

VALIDADE POR PROCESSO DE RESPOSTA NO TESTE DE CLOZE

*Neide de Brito Cunha**

*Acácia Aparecida Angeli dos Santos***

RESUMO

Tendo por base considerações sobre as evidências de validade por processo de resposta, foram analisados neste estudo 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: das crianças que tiveram 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. Embora tenha sido encontrada a evidência de validade por processo de resposta, por meio da avaliação da homogeneidade na distribuição dos tipos de erros, é preciso ampliar o conhecimento sobre as características psicométricas do teste de Cloze.

Palavras-chave: avaliação psicológica; Cloze; compreensão de leitura; psicometria; processo de resposta.

RESPONSE PROCESS VALIDITY IN THE CLOZE TEST

ABSTRACT

Based in the response process validity evidence, this study explored quality of children's mistakes on the Cloze test. The sample was composed by 266 children, boys and girls, aging from 8 to 13 years old, attending at the third and fourth grades of the elementary education of public and private schools of São Paulo, Brazil. Two protocols groups were formed: of the children with high and low means to construct a errors scale, for the least to most serious: blank, phonological, lexical, syntactic and semantic. The results showed that children with high means test presented lexical mistakes, while children with low means made semantic mistakes. Although the response process validity evidence was finding, by the

*Doutora em Psicologia, na área de Avaliação Psicológica Educacional, pela Universidade São Francisco - Itatiba. Professora adjunta doutora da Universidade São Francisco.

E-mail: neidedavilla@uol.com.br

**Professora titular do curso de graduação e do Programa de Pós-Graduação Stricto Sensu em Psicologia da Universidade São Francisco. Endereço: Universidade São Francisco - Itatiba, Faculdade de Psicologia, Programa de Pós Graduação Stricto Sensu em Psicologia. Rua Alexandre Rodrigues Barbosa, 45 – Centro - Itatiba, SP – Brasil. CEP: 13251-900.

E-mail: acacia.angeli@gmail.com.br

homogeneously evaluation on the errors distribution types, it's necessary to extend the knowledge of psychometric characteristics of the Cloze test.

Keywords: psychological assessment; Cloze test; reading comprehension; psychometrics; response process.

INTRODUÇÃO

No processo de construção de um teste, uma escala, um inventário ou um questionário, a validade é um dos temas cruciais, visto que busca demonstrar a utilidade da medida realizada, ou seja, o significado das pontuações obtidas. É precisamente a validade que permitirá realizar inferências e interpretações corretas dos resultados ao aplicar um teste e estabelecer a relação com o construto/variação que se mede (ANASTASI; URBINA, 2000; PASQUALI, 2003; MUÑIZ, 2004; MACÍAS, 2007).

Messick (1980) cita que, ao longo dos anos, foram muitas as acepções atribuídas à validade, a saber: convergente, discriminante, fatorial, substantiva, estrutural, externa, de população, ecológica, temporal, de tarefa, sob as quais é possível delimitar três etapas que ficaram impressas nos padrões de 1974, 1985 e 1999 (AMERICAN PSYCHOLOGICAL ASSOCIATION; AMERICAN EDUCATIONAL RESEARCH ASSOCIATION; NATIONAL COUNCIL ON MEASUREMENT IN EDUCATION, 1966, 1974, 1985, 1999). Para Oliden (2003), a evolução de seu significado desde uma origem pragmática e operacional, até a complexidade da visão que hoje impera, reflete o caráter progressivo da ciência que a envolve. A autora resume da seguinte maneira as três fases:

1ª: etapa operacional dominada por uma visão pragmática que destaca a validade externa. Essa perspectiva diferencia quatro tipos de validade: conteúdo, preditiva, concorrente e de construto (AMERICAN PSYCHOLOGICAL ASSOCIATION; AMERICAN EDUCATIONAL RESEARCH ASSOCIATION; NATIONAL COUNCIL ON MEASUREMENT IN EDUCATION, 1954), que nas edições de 1966 e 1974 foram reduzidas para três, agrupando a validade preditiva e a concorrente na validade de critério;

