

**INSTITUTO NACIONAL DE MATEMÁTICA PURA E APLICADA**

**HELIO FURTADO COSTA**

**PRODUÇÃO E REVISÃO DE ITENS DE MÚLTIPLA ESCOLHA**

Rio de Janeiro

2014

HELIO FURTADO COSTA

**PRODUÇÃO E REVISÃO DE ITENS DE MÚLTIPLA ESCOLHA**

Trabalho de conclusão de curso apresentado ao Instituto Nacional de Matemática Pura e Aplicada (Impa) como requisito parcial à obtenção do título de mestre em matemática.

Orientador: Prof. Eduardo Wagner

Rio de Janeiro

2014

HELIO FURTADO COSTA

**PRODUÇÃO E REVISÃO DE ITENS DE MÚLTIPLA ESCOLHA**

Trabalho de conclusão de curso apresentado ao Instituto Nacional de Matemática Pura e Aplicada (Impa) como requisito parcial à obtenção do título de mestre em matemática.

Data de aprovação: \_\_\_\_/\_\_\_\_/\_\_\_\_

Banca Examinadora:

---

---

---

## **AGRADECIMENTOS**

Agradeço ao professor Eduardo Wagner por gentilmente orientar este trabalho e, especialmente, por ser o responsável pelo meu ingresso na área de produção de itens de múltipla escolha.

Agradeço de maneira muito carinhosa a todos os organizadores do PROFMAT – Mestrado Profissional em Matemática em Rede Nacional, que enfrentaram os desafios típicos do pioneirismo, com o objetivo de melhorar a formação matemática dos professores brasileiros e, por consequência, elevar o nível de proficiência dos estudantes do País, chave para reduzir as desigualdades sociais.

Agradeço ainda aos funcionários do Instituto Nacional de Matemática Pura e Aplicada (Impa), que sempre me ajudaram de maneira muito rápida e gentil.

## RESUMO

Este trabalho apresenta um manual de produção e revisão de itens de múltipla escolha, principalmente para autores iniciantes. Este manual pode ser usado de maneira complementar a outro guia de caráter mais prescritivo, usualmente fornecido pela instituição que organiza uma avaliação em larga escala ou que demanda itens de múltipla escolha para a formação de um banco.

O objetivo principal do manual é transformar parte do conhecimento tácito em explícito, reduzindo a importância da experiência para produzir itens de qualidade. O manual tem ainda o objetivo específico de dar maior profundidade e rapidez às técnicas de geração de distratores, etapa que costuma ser a parte mais difícil e laboriosa quando se deseja criar itens de múltipla escolha.

**PALAVRAS-CHAVE:** avaliação em larga escala; itens de múltipla escolha.

## SUMÁRIO

1. INTRODUÇÃO .....	7
2. DESCRIÇÃO DO PROBLEMA E ESTABELECIMENTO DE OBJETIVOS... 10	
3. O MANUAL DE PRODUÇÃO E REVISÃO DE ITENS DE MÚLTIPLA ESCOLHA .....	14
4. LIMITAÇÕES.....	17
4.1. LIMITAÇÕES DAS AVALIAÇÕES EM LARGA ESCALA.....	17
4.2. LIMITAÇÕES DESTE TRABALHO.....	25
5. CONCLUSÃO.....	28
6. REFERÊNCIAS BIBLIOGRÁFICAS .....	29
ANEXO I – MANUAL DE PRODUÇÃO E REVISÃO DE ITENS DE MÚLTIPLA ESCOLHA .....	33

## 1. INTRODUÇÃO

Até o menos judicioso dos envolvidos com educação no Brasil é consciente dos principais problemas da área. Ainda que, nas últimas décadas, tenhamos praticamente garantido a universalidade do acesso ao ensino por parte das crianças e dos adolescentes, temos problemas quanto à permanência e, principalmente, quanto à qualidade do ensino e da (não) consecução de objetivos pedagógicos por parte dos estudantes. Resultados de testes nacionais e internacionais revelam repetidamente o desempenho pífio dos estudantes brasileiros. De acordo com o *Programme for International Student Assessment (Pisa)* – Programa Internacional de Avaliação de Estudantes –, uma avaliação trienal de âmbito mundial realizada pela Organização para a Cooperação e Desenvolvimento Econômico e aplicada a estudantes na faixa dos 15 anos, o Brasil foi classificado em 58º lugar em matemática, dentre os 65 países e economias que participaram da avaliação de 2012<sup>1</sup>. Na edição anterior do *PISA*, realizada em 2009, o Brasil havia ficado em 57º lugar em matemática, dentre os 65 participantes<sup>2</sup>.

Sendo a desigualdade social um dos piores indicadores econômicos brasileiros e, concomitantemente, considerando a educação como um fator fundamental para a mobilidade social, nossos problemas educacionais têm dimensão ainda mais relevante.

Neste cenário, muitos opinam, modelos redutores são inventados e reinventados. Teorias se alternam sem sucesso, de modo que o desempenho

---

<sup>1</sup> ORGANISATION FOR ECONOMIC CO-OPERATION AND DEVELOPMENT (OECD). **BRAZIL – Country note – Results from PISA 2012**. [2013?]. p. 9. Disponível em: <<http://www.oecd.org/pisa/keyfindings/PISA-2012-results-brazil.pdf>>. Acesso em 15 fev. 2014.

<sup>2</sup> ORGANISATION FOR ECONOMIC CO-OPERATION AND DEVELOPMENT (OECD). **PISA 2009 results: executive summary**. 2010. p. 8. Disponível em: <<http://www.oecd.org/pisa/pisaproducts/46619703.pdf>>. Acesso em 15 fev. 2014.

dos estudantes brasileiros nos testes internacionais vem se mantendo sofrível, o que compromete o desenvolvimento econômico nacional e afeta as aspirações do País na competição global entre as nações.

