

Estudo de validade do questionário de avaliação de consciência metatextual

Neide de Brito Cunha¹

Fatec Bragança Paulista, Bragança Paulista – SP – Brasil

Acácia Aparecida Angeli dos Santos

Universidade São Francisco, Itatiba – SP – Brasil

Resumo: Considerando a escassez de estudos com consciência metatextual no Brasil, elaborou-se um instrumento para medi-la. Foram realizados dois estudos, um de natureza qualitativa para derivar evidências de validade de conteúdo para o questionário de avaliação da consciência metatextual – piloto (QACM-P), demonstradas pela análise de juízes e estrutura do instrumento. No segundo estudo, de caráter quantitativo, foram buscadas evidências de validade baseadas na relação com outras variáveis quanto aos anos escolares e construtos relacionados e à consistência interna dos itens. Participaram 30 estudantes de uma escola pública, sendo dez crianças de cada ano escolar, do 3º ao 5º. Os instrumentos utilizados foram o QACM-P e dois testes de Cloze. Os resultados indicaram que houve evidência de validade de conteúdo pela análise dos juízes e dos erros das crianças. Houve evidência de validade para construtos relacionados, no entanto não houve pelos anos escolares. Estudos futuros serão necessários para o refinamento desse instrumento, antes de ser aplicado em uma amostra maior e mais representativa.

Palavras-chave: compreensão da leitura; consciência metatextual; ensino fundamental; validade do teste; psicométrica.

VALIDITY STUDY OF THE ASSESSMENT QUESTIONNAIRE OF METATEXTUAL AWARENESS

Abstract: Considering the scarcity of studies with metatextual awareness in Brazil, an instrument to measure it was developed. Two studies were conducted, one to derive qualitative evidence of content validity for the assessment questionnaire of metatextual awareness – pilot (QACM-P), demonstrated by the analysis of judges and the structure of the instrument. In the second study, a qualitative, were sought evidence of validity based on relationships with other variables as the school years and related constructs and the internal consistency of items. 30 students participated of a public school, with 10 children in each school year, 3rd to 5th. The instruments used were the QACM-P and two Cloze tests. The results indicated that there was evidence of content validity by analyzing of judges and the children's errors. There was evidence of validity only for related constructs. Future studies are needed to refine this instrument, before being applied in a larger and more representative sample.

Keywords: reading comprehension; metatextual awareness; basic education; test validity; psychometric.

¹ **Endereço para correspondência:** Neide de Brito Cunha, Centro Paula Souza, Fatec Bragança Paulista, Rua das Indústrias, 130, Distrito Industrial IV, Uberaba – Bragança Paulista – SP – Brasil. CEP: 12926-674. E-mail: atendimento@fatecbpaulista.edu.br.

ESTUDIO DE VALIDEZ DEL CUESTIONARIO DE EVALUACIÓN DE LA CONSCIENCIA METATEXTUAL

Resumen: Considerando la escasez de estudios con consciencia metatextual en Brasil, se elaboró un instrumento para medirla. Fueron realizados dos estudios, un cualitativo para derivar evidencias de validez de contenido para el cuestionario de evaluación de la consciencia metatextual – piloto (QACM-P), demostradas por el análisis de jueces y estructura del instrumento. En el segundo estudio, cuantitativo, fueron buscadas evidencias de validez basadas en la relación con otras variables cuanto a los años escolares y constructos relacionados y la consistencia interna de los ítems. Participaron 30 estudiantes de una escuela pública, siendo diez niños de cada año escolar, de 3º a 5º. Los instrumentos utilizados fueron el QACM-P y dos testes de Cloze. Los resultados indicaron evidencia de validez de contenido por el análisis de jueces y de los errores de los niños. Hubo evidencia de validez solamente para constructos relacionados. Estudios futuros serán necesarios para el refinamiento de ese instrumento, antes que sea aplicado en una muestra mayor y más representativa.

Palabras clave: comprensión de lectura; consciencia metatextual; enseñanza básica; validez de test; psicometría.

Entre 1950 e 1960, surgiu o termo metalinguística, utilizado por linguistas para designar a atividade cognitiva que tem como objeto a própria linguagem. Dependendo do foco, são tratadas as atividades: metafonológicas, metassintáticas, metassemânticas, metapragmáticas e metatextuais, sendo estas duas últimas ainda pouco investigadas (Gombert, 1997).

A aprendizagem de textos envolve, de maneira complementar, conhecimentos implícitos e explícitos. Por meio das situações de aprendizagem, no contexto escolar, a linguagem passa a ser tratada como um objeto a ser refletido e aprendido, gerando conhecimento explícito que permite realizar não apenas atividades epilinguísticas, mas também as metalinguísticas (Gombert, 1997; Mota, 2009).