2ª: estágio teórico marcado pela importância concedida à teoria psicológica, no qual se adota uma visão integradora. Impôs-se a análise da validade de construto como um conceito unificador que abarcava aspectos de conteúdo e de relações com outras variáveis. Supôs-se o reconhecimento da validade como processo único de coleta de evidências por meio de estratégias de investigação diferentes, relacionadas com o construto, com o conteúdo e com o critério;

3ª: fase atual, a qual se pode denominar contextual, na qual se amplia a aceção anterior e se delimita o conceito de “uso proposto”. Seu objetivo seria dotar os testes de avais tanto científicos como éticos. Nessa nova revisão não se encontram referências a diferentes formas de validade e se incorpora

à conotação teórica um aspecto até então esquecido, o uso. Assim, já não se considera suficiente somente a justificativa das pontuações, é necessário também lastrear os fundamentos teóricos no contexto externo, em relação ao “propósito” ou à “interpretação proposta”.

O objetivo dos estudos de validade, sob essa perspectiva atual, é coletar evidências suficientes que possam prestar uma base científica à interpretação das pontuações em uso concreto, que podem provir de diversas fontes. A importância outorgada a cada uma delas dependerá dos objetivos do teste. Os últimos estândares diferenciam entre fontes relacionadas com o conteúdo, o processo de respostas, a estrutura interna, as relações com outras variáveis e as consequências do teste. As fontes são reunidas em dois grupos, sendo as de evidência interna aquelas que supõem uma análise individualizada dos itens e externa quando analisam o teste em conjunto (OLIDEN, 2003).

Considerando as fontes de evidência interna, para o propósito deste estudo será explorada a evidência de validade baseada no processo de resposta. Os métodos considerados para investigá-la implicam análises de protocolos, entrevistas e também o uso de modelos componenciais. Algumas considerações sobre esses métodos são descritas a seguir.

Segundo Prieto e Delgado (1999), a influência exercida pela psicologia cognitiva sobre a psicometria tradicional levou a uma re-análise da medição por meio de testes, que deu maior importância à utilidade do construto na sua representação. Na busca de instrumentos de medida que se ajustassem a um marco que não fosse estritamente estatístico, o estudo dos processos cognitivos envolvidos na resolução dos itens passou a ser uma importante fonte de informação.

A metodologia inclui predominantemente a análise de protocolos de resposta, entrevistas ou outros procedimentos que permitam a análise individualizada do par sujeito/item. Para Oliden (2003) 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.

Em acréscimo, Macías (2007) afirma que a evidência baseada no processo de resposta reúne argumentos que evidenciam as consistências entre as respostas ao instrumento (evidência empírica) e os processos que, desde a teoria, se estabelece que assumem os avaliados nas tarefas propostas. Os argumentos relacionados com esse aspecto se encontram no marco conceptual; assim, os processos seguidos pelos avaliados para chegar às respostas que são formuladas devem ser explicitados.

Tendo por base essas considerações sobre as evidências de validade por processo de resposta, serão analisadas neste estudo as respostas a um teste de Cloze. É importante lembrar que o Cloze é uma técnica que vem sendo utilizada por pesquisadores da área de linguagem, principalmente a partir da década de

1970 e foi originalmente proposto por Taylor (1953), como um meio de avaliar a compreensibilidade de textos escritos. Como procedimento de ensino, o Cloze é muito versátil em formatos e conteúdos, permitindo ao leitor tomar consciência da interatividade entre seu eu e o texto e de sua participação ativa na reconstrução do significado pretendido pelo autor.

Esse procedimento consiste, em sua versão original, na apresentação de uma seleção de prosa da qual se omitem todos os quintos vocábulos, substituindo-os por um traço, proporcional ao tamanho da palavra omitida, que deve ser recuperada pelo leitor para restituir o sentido completo à seleção. Tendo como pressuposto a inter-relação entre a experiência do mundo e da linguagem que o leitor tem armazenadas na sua memória, com as respostas semânticas e sintáticas que fornece o contexto para completar o significado, as palavras omitidas passam a constituir espaços de observação direta do processo da compreensão, úteis tanto para o educador como para o pesquisador das ciências da linguagem (CONDEMARÍN; MILICIC, 1994).

