SISTO, F. F.; NORONHA, A. P. P. TONI 3 - Forma A e Teste de Cloze: evidências de validade. *Psicologia: Teoria e Pesquisa*, v. 26, n. 3, p. 399-405, 2010.

SUEHIRO, A. C. B.; MAGALHAES, M. M. S. Relação entre medidas de avaliação da linguagem escrita em estudantes do ensino fundamental. *Psico-USF* [online], Itatiba, v. 19, n. 3, p. 489-498, 2014. [CrossRef](#)

SUEHIRO, A. C. B.; SANTOS, A. A. A. Compreensão de leitura e consciência fonológica: evidências de validade de suas medidas. *Estudos de psicologia (Campinas)* [online], Campinas, v. 32, n. 2, p. 201-211, 2015. [CrossRef](#)

Recebido em: 24 de março de 2017

Aceito em: 4 de setembro de 2018