As atividades epilinguísticas têm por objetivo proporcionar ao usuário da língua oportunidade para refletir sobre os recursos expressivos de que faz uso tanto no plano oral como no escrito. Quando ele realiza tais atividades, sua atenção se volta para a reflexão sobre os recursos que estão sendo utilizados no processo comunicativo em questão. Quanto às atividades metalinguísticas, os Parâmetros Curriculares Nacionais (PCN) as descrevem como a utilização ou construção de uma metalinguagem que possibilite falar sobre a língua. Quando parte integrante de uma situação didática, a atividade metalinguística se desenvolve no sentido de possibilitar ao aluno o levantamento de regularidades de aspectos da língua, a sistematização e a classificação de suas características específicas (Parâmetros Curriculares Nacionais, 1997).

Estudiosos da linguagem consideram duas dimensões do texto: os tipos e os gêneros textuais. Os tipos são definidos pela sua estrutura e por aspectos formais, obedecendo a uma sequência lógica, aspectos lexicais, sintáticos e verbais. Já os gêneros textuais estão presentes nas situações de comunicação determinadas dentro de um dado contexto, sendo consideradas as especificidades do momento histórico em que ocorrem, podendo ser determinados pelo ambiente e pelo suporte em que estão (Marcuschi, 2002).

Marcuschi (2002) considera que os tipos textuais se realizam em todos os gêneros, e, muitas das vezes, o mesmo gênero é realizado em dois ou mais tipos. Nesse sentido, usou o exemplo de uma publicidade que pode estar no formato de um poema ou de uma lista de produtos em oferta, mas mesmo assim serão percebidos como publicidade e poderão propiciar a compra daquele produto por parte de clientes ou usuários.

O trabalho com gêneros textuais é entendido, neste artigo, como a consciência metatextual que, segundo a definição de Gombert (1992), é uma atividade realizada por um indivíduo que trata o texto como um objeto de análise, cujas propriedades podem ser examinadas a partir de um monitoramento intencional em que o indivíduo focaliza de maneira deliberada sua atenção no texto, e não em seus usos. Com base nessa definição é que foram considerados os gêneros textuais neste estudo para escolha dos textos que fizeram parte do instrumento desenvolvido para medir a consciência metatextual.

No Brasil, a consciência metatextual tem sido examinada com diferentes objetivos. Spinillo (2009) fez um levantamento das pesquisas que indicou que alguns autores observaram os critérios utilizados por crianças para definir histórias. Outros autores consideraram também outros tipos de texto, como a carta e a notícia de jornal, para analisar os critérios que as crianças adotavam para identificá-los; a possível progressão em relação a esses critérios com o avanço da idade, da escolaridade e da classe social; o papel desempenhado por intervenções específicas sobre a consciência metatextual e a produção de textos (Albuquerque & Spinillo, 1997, 1998; Ferreira & Spinillo, 2003; Spinillo & Melo, 2008; Spinillo & Pratt, 2005).

Na França, Rouet e Eme (2002) estudaram a consciência metatextual com crianças e universitários e identificaram que ela envolve desde o reconhecimento de características mais básicas (fonema e morfema, por exemplo) até as mais elaboradas, como as semânticas e as do texto. Também identificaram índice positivo de correlação entre o conhecimento sobre textos, ou consciência metatextual, e seu desempenho em tarefas de compreensão, tanto em crianças como em universitários. Os materiais para a investigação incluíram um questionário de conhecimento metatextual e tarefas de compreensão de leitura. No experimento com crianças, os autores pediram a 42 estudantes do 3º ano e 42 do 5º ano para responderem ao questionário que era apresentado oralmente pelo experimentador. Os resultados mostraram um desenvolvimento significativo do conhecimento de aspectos de conhecimento textual, de atividades de planejamento e regulação entre o 3º e 5º anos. No segundo experimento, foi usado um procedimento similar para investigar o conhecimento metatextual de universitários. A maioria dos participantes mostrou conhecimentos elaborados sobre aspectos básicos de textos como: títulos, parágrafos e sumários.

Sobre a possibilidade de encontrar diferenças entre o conhecimento metatextual de alunos em diferentes etapas de escolarização, Spinillo (2009) ressalta que seria uma importante contribuição da psicologia para a educação, no que diz respeito ao desenvolvimento cognitivo, oferecer informação sobre o progresso na aquisição de um dado

conceito, de forma que o conhecimento da criança pudesse ser situado perante essa progressão ao mesmo tempo que pudessem ser apontadas as dificuldades que enfrenta. Para essa identificação de dificuldades, são necessários instrumentos de avaliação que forneçam informações válidas e fidedignas para que possam ser comparados os ganhos efetivos em relação ao que se ensina (Urbina, 2007).