A validade de construto do Cloze, desde uma perspectiva linguística, tem como ponto de partida o fato de que em cada língua, em razão de sua estrutura, existem contingências estatísticas que determinam a probabilidade de que certas palavras tendem a seguir outras. Por exemplo, em inglês, o adjetivo deve preceder o substantivo e em espanhol e em português o artigo deve preceder o substantivo. Além disso, o emprego das palavras está condicionado parcialmente pelo tema. Assim, se o conteúdo se refere à leitura, é provável que num texto as palavras “compreensão e legibilidade” se encontrem em quantidade maior que as palavras “porcas e parafusos”, que serão mais prováveis num texto sobre mecânica. A frequência do uso das palavras também é influenciada por certos modismos culturais que funcionam como expressões ou ditos, como, por exemplo, “é duro como o aço”. Assim, um indivíduo familiarizado com sua língua pode predizer as palavras que seguem a outras (CONDEMARÍN; MILICIC 1994).

Estudos com o instrumento indicam a existência de diferenças de desempenho no teste de Cloze relacionadas com o assunto abordado. Isso levou a crer que a familiaridade com o assunto aumenta a probabilidade de acerto das respostas dadas, introduzindo uma variável interveniente na comparação dos desempenhos (COHEN, 1975; PAGE, 1975; SANTOS, 1981). Já Abraham e Chappelle (1992) pesquisaram os fatores intrínsecos ao texto, que interferem na sua compreensão e obtiveram no seu trabalho as seguintes predições de dificuldade do item: 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.

Na mesma direção dos resultados de Abraham e Chappelle (1992), a pesquisa de Santos et al (2002) indicou que o tipo de processamento requerido e a habilidade avaliada pelo teste de Cloze dependem basicamente do método adotado na criação das lacunas. Desse modo, o acerto em determi-

nado tipo de lacuna pode depender do uso do contexto, reforçando a posição dos pesquisadores que salientam o processamento global contextual como elemento central na elaboração das respostas, enquanto o acerto de outro tipo de lacuna aparece mais 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.

A pesquisa de Santos et al (2002) também verificou, assim como a anterior de Santos (1990), que os adjetivos, verbos e advérbios estão entre as lacunas mais difíceis de serem completadas por serem classificadas como palavras de conteúdo, relativas à função coesiva do discurso. Em relação ao processamento cognitivo, as palavras de conteúdo requerem mais o processamento semântico, com a recuperação das palavras da memória de longo-prazo, em oposição ao processamento sintático, que requer a recuperação das palavras a partir da estrutura sintática das frases, nas quais a palavra omitida se insere. As preposições e os artigos foram as lacunas mais fáceis de serem preenchidas, pois formam unidades com as palavras imediatamente adjacentes e são recuperadas mais facilmente, o que confirmou uma maior facilidade de processamento do tipo *bottom-up* (KATO, 1987; BITAR, 1989). Os substantivos apresentaram dificuldade média, mas houve uma grande dispersão dos índices de dificuldade.

Com as contribuições da psicologia cognitiva e da teoria do processamento da informação, também da neuropsicologia cognitiva e da neurolinguística, surgiu o modelo da psicologia da leitura (GARCÍA, 1998). 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. A descrição breve de cada um deles é apresentada a seguir de modo sucinto.

O módulo perceptivo contém os processos de extração de informação, que têm a ver com a memória icônica e processos de memória de trabalho nos quais se efetuam tarefas de reconhecimento e análise linguística. Já o módulo léxico consiste na recuperação do conceito associado à unidade linguística, ou recuperação léxica, mediante dois caminhos, a saber, um caminho direto, também conhecido como visual ou ortográfico que permite a conexão do significado com os sinais gráficos, por intervenção da memória global das palavras. O outro caminho, indireto ou fonológico, recupera a palavra por meio da aplicação das regras de transformação de grafema a fonema, o que leva à descoberta do significado.

