



## Avaliação da consciência metatextual: um instrumento de medida

Neide de Brito Cunha\*, Acácia Aparecida Angeli dos Santos\*\*

\*Universidade do Vale do Sapucaí, Minas Gerais, Brasil, \*\*Universidade São Francisco, São Paulo, Brasil

### Resumo

Com base na Psicologia Cognitiva foi construído o Questionário de Avaliação da Consciência Metatextual (QACM). Foi realizado um estudo piloto que indicou evidência de validade de conteúdo pela análise dos juízes e pelos erros das crianças e evidências de validade entre medidas que avaliam construtos relacionados, no caso com a compreensão de leitura, medida por testes de Cloze. O instrumento foi reelaborado e outra pesquisa foi realizada com 315 alunos do 3º ao 5º ano do ensino fundamental, de escolas públicas. Foram evidenciadas as validades de construtos convergentes e de critério, distinguindo os alunos pelo seu nível de escolaridade.

Palavras chave: consciência metatextual, avaliação psicológica, estudos de validade.

A linguagem passa a ser tratada como um objeto a ser refletido e aprendido, no contexto escolar, gerando conhecimento explícito que permite realizar atividades metalinguísticas além das epilinguísticas. Essas últimas atividades objetivam proporcionar ao usuário da língua oportunidade para refletir sobre os recursos expressivos que usa tanto no plano oral como no escrito, voltando sua atenção para a reflexão sobre os recursos que estão sendo utilizados no processo comunicativo em questão (Gombert, 1997; Mota, 2009).

Os Parâmetros Curriculares Nacionais (PCN) descrevem as atividades metalinguísticas como a utilização ou construção de uma metalinguagem que possibilite falar sobre a língua. Quando parte integrante de uma situação didática, elas se desenvolvem no sentido de possibilitar ao aluno o levantamento de regularidades de aspectos da língua, a sistematização e a classificação de suas características específicas (Brasil, 1997).

Há duas dimensões do texto: os tipos e os gêneros textuais, segundo estudiosos da linguagem. Os tipos são definidos pela sua estrutura e aspectos formais, obedecendo a uma sequência lógica, aspectos lexicais, sintáticos e verbais. Quanto aos gêneros textuais estão presentes nas situações de comunicação determinadas dentro de um dado contexto. Nesse sentido, os tipos textuais se realizam em todos os gêneros, ocorrendo, muitas das vezes, o mesmo gênero sendo realizado em dois ou mais tipos (Marcuschi, 2002).

Considerando a progressão do desenvolvimento da consciência metatextual e a partir da ideia de que a instrução desempenha papel fundamental neste 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). 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 experimental dos testes (Santos & Oliveira, 2010).

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. O conceito contemporâneo de validade diz respeito ao grau de evidências que um instrumento apresenta para sustentar as interpretações pretendidas (Macías, 2007; Urbina, 2007). As evidências 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 (AERA/APA/NCME, 1999).

Vários métodos empíricos podem ser utilizados para buscar a evidência de validade de conteúdo, sendo que 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), que foi o procedimento adotado no presente estudo.

Com base nessas considerações foram estudados 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 as questões, foram realizados três estudos, descritos a seguir.

### Estudo 1

#### Objetivos

Foram definidos os seguintes objetivos: 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.

## 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. Para responder 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

1. Ficha de Avaliação do Questionário de Avaliação da Consciência Metatextual- Piloto: composta pelos seguintes dados: Nome, Idade, Maior Formação e Atuação. Nela os juízes, deveriam avaliar a linguagem, a temática e o conteúdo e assinalar se estavam adequados ou se exigiam modificações e, em caso positivo, especificá-las.
2. Questionário de Avaliação da Consciência Metatextual – Piloto - constituído por 11 textos, apresentados de maneira que não se configurassem os suportes dos mesmos, 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 dos mesmos.

Alguns textos para elaboração do questionário foram obtidos do programa Ler e Escrever da Secretaria da Educação do Governo do Estado de São Paulo, 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, sendo que vários deles sofreram adaptações. Foram considerados os tipos de textos sugeridos nos Parâmetros Curriculares Nacionais (Brasil, 1997) para o primeiro ciclo do Ensino Fundamental, Volume 2 – Língua Portuguesa.

## Procedimentos

Após a autorização da Secretaria da Educação, foi solicitada a aprovação pelo Comitê de Ética, sob Protocolo de CAAE: 0277.0.142.000-10 e enviado por e-mail a três juízas o Questionário de Avaliação da Consciência Metatextual – Piloto (QACM-P) e o protocolo para sua 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.

Foi solicitado aos pais dos alunos autorização para que seus 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, era entregue o QACM-P, para o qual era explicado que havia 11 textos e que deviam ler cada um 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 a explicação dos estudantes. O tempo de aplicação não ultrapassou 30 min. 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 dos juízes 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 dos juízes 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, foi analisado 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

A análise pelos três juízes, quanto à linguagem, à temática e ao conteúdo do QACM-P indicou 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, segundo os erros das crianças, para verificar se os itens seriam suficientes e apropriados, com base nas categorias criadas. Os acertos também foram considerados para os 11 textos do questionário. Esses dados aparecem na Tabela 1.

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, foram analisadas as respostas dos seis professores dos anos investigados nesta pesquisa. 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.