Considerando a progressão do desenvolvimento da consciência metatextual e a partir da ideia de que a instrução desempenha papel fundamental nesse desenvolvimento, entende-se que explorar suas relações com outras habilidades, como a compreensão de leitura, poderia oferecer contribuição empírica para entendê-la. Para avaliar a compreensão de leitura, no Brasil, um dos instrumentos mais utilizados é a técnica de Cloze, conforme constataram Suehiro e Cunha (2012). Ela 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. No presente estudo, os dois textos usados com essa técnica para a avaliação da leitura possuem dados que evidenciam suas qualidades psicométricas, como a existência de validade preditiva dos testes (Santos & Oliveira, 2010).

Da mesma forma que acontece com relação à compreensão de leitura, para que instrumentos de avaliação eficazes sejam disponibilizados para o uso, é necessário que haja evidências empíricas de sua validade e precisão. No processo de construção de um teste, uma escala, um inventário ou um questionário, a validade é um dos temas essenciais, visto que busca demonstrar a utilidade da medida realizada, ou seja, o significado das pontuações obtidas. É justamente a validade que permitirá realizar inferências e interpretações corretas dos resultados quando se aplica um teste e se estabelece a relação com o construto/variável que se mede (Macías, 2007; Urbina, 2007).

O conceito contemporâneo adotado deixa de entender a validade como propriedade do instrumento, assumindo-a como um conceito mais amplo, que diz respeito ao grau de evidências que um instrumento apresenta para sustentar as interpretações pretendidas (Urbina, 2007). As evidências, então, são investigadas em diferentes fontes de validade, organizadas em cinco grandes eixos: conteúdo do teste, processo de resposta do teste, estrutura interna do teste, relações com outras variáveis e consequências do teste (American Educational Research Association, American Psychological Association, & National Council on Measurement in Education, 1999).

No processo de construção de um novo instrumento, a fonte de validade referente ao conteúdo deve ser feita precocemente, para examinar se os conceitos abstratos, que originaram o teste, estão representados adequadamente, como expresso por indicadores observáveis e mensuráveis da qualidade dos itens. Vários métodos empíricos podem ser utilizados para buscar a evidência de validade de conteúdo, e um deles se refere a um painel de juízes, especialistas da área, a fim de que comparem os objetivos do teste com seu conteúdo, avaliando a pertinência de cada item. O resultado do painel de juízes é usado para definir os itens que serão mantidos, revisados ou eliminados (Exner, 1993), procedimento adotado no presente estudo.

Com base nessas considerações apresentadas sobre a escassez de pesquisas com consciência metatextual e sobre os procedimentos para a busca de evidências de validade para um instrumento de medida, considera-se oportuno estudar os parâmetros psicométricos do questionário de avaliação da consciência metatextual – piloto (QACM-P). Os problemas que se colocam neste trabalho são:

- As questões propostas contêm os gêneros textuais ensinados no primeiro ciclo do ensino fundamental?
- Há evidências de validade de conteúdo para o QACM-P?
- Há progressão da consciência metatextual de acordo com o ano escolar?
- Há correlação entre consciência metatextual e compreensão de leitura, evidenciando que os processos cognitivos são compartilhados?
- Há consistência interna dos itens?

Para responder a essas questões, foram realizados dois estudos, descritos a seguir.

Estudo I

Objetivos

Foram definidos os seguintes objetivos para o estudo 1: derivar evidências de validade de conteúdo do QACM-P com base na análise de juízes; analisar a estrutura do instrumento com base nas respostas das crianças; verificar variações de nomenclatura para os gêneros textuais de acordo com as respostas de professores das turmas analisadas.

Método

Participantes

Participaram como juízas para o conteúdo do QACM-P três professoras universitárias, de um curso de Pedagogia de uma universidade do interior do Estado de São Paulo, que ministravam disciplinas relacionadas à alfabetização e língua portuguesa. Duas delas eram formadas em Letras e doutorandas, e uma era formada em Psicologia, com doutorado em Educação. Para responderem aos instrumentos, foram indicados por seus professores 30 alunos do 3º ao 5º ano do ensino fundamental, de uma escola pública do interior do Estado de São Paulo. Do total, 16 eram meninos e do total geral 10 do 3º, 11 do 4º e 9 do 5º ano. Finalmente, os professores desses alunos, que lecionavam nos anos investigados, também responderam ao QACM-P, sendo um homem e cinco mulheres.

Instrumentos

- *Ficha de avaliação do QACM-P*: composta pelos seguintes dados: nome, idade, maior formação e atuação. Nessa ficha, as juízas deveriam avaliar a linguagem, a

temática e o conteúdo, e assinalar se estavam adequados ou se exigiam modificações e, em caso positivo, especificá-las.