O terceiro módulo, denominado de 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. A compreensão ocorre à medida que o leitor vai realizando as relações entre as proposições nos dois níveis (KINTSCH; VAN DIJK, 1978; GARCÍA, 1998).

O resgate dos estudos aqui apresentados fornece elementos que possibilitam analisar as respostas às lacunas em testes de Cloze compatíveis com o que é conceitualmente esperado para investigar a validade por processo de resposta. Nesse sentido, esta pesquisa tem como objetivo explorar as diferenças qualitativas nos erros apresentados num teste de Cloze em crianças com pontuações extremas e encontrar evidências de validade por processo de resposta.

MÉTODO

Participantes

Participaram deste estudo 266 crianças, de ambos os sexos, sendo 140 (52,6%) do sexo masculino e 126 (47,4%) do feminino. Do total da amostra, 118 (44,4%) frequentavam a 3ª série e 148 (55,6%) a 4ª, sendo 66 (24,8%) de escolas particulares, 189 (71,1%) de públicas, quatro do Sesi (1,5%) e sete não informaram. As idades variaram de 8 a 13 anos ($M=9,42$; $DP=0,808$). Cabe salientar que todas as escolas eram do interior do estado de São Paulo.

Instrumento

Teste de Cloze (Santos, 2005) – foi utilizado um texto estruturado na forma do Cloze tradicional com os quintos vocábulos omitidos, substituindo-os por um traço proporcional ao tamanho da palavra omitida, que deveria ser recuperada pelo leitor para restituir o sentido completo à seleção. Esse texto foi especialmente montado para ser utilizado com crianças da faixa etária do ensino fundamental e também apresenta evidências de validade de critério. Para tanto, foi submetido a um estudo piloto com 314 alunos de ambos os sexos, de 2ª a 4ª 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 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 CFP (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, para a terceira série 0,70 e para a quarta série 0,72.

Cr terios de Corre o

Foi atribuído um ponto para cada acerto, sendo consideradas corretas as palavras que tiverem sido as mesmas usadas pelo autor do texto. Essa forma de corre o   denominada literal ou verbatim e   recomendada por evitar o subjetivismo na avalia o. Assim, s  pode haver 15 acertos , correspondentes ao n mero de omiss es.

Procedimento

A presente pesquisa foi analisada e aprovada pelo Comit  de  tica da Universidade S o Francisco. A coleta de dados ocorreu em institui es cujas diretoras autorizaram sua realiza o e com aqueles alunos cujos pais assinaram o Termo de Consentimento Esclarecido, permitindo a participa o do(a) filho(a). A aplica o ocorreu em situa o de sala de aula, pela professora, num dia normal de atividades escolares. Tamb m foi fornecido um roteiro  s professoras com instru es para a aplica o.

RESULTADOS

Tendo em vista os objetivos pretendidos, recorreu-se a provas de estat tica descritiva para se analisar os dados obtidos no presente estudo. A m dia de acertos do total de alunos (N= 266) foi de 9,91, com um desvio padr o de 3,79. Houve sujeitos que obtiveram os n meros m nimo (n=7) e m ximo (n=29) de pontos permitidos pelo instrumento.

Foram formados os seguintes grupos de protocolos: com at  6 pontos, que somaram 46; de 7 a 13 pontos, que somaram 167; com 14 pontos ou mais, que somaram 53 protocolos. Em seguida, os testes de Cloze das crian as que tiveram as melhores e as piores pontua es foram utilizados para a constru o de uma escala dos tipos de erros cometidos. Para tanto, foram eliminados os protocolos das crian as que n o tiveram qualquer acerto, os que estavam ileg veis e os com 14 acertos, porque neles, sistematicamente, os erros aconteceram com a palavra “do”. Tamb m foram eliminados os protocolos com 15 acertos. Assim obteve-se o n mero de 22 protocolos da categoria “m dias altas”. Em seguida, foram sorteados 22 protocolos de alunos classificados na categoria “m dias baixas”.