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

Gêneros textuais	Total Erro A Gênero aproximado	Total Erro B Característica do gênero	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

## Estudo 2

### Objetivos

Pretendeu-se no Estudo 2 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

1. Questionário de Avaliação da Consciência Metatextual –P – já descrito no Estudo 1.

2. Testes de Cloze “A Princesa e o Fantasma” e “Uma vingança infeliz”.

São textos 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 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 segunda a quarta séries (atuais 3º a 5º anos) do Ensino Fundamental.

### 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. O tempo de aplicação não ultrapassou 20 minutos para cada texto.

### Critérios de Correção

Para o QACM-P foi atribuído um ponto para cada acerto. O número total de acertos poderia alcançar 11 pontos, correspondentes ao número de questões do instrumento. Para os testes de Cloze, foi atribuído um ponto para cada acerto, sendo consideradas corretas as palavras que tivessem sido as mesmas usadas pelo autor do texto. Assim, o número de acertos máximos para cada teste foi de 15, correspondentes ao número de omissões.

## Resultados

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 ( $W=0,870$ ;  $p=0,002$ ). Foi analisada 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 2.

Tabela 2. Comparação entre as pontuações dos três anos pela prova de Kruskal-Wallis.

	Ano	N	Média dos ranks	X <sup>2</sup> (2,27)	p
Cloze 1	3º	10	7,45	13,610	0,001
	4º	11	17,82		
	5º	9	21,61		
Cloze 2	3º	10	7,85	13,456	0,001
	4º	11	16,82		
	5º	9	22,39		
C. Metatextual	3º	10	11,25	4,365	0,113
	4º	11	16,14		
	5º	9	19,44		

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 alpha de Cronbach para o estudo da consistência interna do QACM-P. O valor obtido foi o de  $\alpha=0,81$ .

### Estudo 3

#### Objetivo

Os objetivos estabelecidos foram: averiguar a consciência metatextual por meio do QACM; analisar sua validade pela relação com medidas de compreensão de leitura e o critério de diferenciação do nível escolar.

#### Participantes

Participaram 315 alunos do 3º ao 5º ano do ensino fundamental, de escolas públicas, cujas idades variaram de 8 a 13 anos, sendo 169 meninos.

#### Instrumentos

Foram utilizados o QACM reformulado com 11 questões fechadas e quatro alternativas de resposta e os testes de Cloze já descritos no Estudo 2.

#### Procedimentos

Foram os mesmos já descritos no Estudo 2.

#### Resultados

Recorreu-se novamente a provas da estatística não paramétrica para as análises, sendo, portanto, as médias resultantes do ranking de postos. Os valores obtidos com a prova de Kruskal Wallis para todas as medidas são apresentados na Tabela 3.

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

	Ano	N	Média dos ranks	$\chi^2$ (gl=2)	p
Consciência Metatextual	3º	69	102,94	36,862	<0.001
	4º	123	162,85		
	5º	123	184,04		

Houve diferença significativa quanto ao nível de significância entre as médias dos postos, o que mostra que as diferenças não ocorreram por acaso. Ainda com relação às evidências de validade, foi verificada a relação entre medidas que avaliam construtos relacionados. A prova de correlação de *Spearman* revelou índices positivos e significativos, de magnitude moderada, entre os escores do QACM e o Cloze 1 ( $r = 0,68$ ;  $p < 0,001$ ) e o Cloze 2 ( $r = 0,57$ ;  $p < 0,001$ ) (Dancey & Reidy, 2006). No entanto, verificou-se que 84 crianças (26,66%) conseguiram alcançar o teto da pontuação, sendo 13 do 3º ano, 30 do 4º e 41 do 5º ano, o que indica que, à medida que a escolarização avança, aumenta o nível da consciência metatextual das crianças. Esse fato não era esperado.

#### Discussão e considerações finais

No estudo 1, a análise dos juízes para evidenciar validade de conteúdo foi satisfatória quanto à linguagem, a temática e o conteúdo do QACM-P. Dessa análise foram considerados os itens que deveriam ser mantidos e os que deveriam ser revisados, sendo que nenhum deles foi indicado para ser eliminado (Exner, 1993). 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: 1 para bula e 4 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.

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. 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. Esses resultados podem ser explicados pela falta de uso de estratégias de compreensão e processamento inadequado da informação.

No Estudo 3, embora o QACM, embora tenha demonstrado ser sensível para captar o progresso das crianças em identificar os gêneros textuais com o avanço da escolaridade, teve um nível de acertos acima do esperado. Esse fato indica que o instrumento pode estar defasado quanto ao número de questões, com necessidade de que sejam incluídos mais gêneros textuais, e também quanto ao número de alternativas para escolha da resposta correta.

#### Referências

- American Educational Research Association, American Psychological Association and National Council on Measurement in Education. (1999). *Standards for educational and psychological testing*. Washington, DC: American Psychological Association.
- Brasil. (1997). Secretaria de Educação Fundamental. *Parâmetros Curriculares Nacionais: Português*. Brasília: MEC/SEF.
- Exner, C. E. (1993). Content validity of the in-hand manipulation test. *The American Journal of Occupational Therapy*, 47(6), 505-513.
- Gombert, J. E. (1997). Metalinguistic development in first language acquisition. In L. van Lier, & D. Corson. *Encyclopedia of language and education* (pp. 43- 52, v. 6). Netherlands: Kluwer Academic Publishers.
- 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.
- 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.

- Spinillo, A. G. (2009). A consciência metatextual. In M. Mota (Org.), *Desenvolvimento metalinguístico: questões contemporâneas* (pp. 77-114). São Paulo: Casa do Psicólogo.
- 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. 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.