- **QACM-P:** constituído por 11 textos, apresentados de maneira que não se configurassem os suportes deles, como gibi, jornal, tela de computador etc. Foi elaborado com questões abertas, para que os estudantes pudessem escrever os tipos de textos numa linha abaixo destes.

Alguns textos para elaboração do questionário foram obtidos do programa “Ler e Escrever”, da Secretaria da Educação do Governo do Estado de São Paulo, e do “Programa de Formação de Professores Alfabetizadores”, da Secretaria de Educação Fundamental do Ministério da Educação, módulo 1, de janeiro de 2001. Outros foram inventados, como lista e *e-mails* ou retirados da internet, e vários deles sofreram adaptações. Foram considerados os tipos de textos sugeridos nos Parâmetros Curriculares Nacionais (1997) para o primeiro ciclo do ensino fundamental, volume 2 – “Língua Portuguesa”.

Quadro 1. Exemplo de um item do QACM-P

Leia o texto e escreva na linha abaixo a que gênero textual ele pertence.

Dedo mindinho

Seu vizinho,

Pai de todos

Fura-bolos

Cata-piolhos.

Fonte: Elaborado pelas autoras.

Procedimentos

Após a autorização da Secretaria da Educação, solicitou-se a aprovação pelo Comitê de Ética, sob protocolo de Certificado de Apresentação para Apreciação Ética (Caae) nº 0277.0.142.000-10, e enviaram-se, por *e-mail*, a três juízas o QACM-P e o protocolo para avaliação. Foi-lhes solicitado que analisassem o conteúdo dos itens do questionário, considerando a linguagem e a temática. Deveriam, também, avaliar se esses tópicos estavam “adequados” ou se “exigiam modificações (com especificações)”. Elas anotaram, no próprio instrumento, seus comentários e sugestões. O tempo médio entre a entrega e devolução dos protocolos foi de uma semana.

Solicitou-se aos pais dos alunos autorização para que os filhos participassem da pesquisa, e, aos que concederam, foram entregues os Termos de Consentimento Livre e Esclarecido para assinatura. Os pais e os alunos foram assegurados de que a pesquisa não influenciaria na nota, não afetaria o seu desempenho acadêmico na escola, além de ser de caráter estritamente confidencial.

Em seguida, foi entregue o QACM-P. Ao receberem o instrumento com 11 textos, os participantes foram orientados a lê-los e escrever o nome do tipo do texto lido. Após cada resposta, era solicitado que justificassem oralmente a resposta dada, procedimento que foi gravado, caso houvesse divergência e/ou complementação da resposta escrita e da explicação dos estudantes. O tempo de aplicação não ultrapassou 30 minutos para esse instrumento.

Os professores das turmas nas quais ocorreu a coleta responderam ao QACM-P na sala dos professores, no horário de intervalo deles, conforme orientação da diretora da unidade escolar. A mesma pesquisadora aplicou coletivamente os questionários que foram respondidos em, no máximo, 15 minutos. Decidiu-se que esse procedimento poderia reforçar a análise das juízas quanto à adequação dos textos e variedade de nomenclaturas utilizadas para os gêneros textuais constantes no QACM-P.

Análise dos dados

Primeiramente, foram analisadas as respostas das juízas para verificar o potencial do instrumento. Depois, para os QACM-P respondidos pelos alunos, foram analisados seus erros, com base nas seguintes categorias, consideradas dos menos para os mais graves: erro A – gênero textual aproximado; erro B – característica do gênero textual; erro C – não é gênero textual. Com base nas justificativas das crianças para suas respostas, analisou-se o potencial do instrumento, considerando se os tipos de texto eram conhecidos, de fácil compreensão e com linguagem adequada.

Quanto aos QACM-P respondidos pelos professores dos anos investigados, foi principalmente observada a variedade de nomenclaturas utilizadas para os diferentes gêneros textuais utilizados. Esse procedimento serviu de base para a elaboração de uma relação de termos que seriam considerados sinônimos válidos na correção dos testes das crianças.

Resultados

De acordo com os objetivos propostos, primeiramente foi verificada a análise pelas três juízas, que avaliaram a linguagem, a temática e o conteúdo do QACM-P. Uma professora fez uma proposta de alteração que foi acatada: “Ao invés do nome ‘gibi’, sugiro o termo ‘história em quadrinhos’ que mais é utilizada no contexto de sala de aula”. Outra fez a seguinte observação:

A única sugestão que tenho é que os elementos paratextuais sejam apresentados. Por exemplo, a *e-mail*. Acho que deve aparecer uma caixinha parecida com a da internet. Não sei se é essa a intenção de vocês, ou se vocês querem que os alunos reconheçam o gênero apenas através da linguagem utilizada.