A classifica o dos tipos de erros, dos menos aos mais graves, ficou assim estabelecida: 1 – branco; 2 – erro fonol gico (incluindo erros ortogr ficos e de acentua o); 3 – erro lexical (uso de sin nimo); 4 – erro sint tico e 5 – erro sem ntico. A Tabela 1 mostra as estat sticas descritivas dos tipos de erros cometidos no Cloze, pelas crian as com m dias altas.

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

Palavra correta	Tipo de Erro – Alunos com médias altas					Total
	Branco	Fonológico	Lexical	Sintático	Semântico	
um	0	0	0	2	0	2
lá	2	1	3	0	5	11
misterioso	0	0	16	0	1	17
ela	0	1	0	1	2	4
era	0	0	0	0	0	0
princesa	0	2	0	0	2	4
ficou	0	0	3	6	3	12
de	0	0	0	1	0	1
que	0	0	0	0	1	1
do	0	0	0	12	2	14
o	1	0	0	1	1	3
a	0	0	0	0	0	0
e	0	0	0	1	0	1
caixinha	0	0	4	0	3	7
amor	0	0	1	0	2	3
Total	3	4	27	24	22	80

Para esses alunos, os erros lexicais foram os mais cometidos (n= 27), seguidos pelos erros sintáticos (n=24) e pelos semânticos (n=22). Os erros fonológicos foram os menos cometidos, sendo que as palavras que mais proporcionaram erros lexicais foram “misterioso” (n=16), “do” (n=12), que levou aos erros sintáticos, e “lá” (n=5) aos semânticos. Nenhuma criança errou nas palavras “era” e “a”. Verifica-se que esses alunos praticamente não deixaram lacunas em branco (n=3).

Dando prosseguimento às análises, foi efetuada a Prova de Qui-quadrado para avaliar a homogeneidade na distribuição dos tipos de erros cometidos por essas crianças. Essas análises são apresentadas na Tabela 2.

Tabela 2 – Prova de Qui-quadrado 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	3	3,75	16,00	10,56
Fonológico	4	5,00	16,00	9,00
Lexical	27	33,75	16,00	7,56
Sintático	24	30,00	16,00	4,00
Semântico	22	27,50	16,00	2,25
TOTAL	80	100,00	80,00	33,38

O grau de liberdade para a prova foi de 4, o índice χ^2 foi de 33,38 e 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 10,56 para o branco; 9,00 para o erro fonológico; 7,56 para o lexical; 4,00 para o sintático e 2,25 para o semântico. A seguir, estão demonstradas as estatísticas descritivas dos tipos de erros cometidos no Cloze, pelas crianças com médias baixas, na Tabela 3.

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

Palavra correta	Tipo de Erro – Alunos com médias baixas					Total
	Branco	Fonológico	Lexical	Sintático	Semântico	
um	0	5	1	2	0	8
lá	2	1	0	0	11	14
misterioso	0	2	17	1	1	21
ela	0	3	1	2	8	14
era	1	0	0	3	4	8
princesa	0	7	1	2	2	12
ficou	1	0	0	6	10	17
de	1	0	0	7	3	11
que	4	1	0	0	3	8
do	2	0	0	16	1	19
o	2	0	0	8	4	14
a	0	0	0	3	2	5
e	3	0	1	5	2	11
caixinha	0	4	4	0	4	12
amor	1	0	0	0	6	7
Total	17	23	25	55	61	181

Os critérios adotados para classificação dos erros desses alunos coincidiram com a ordem crescente dos totais de erros cometidos. Assim, o total de erros fonológicos foi de 23, o de lexicais foi de 25, o de sintáticos foi de 55 e o de semânticos foi de 61.

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

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

Tipos de erros	F	fo (%)	fe	(F-fe) ² /fe
Branco	17	9,39	36,20	10,18
Fonológico	23	12,71	36,20	4,81
Lexical	25	13,81	36,20	3,47
Sintático	55	30,39	36,20	9,76
Semântico	61	33,70	36,20	16,99
TOTAL	181	100,00	181,00	45,22

O índice χ^2 foi de 45,22, 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.

DISCUSSÃO

A classificação dos tipos de erros, dos menos aos mais graves, exceto para o “branco”, teve por base as divisões da Linguística, que coincidem com as de García (1998), no modelo da psicologia da leitura. O erro fonológico significou que a criança conseguiu recuperar a mesma palavra utilizada pelo autor, mas com problemas de ortografia ou acentuação. Ficou, portanto, caracterizado o erro na grafia da palavra, mas o não comprometimento da compreensão do texto.

Do mesmo modo que o erro fonológico, o lexical indicou compreensão, pois houve recuperação do conceito associado à unidade linguística, mas com uso de um sinônimo ou palavra que não compromete o contexto. Assim, foi possível, por meio dos léxicos, acessar o sistema semântico, o que permitiu a compreensão da leitura (GARCÍA, 1998). Os erros sintáticos indicaram problemas com a consideração da ordem e do significado das palavras, com o papel das palavras funcionais e com o uso dos sinais de pontuação.

Para analisar a distribuição dos tipos de erros cometidos no Cloze pelas crianças com médias altas foi realizada a prova de qui-quadrado. Ficou demonstrado que não houve homogeneidade na distribuição, o que gerou um índice estatisticamente significativo ($p < 0,001$).

Os erros lexicais com a palavra “misterioso” foram os mais cometidos, indicando que a compreensão não foi prejudicada, pois essas crianças usaram sinônimos. Cabe analisar que essa palavra é a mais extensa a ser recobrada, com cinco sílabas, o que dificultaria sua recuperação, de acordo com as predições de dificuldade do item de (ABRAHAM; CHAPELLE, 1992).

Em seguida, apareceram os erros sintáticos, na palavra “do”, por desconsideração da contração da preposição “de” com o artigo “o”. Vale destacar que foi essa palavra que levou ao erro as crianças que alcançaram o escore de 14 pontos no Cloze, porque sistematicamente colocaram “de”. As preposições estão entre as palavras mais fáceis de serem completadas nos testes de Cloze, no entanto, no texto utilizado, o uso da contração “de” + “o” dificultou sua recuperação pelas crianças.

Depois vieram os erros semânticos, que ocorreram em maior número com a palavra “lá”, que por ser um advérbio é considerada uma palavra de conteúdo, portanto, de difícil preenchimento. Os erros fonológicos foram os menos cometidos e nenhuma dessas crianças errou nas palavras “era” e “a”, por serem palavras curtas e cujas pistas do contexto estavam muito próximas, respectivamente nas frases “o seu fantasma ‘era’ um príncipe” e “a princesa armou um plano” (SANTOS, 1990; ABRAHAM; CHAPELLE, 1992; SANTOS et al., 2002).

Observou-se que os erros dos alunos com médias altas aconteceram mais na recuperação da exata palavra utilizada pelo autor do texto. Para elas, os erros sintáticos e semânticos apareceram quase na mesma proporção. Elas quase não cometeram erros fonológicos e também quase não deixaram lacunas em branco. Infere-se que essas crianças conseguiram atribuir sentido à leitura, porque dispunham de recursos cognitivos suficientes para fazê-lo.

Considera-se que o leitor proficiente utiliza os dois métodos de processamento do conteúdo do material escrito, o *bottom-up*, no qual o processamento é dependente do texto, deixando de lado o sujeito e a situação de enunciação para construir o significado a partir da análise e síntese do significado das partes. O outro método é o ascendente, linear e indutivo, ou *top-down*, mais relacionado à área de psicologia cognitiva, depende do leitor, pois parte do conhecimento de mundo do indivíduo, para a decifração da palavra, de modo que o texto seria apenas um confirmador das hipóteses do leitor (KATO, 1987; BITAR, 1989).