Nesse caso, não acatamos a sugestão, visto que o objetivo do instrumento é justamente o contrário, ou seja, que o respondente reconheça o texto sem o seu suporte.

Por fim, a última juíza considerou que todos os itens da estrutura do instrumento estavam adequados.

De acordo com o segundo objetivo, foram buscadas evidências de validade de conteúdo do QACM-P, demonstradas com base em dados sobre a estrutura do instrumento. Analisaram-se os erros das crianças para verificar se os itens seriam suficientes e apropriados, com base nas seguintes categorias, consideradas dos menos para os mais graves: erro A – gênero textual aproximado; erro B – característica do gênero textual; erro C – não é gênero textual. Os acertos também foram considerados para os 11 textos do questionário. Esses dados aparecem na Tabela 1.

Tabela 1 – Erros e acertos das crianças no QACM-P

Gêneros textuais	Total Erro A: gênero textual aproximado	Total Erro B: característica do gênero textual	Total Erro C: não é gênero textual	Total de erros	Total de acertos
Quadrinho	2	0	11	13	17
Parlenda	8	11	7	26	4
Poema	1	5	9	15	15
Adivinha	0	0	5	5	25
Regra	0	6	18	24	6
E-mail	9	4	12	25	5
Notícia	2	16	7	25	5
Conto	2	1	7	10	20
Bula	0	18	11	29	1
Receita	0	0	5	5	25
Lista	0	0	13	13	17
Total geral	24	61	105	190	140

Fonte: Elaborada pelas autoras.

Os erros de tipo A (gênero textual aproximado) – com total de 24 – ocorreram mais para *e-mail* (N = 9) e parlenda (N = 8). As crianças responderam, por exemplo, “carta” para *e-mail* e “poesia” para parlenda.

Os erros do tipo B (característica do gênero textual) – com total de 61 – aconteceram mais para bula (N = 18) e notícia (N = 16). Exemplos das respostas das crianças para esse tipo de erro são: “de remédio” para bula e “informação” para notícia.

O maior número de erros foi correspondente ao tipo C (não é gênero textual), com total de 105, e a maior quantidade aconteceu para regra (N = 18) e lista (N = 13). São exemplos de respostas erradas: “serve para ensinar”, no caso da regra, e “supermercado” para a lista.

Quanto aos acertos, observa-se que ocorreram mais para adivinha (N = 25), receita (N = 25) e conto (N = 20). No total, foram 140 acertos, número que corresponde a 42,4%

do total de 330, correspondentes a 11 alternativas para 30 alunos. Quanto ao total de erros, que poderia ser também de 330, foi de 190, correspondendo a 57,5% do total.

Quanto ao terceiro objetivo, analisaram-se as respostas dos seis professores dos anos investigados nesta pesquisa. Eles também tiveram erros, porém poucos e do tipo B. Como exemplos, podem-se citar “instrução” para regra e “literário” para poesia. Nesse caso, não foi considerado que houve problemas que sugerissem alteração quanto às nomenclaturas utilizadas, visto que os erros do tipo B (característica do gênero textual) não são considerados tão graves em relação à consciência metatextual.

Estudo 2

Objetivos

No estudo 2, pretendeu-se averiguar as evidências de validade para o QACM-P, baseadas na relação com outras variáveis, a saber, os anos escolares e construtos relacionados. Objetivou-se também analisar a consistência interna dos itens.

Instrumentos

- QACM-P: já descrito no estudo 1.
- Testes de Cloze: utilizaram-se os textos “A princesa e o fantasma” e “Uma vingança infeliz”, usados para a avaliação da compreensão de leitura com base no que preconiza Taylor (1953), referindo-se à possibilidade de omissão de vocábulos substituídos por um traço. Os respondentes preenchem as lacunas com a palavra que supõem dar mais sentido ao texto. As histórias aqui utilizadas foram construídas por Santos (2005) e possuem evidência de validade de critério e índices de precisão adequados, estabelecidos por um estudo com alunos de 2ª a 4ª séries (atualmente do 3º ao 5º ano).

Procedimentos

Para o QACM-P, os procedimentos já foram descritos no estudo 1. Para os testes de Cloze, a coleta foi feita individualmente com as crianças no refeitório da escola. Nessa fase, as crianças deveriam ler o texto até o fim e depois escrever a palavra que melhor completasse seu sentido. Foi alertado que o tamanho do traço correspondia ao tamanho da palavra que cabia em cada espaço. O tempo de aplicação não ultrapassou 20 minutos para cada texto.