Por fim, foram analisados os erros cometidos pelas crianças com médias baixas. Evidenciou-se, por meio da prova de qui-quadrado, uma distribuição heterogênea dos tipos de erros cometidos, que gerou um índice estatisticamente significativo ($p < 0,001$).

Os critérios adotados para classificação dos erros desses alunos coincidiram com a ordem crescente dos totais de erros cometidos, a saber: fonológicos, lexicais, sintáticos e semânticos. Esse resultado demonstra que os alunos com médias baixas cometeram mais erros semânticos. Assim sendo, infere-se que esses alunos têm dificuldades na compreensão de textos, pois os erros semânticos caracterizam dificuldades com os conhecimentos prévios, referentes ao significado declarativo e procedimental, para integrar as palavras num todo coerente. Há de se considerar que as estruturas do discurso são interpretadas como um jogo de proposições, ordenado por várias relações semânticas entre elas. As relações das proposições na microestrutura do texto requerem conhecimentos anteriores mais específicos, enquanto, na macroestrutura, a natureza das proposições é mais global. A compreensão ocorre à medida que o leitor vai realizando as relações entre as proposições nos dois níveis, conclui-se, então, que esse processo não aconteceu para essas crianças (KINTSCH; VAN DIJK, 1978; GARCÍA, 1998).

A palavra “princesa” foi a que mais proporcionou erros fonológicos, encontrando-se, principalmente, erros de troca de “s” por “z”, o que comprova a complexidade na representação do grafema “s”, visto que é um símbolo que representa diferentes fones. Dentre os erros lexicais, a palavra “misterio-

so” várias vezes foi substituída por “homem”, não desfazendo a combinação substantivo-adjetivo, pois depois desta palavra estava “estrangeiro”, considerando-se uma análise morfossintática.

Essas palavras, que proporcionaram erros fonológicos e lexicais, são respectivamente substantivo e adjetivo. Há de lembrar que pesquisas apontaram que os substantivos apresentam dificuldade média no preenchimento de lacunas nos testes de Cloze, enquanto os adjetivos, os verbos e os advérbios são os mais difíceis de serem completados, por serem classificados como palavras de conteúdo (SANTOS, 1990; ABRAHAM; CHAPELLE, 1992; SANTOS et al, 2002).

A palavra “do” levou aos erros sintáticos, assim como aconteceu com as crianças com médias altas. Quanto aos erros semânticos, foram mais erradas as palavras “lá” e “ficou”, que são respectivamente advérbio e verbo e são classificadas como palavras de conteúdo, portanto, de difícil preenchimento (SANTOS, 1990; SANTOS et al, 2002).

Cabe aqui retomar as considerações de Abraham e Chappelle (1992), de que há fatores intrínsecos ao texto que interferem na sua compreensão, assim, a recuperação de palavras de conteúdo, o local da pista do contexto, o número de sílabas no período, a extensão da palavra a ser encontrada, o número de respostas possíveis para a lacuna e o número de formas possíveis a serem consideradas podem predizer a dificuldade do item. Esses aspectos merecem ser mais bem explorados em pesquisas que utilizem o teste de Cloze.

CONSIDERAÇÕES FINAIS

É importante ressaltar que há fatores intrínsecos ao texto que interferem na sua compreensão. Assim, a recuperação de palavras de conteúdo, o local da pista do contexto, o número de sílabas no período, a extensão da palavra a ser encontrada, o número de respostas possíveis para a lacuna e o número de formas possíveis a serem consideradas podem predizer a dificuldade do item. Esses aspectos merecem ser mais bem explorados em pesquisas que utilizem o teste de Cloze.

Percebe-se, nesta análise qualitativa dos erros no Cloze, que houve a necessidade de recorrer à Psicologia e à Linguística para a compreensão de determinados mecanismos presentes na produção discursiva escrita. Assim sendo, sugere-se que mais estudos sejam realizados com essa possibilidade de interação entre as áreas.

Embora tenha sido encontrada a evidência de validade por processo de resposta, por meio da avaliação da homogeneidade na distribuição dos tipos de erros por categoria, é preciso que mais evidências de validade sejam pesquisadas para propiciar a ampliação do conhecimento sobre as características psicométricas do teste de Cloze.

Espera-se que a realização de outras pesquisas supere as limitações deste estudo, principalmente no que se refere ao tamanho da amostra. Além disso, elas poderão contribuir para a ampliação dos conhecimentos na área e para aprimorar os procedimentos técnicos empregados na avaliação psicológica. Deve haver

incremento na construção e a validação dos instrumentos psicológicos utilizados no Brasil, para tentar solucionar o problema da precariedade deles e de seu suprimento com a utilização de instrumentos estrangeiros.

REFERÊNCIAS

ABRAHAM, R. G.; CHAPELLE, 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.

ANASTASI, A.; URBINA, S. *Testagem Psicológica*. Tradução M. A. V. Veronese. Porto Alegre: Artes Médicas, 2000.

AMERICAN PSYCHOLOGICAL ASSOCIATION, AMERICAN EDUCATIONAL RESEARCH ASSOCIATION AND NATIONAL COUNCIL ON MEASUREMENT IN EDUCATION. Technical recommendations for psychological tests and diagnostic techniques. *Psychological bulletin*, Washington, DC, v. 51, n. 2, Part 2, jul./dez. 1954.

AMERICAN PSYCHOLOGICAL ASSOCIATION; AMERICAN EDUCATIONAL RESEARCH ASSOCIATION; NATIONAL COUNCIL ON MEASUREMENT IN EDUCATION. *Standards for educational and psychological Testing*. Washington, DC: APA, 1966, 1974, 1985, 1999.

BITAR, M. L. *Eficiência dos instrumentos de avaliação em leitura*. 1989. Dissertação (Mestrado) - Pontifícia Universidade Católica de São Paulo, São Paulo, 1989.

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

CONSELHO FEDERAL DE PSICOLOGIA Resolução nº 002/2003. Brasília: CFP, 2003. Disponível em: <<http://www.pol.org.br>>. Acesso em: 21 ago. 2006.

CONDEMARÍN, M.; MILICIC, N. *Test de Cloze: Procedimiento para el desarrollo y la evaluación de la comprensión lectora*. Santiago de Chile: Andrés Bello, 1994.

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

KATO, M. *Aprendizado de leitura*. São Paulo, Martins Fontes, 1987.

KINTSCH, W.; DIJK, T. A. van. Toward a model of text comprehension and production. *Psychological Review*, Pittsburgh, v. 85, n. 5, p. 363-394, Sept. 1978.

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. Test validity and the ethics of assessment. *American Psychologist*, Washington, DC, v. 35, p. 1012-1027, 1980.

MUÑIZ, J. La validación de los testes. *Metodología de las ciencias del comportamiento*, Barcelona, v. 5, n. 2, p. 121-141, jul./dez. 2004.

OLIDEN, P. E. Sobre la validez de los tests. *Psicothema*, Oviedo, v. 15, n. 2, p. 315-321, maio/jul. 2003.

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, winter 1975.

PASQUALI, L. *Psicometria: teoria dos testes na psicologia e na educação*. Petrópolis: Vozes, 2003.

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

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)—Pontificia Universidade de Campinas, Campinas, 1981.

SANTOS, A. A. A. *Leitura entre Universitários: diagnóstico e remediação*. 1990. Tese (Doutorado)—Universidade de São Paulo, São Paulo, 1990.

SANTOS, A. A. A. *Evidências de validade de critério para o Teste de Cloze*. Manuscrito não publicado. 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.

TAYLOR, W. L. Cloze procedure: a new tool for measuring readability. *Journalism Quarterly*, San Jose, v. 30, p. 415-433, 1953.

Recebido em: março de 2009

Aceito em: outubro de 2009