Critérios de correção

No caso do QACM-P, foi atribuído um ponto a cada acerto. O número total de acertos poderia alcançar 11 pontos, correspondentes ao número de questões do instrumento.

Quanto aos testes de Cloze, atribuiu-se um ponto a cada acerto, e consideraram-se corretas as palavras usadas pelo autor do texto. Essa forma de correção é denominada

literal ou verbatim, cujo objetivo é evitar o subjetivismo na avaliação. Assim, o número de acertos máximos para cada teste foi de 15, correspondentes ao número de omissões.

Resultados

De acordo com o primeiro objetivo, foram averiguados os escores médios do QACM-P e das provas de avaliação da compreensão de leitura (Cloze 1 e 2). Os resultados das análises descritivas são mostrados na Tabela 2.

Tabela 2. Estatísticas descritivas das respostas aos instrumentos

	N	Mínimo	Máximo	Média	Desvio padrão
QACM	30	0	7	4,63	2,26
Cloze 1	30	0	12	6,5	3,87
Cloze 2	30	0	12	5,1	3,78

Fonte: Elaborada pelas autoras.

Para o QACM-P, o máximo de acertos correspondeu a sete pontos, de um máximo de 11 que poderia ser atingido para o questionário. Cabe lembrar que as justificativas das crianças para cada resposta ao QACM-P foram gravadas a fim de verificar se elas poderiam expressar oralmente respostas diferentes das escritas, no entanto constatou-se que a maioria não conseguiu explicar de outra maneira o que já havia afirmado. São exemplos de respostas orais: “é conto porque conta uma coisa”; “é história em quadrinhos porque é o Cebolinha”; “não sei como chama isso”; “é supermercado porque tem coisa que compra lá”.

Quanto à compreensão de leitura, pode-se observar que, tanto para o Cloze 1 quanto para o 2, o mínimo (zero) e o máximo (12) de acertos foram os mesmos, já que o escore total seria de 15 pontos para cada teste. Quanto às médias de acertos das crianças, Cloze 1 ($M = 6,5$; 43,33%), Cloze 2 ($M = 5,1$; 34%) e QACM-P (4,64; 42,18%), constata-se um nível dificuldade maior no Cloze 2 e índice acima da média somente no Cloze 1 para essa amostra.

Foram utilizadas provas de estatísticas não paramétricas, visto que o resultado da prova de Shapiro-Wilk para a análise da curva da pontuação dos escores da consciência metatextual permitiu recusar a hipótese de normalidade ($W = 0,870$; $p = 0,002$). Assim, para as análises inferenciais entre as médias dos postos aferidas pelos dois instrumentos, assumiu-se o nível de significância $p \leq 0,05$. Assumindo os dados como ordinais, adotou-se o teste de Kruskal-Wallis para a comparação de grupos, bem como a prova de correlação de Spearman para a análise das pontuações nas duas medidas.

Primeiramente, contudo, optou-se por analisar a consistência interna dos itens, e o resultado obtido mostrou que o valor era satisfatório ($\alpha = 0,80$). A busca pela evidência de validade de critério do QACM-P foi investigada com base na separação dos anos

escolares, o que também foi realizado com os dois testes de Cloze. Esses resultados são apresentados na Tabela 3.

Tabela 3. Comparação entre as pontuações dos três anos pela prova de Kruskal-Wallis

	Ano	N	Média dos ranks	χ^2 (2,27)	p
Cloze 1	3°	10	7,45	13,610	0,001
	4°	11	17,82		
	5°	9	21,61		
Cloze 2	3°	10	7,85	13,456	0,001
	4°	11	16,82		
	5°	9	22,39		
C. metatextual	3°	10	11,25	4,365	0,113
	4°	11	16,14		
	5°	9	19,44		

Fonte: Elaborada pelas autoras.

Quanto ao QACM-P, embora as médias dos postos tenham aumentado com a evolução dos anos, não houve diferença significativa, podendo ser atribuída ao acaso. Houve diferença significativa quanto aos níveis de significância entre as médias dos postos dos instrumentos, com relação aos anos escolares para ambas as provas de Cloze 1 e 2 ($p < 0,001$).

Ainda com relação às evidências de validade, foi avaliada a relação entre construtos relacionados. A prova de correlação de Spearman revelou índices positivos e de magnitude forte (Dancey & Reidy, 2006) entre os escores do Cloze 1 e o QACM-P ($\rho = 0,73$), e entre o Cloze 2 e o QACM-P ($\rho = 0,70$).

Por fim, usou-se o alfa de Cronbach para o estudo da consistência interna do QACM-P. O valor obtido foi o de alfa = 0,81.

Discussão

No estudo 1, a análise das juízas para evidenciar validade de conteúdo foi satisfatória quanto à linguagem, à temática e ao conteúdo do QACM-P. Dessa análise, foram considerados os itens que deveriam ser mantidos e os que deveriam ser revisados, e nenhum deles foi indicado para ser eliminado (Exner, 1993). Cabe lembrar que os elementos paratextuais não foram apresentados para que as crianças reconhecessem o gênero apenas por meio da linguagem utilizada, atestando o conhecimento sobre as características peculiares de cada um.

Quanto à validade de conteúdo com base nos erros das crianças, nas três categorias, dos menos aos mais graves, ficou evidenciado que, em todos os casos, houve acertos,

mesmo quando foram poucos como: um para bula e quatro para parlenda. Em ambos os casos, houve muitos erros, porém de tipo B (característica do gênero textual), demonstrando que as crianças conseguiam identificar especificidades dos textos. O número maior de erros em gêneros como bula (com 29 no total), parlenda (com 26), *e-mail* e notícia (com 25) indica que eles precisam ter mais situações de instrução para alcançarem o nível de acertos de receita e adivinha (com 25 cada) e conto (com 20). Isso permitiria que os alunos realizassem não somente atividades epilinguísticas, como também as metalinguísticas, por meio de intervenções exclusivas sobre a consciência metatextual e a produção de textos (Parâmetros Curriculares Nacionais, 1997; Gombert, 1997; Spinillo, 2009).

Nesse sentido, quanto aos totais de erros e acertos das crianças no QACM-P, considerou-se que não havia itens que não pudessem ser identificados ou que tivessem um número muito alto, a ponto de alcançar efeito teto. Assim, considerou-se que os itens estavam apropriados para medir a consciência metatextual. As respostas dos professores ao instrumento pareceram comprovar essas afirmações.

No estudo 2, pelos escores obtidos nos instrumentos, pôde-se constatar a dificuldade das crianças, tanto para reconhecer os gêneros textuais como na compreensão de leitura, que ficaram com a média abaixo dos 50% de acertos, como verificado nos estudos de Lima (2010) e Santos, Sisto e Noronha (2010). Esses resultados podem ser explicados pela falta de uso de estratégias de compreensão e processamento inadequado da informação. Esse fato foi constatado, ainda, por meio da expressão oral, nas respostas gravadas para o QACM-P.

Os testes de Cloze usados, tal como em estudos anteriores, continuaram discriminando os anos escolares (Santos, 2005; Santos & Oliveira, 2010). Os escores do QACM-P, embora tenham demonstrado ser sensíveis para captar o progresso das crianças para identificar os gêneros textuais, pelo aumento das médias com a evolução dos anos, não obtiveram diferenças estatisticamente significativas.

Ressalta-se, contudo, a identificação da evidência de validade pela comparação entre medidas que avaliam construtos relacionados, a saber, a consciência metatextual e a compreensão de leitura. O índice forte de correlação identificado entre as medidas comprova a hipótese de que os processos cognitivos envolvidos nas duas medidas são compartilhados. Pelo índice identificado de consistência interna, é possível afirmar que todos os itens medem o mesmo construto, a saber, a consciência metatextual.

Diante dos resultados alcançados, o conhecimento das crianças pôde ser situado nas habilidades medidas. Além disso, identificaram-se algumas das dificuldades que enfrentaram para responderem aos instrumentos (Mota, 2009; Spinillo, 2009). Assim sendo, ficou demonstrada a importância desse diagnóstico e de investir mais no aprimoramento do questionário para avaliar a consciência metatextual.

Uma possível explicação para que não tenha sido evidenciada a validade de critério pelos anos escolares é o formato do instrumento, com questões abertas. Talvez um questionário com questões fechadas favoreça a distinção dos gêneros pelos estudantes, assim como a aplicação e correção dele sejam mais eficientes e objetivas, sem que

haja a necessidade de considerar os sinônimos e os problemas de interpretação da escrita das crianças.

Nesse sentido, há necessidade de refinamento desse instrumento, embora tenha havido aprovação das juízas e consistência interna adequada. Outra limitação do estudo pode ter sido o número reduzido de alunos e o fato de serem de uma só escola, o que pode ter contribuído para que não fosse encontrada a evidência de validade de critério pelos anos escolares no QACM-P. Espera-se que com o aprimoramento do instrumento, em estudo posterior, sejam sanados os problemas encontrados e que o questionário possa ser útil para o diagnóstico de dificuldades encontradas pelas crianças com respeito à consciência metatextual, para que os professores possam intervir eficazmente para que haja o crescimento linguístico adequado para cada ano escolar.

Referências

- Albuquerque, E. B. C., & Spinillo, A. G. (1997). O conhecimento de crianças sobre diferentes tipos de texto. *Psicologia: Teoria e Pesquisa, 13*, 329-338.
- Albuquerque, E. B. C., & Spinillo, A. G. (1998). Consciência textual em crianças: critérios adotados na identificação de partes de textos. *Revista de Estudios e Investigación en Psicología y Educación, 3*, 145-158.
- American Educational Research Association, American Psychological Association & National Council on Measurement in Education. (1999). *Standards for educational and psychological testing*. Washington: American Psychological Association.
- Dancey, C. P., & Reidy, J. (2006). *Estatística não paramétrica para psicologia*. Porto Alegre: Artmed.
- Exner, C. E. (1993). Content validity of the in-hand manipulation test. *The American Journal of Occupational Therapy, 47*(6), 505-513.
- Ferreira, A. L., & Spinillo, A. G. (2003). Desenvolvendo a habilidade de produção de textos em crianças a partir da consciência metatextual. In M. R. Maluf (Org.). *Metalinguagem e aquisição da escrita: contribuições da pesquisa para a prática da alfabetização* (pp. 119-148). São Paulo: Casa do Psicólogo.
- Gombert, J. E. (1992). *Metalinguistic development*. Hertfordshire: Harvester Wheatsheaf.
- Gombert, J. E. (1997). Metalinguistic development in first language acquisition. In L. van Lier, & D. Corson. *Encyclopedia of language and education* (Vol. 6, pp. 43-52). Netherlands: Kluwer Academic Publishers.
- Lima, T. H. (2010). *Cloze e escala de avaliação de estratégias de aprendizagem: evidências de validade*. Dissertação de mestrado, Universidade São Francisco, Itatiba, SP, Brasil.

- Macías, A. B. (2007). La consulta a expertos como estrategia para la recolección de evidencias de validez basadas en el contenido. *Investigación Educativa*, 7, 5-14.
- Marcuschi, L. A. (2002). Gêneros textuais: definição e funcionalidade. In A. P. Dionísio, A. R. Machado & M. A. Bezerra (Orgs.). *Gêneros textuais e ensino* (pp. 19-36). Rio de Janeiro: Lucerna.
- Mota, M. M. P. E. (2009). Desenvolvimento metalinguístico. In M. Mota (Org.). *Desenvolvimento metalinguístico: questões contemporâneas* (pp. 9-18). São Paulo: Casa do Psicólogo.
- Parâmetros Curriculares Nacionais: Português* (1997). Brasília: MEC, SEF.
- Rouet, J. F., & Eme, E. (2002). The role of metatextual knowledge in text comprehension: issues in development and individual differences. In P. Chambers, M. Izaute & P. J. Marescaux (Eds.). *Metacognition: process, function and use* (pp. 121-134). Amsterdam: Kluwer.
- Santos, A. A. A. (2005). *O teste de Cloze como instrumento de avaliação da compreensão em leitura* (Relatório Técnico). Itatiba: Universidade São Francisco.
- Santos, A. A. A., & Oliveira, E. Z. (2010). Avaliação e desenvolvimento da compreensão da leitura no ensino fundamental. *Psico-USF*, 15(1), 81-91.
- Santos, A. A., Sisto, F. F., & Noronha, A. P. P. (2010). Toni 3 – forma A e teste de Cloze: evidências de validade. *Psicologia: Teoria e Pesquisa*, 26(3), 399-405.
- Spinillo, A. G. (2009). A consciência metatextual. In M. Mota (Org.). *Desenvolvimento metalinguístico: questões contemporâneas* (pp. 77-114). São Paulo: Casa do Psicólogo.
- Spinillo, A. G., & Melo, K. (2008). The production of written stories and metatextual awareness: an intervention study with elementary school children. *International Conference of the EARLI Special Interest Group on Writing*, Sweden, Lund, 11.
- Spinillo, A. G., & Pratt, C. (2005). Sociocultural differences in children's genre knowledge. In T. Kostouli (Org.). *Writing in context(s): textual practices and learning processes in sociocultural settings* (pp. 27-48). New York: Springer.
- Suehiro, A. C. B., & Cunha, N. B. (2012). Avaliação da leitura e escrita em crianças: produção científica brasileira da primeira década de 2000. In E. Boruchovitch, A. A. Santos & E. Nascimento (Orgs.). *Avaliação psicológica nos contextos educativo e psicossocial* (pp. 105-124). São Paulo: Casa do Psicólogo.
- Taylor, W. L. (1953). Cloze procedure: a new tool for measuring readability. *Journalism Quarterly*, 30, p. 415-433.
- Urbina, S. (2007). *Fundamentos da testagem psicológica*. Porto Alegre: Artmed.

Submissão: 20.2.2013

Aceitação: 23.10.2013