Os educadores sabem – ou deveriam saber – que é necessário avaliar constantemente. Sem um diagnóstico adequado, haverá desperdício de tempo e recursos financeiros, talvez os elementos mais valorizados no mundo contemporâneo.

Neste cenário, destaca-se a importância das avaliações em larga escala, que têm por objetivo diagnosticar o estado da educação por meio de testes aplicáveis a grandes contingentes de examinados e, mais recentemente, têm sido cada vez mais usadas no Brasil também como porta de entrada do ensino superior público e gratuito, por meio do Exame Nacional do Ensino Médio (Enem)<sup>3</sup>. Salvo poucas exceções, estas avaliações são integradas de itens (ou seja: questões) de múltipla escolha, já que a apuração de resultados é rápida e levada a cabo por máquinas. Eventualmente há também uma prova de redação em língua portuguesa. Vale dizer que, ainda que avaliações qualitativas possam revelar aspectos importantes dos processos de ensino e aprendizagem, as avaliações em larga escala, que têm caráter quantitativo, não têm substituto conhecido quando se trata de grupos numerosos. Além disso, evita-se a subjetividade, que poderia comprometer o processo.

Entretanto, para o sucesso de uma avaliação em larga escala, é necessário dispor de um grupo de profissionais capazes de atuar nas diversas etapas deste procedimento complexo. Em especial, são requeridos docentes experientes e altamente treinados para o preparo das provas usadas nas

---

<sup>3</sup> “O Exame Nacional do Ensino Médio (ENEM) deve, progressivamente, compor o Sistema de Avaliação da Educação Básica (SAEB), assumindo as funções de: I - avaliação sistêmica, que tem como objetivo subsidiar as políticas públicas para a Educação Básica; II - avaliação certificadora, que proporciona àqueles que estão fora da escola aferir seus conhecimentos construídos em processo de escolarização, assim como os conhecimentos tácitos adquiridos ao longo da vida; III - avaliação classificatória, que contribui para o acesso democrático à Educação Superior”, conforme o artigo 21 da Resolução CNE/CEB nº 2/2012.

avaliações em larga escala: elas são constituídas de itens que demandam certas competências técnicas e pedagógicas muito particulares em sua produção. Os elaboradores dos testes devem não apenas conhecer profundamente os assuntos que são objetos dos testes, mas também é exigida a obediência a um conjunto elevado de normas, bem como criatividade na criação dos itens dos exames.

Apesar disso, uma apreciação rápida de provas aplicadas em concursos públicos revela que ainda há um longo caminho a ser percorrido. O item abaixo, integrante da Prova de Conhecimentos Gerais de Direito do X Concurso para Provimento de Cargos de Juiz do Trabalho Substituto da 12ª Região (Santa Catarina), datada de 2 de setembro de 2000, é um exemplo onde o leitor atento irá encontrar diversos exemplos do que não deve ser feito.

Assinale a alternativa mais correta. São fontes do direito administrativo:

- (A) a lei, a doutrina, a jurisprudência e os costumes;
- (B) a lei, a doutrina e os costumes;
- (C) a lei, os costumes, os contratos administrativos e os tratados;
- (D) a lei, a doutrina, a jurisprudência e os contratos administrativos;
- (E) n. d. a.

Destaca-se, no item acima, a busca pela “alternativa mais correta”, a presença de opções de resposta onde uma contém outra – se a opção A está correta, a opção B também está – e a presença de “nenhuma das anteriores” na opção E, que torna o conjunto de opções de respostas desequilibrado. Ademais, cabe ainda observar que o conteúdo do item é tipicamente livresco, não havendo qualquer indício de criatividade.

## 2. DESCRIÇÃO DO PROBLEMA E ESTABELECIMENTO DE OBJETIVOS

Em se aceitando o diagnóstico apresentado, no sentido de que é necessário ainda desenvolver muito a qualidade da produção de itens de múltipla escolha no Brasil, vale investigar o motivo das dificuldades encontradas na área.

Um aspecto relevante é a carência de material bibliográfico. Faltam, em língua portuguesa, obras que tenham por objetivo treinar profissionais para escrever e revisar itens de múltipla escolha. Existem, todavia, manuais de produção que, muitas vezes, são restritos à metodologia e às normas eleitas pela instituição responsável pela formação do banco de itens, mas não aprofundam a discussão acerca da qualidade dos itens a serem produzidos. Pode ser útil considerar que a situação nos EUA, em 1998, era aproximadamente a mesma, de acordo com Osterlind, que afirmou, à época, que havia abundante bibliografia acerca de teorias psicométricas e de avaliação educacional, como a Teoria Clássica dos Testes (TCT) e a Teoria da Resposta ao Item (TRI), bem como da construção de testes e da análise de resultados de testes, mas havia propriamente pouca informação sobre produção de itens<sup>4</sup>.

Além disso, parece que o treinamento oferecido aos produtores em geral não consegue dar conta dos principais problemas que os autores iniciantes enfrentam. O Guia de Elaboração de Itens preparado pelo Centro de Políticas Públicas e Avaliação da Educação da Universidade Federal de Juiz de Fora enumera algumas características necessárias ao produtor de itens: a) o domínio da área de conhecimento a ser avaliada; b) o entendimento acerca dos processos de desenvolvimento e aprendizagem que caracterizam os estudantes para os quais os itens serão produzidos; c) o domínio da linguagem

---

<sup>4</sup> OSTERLIND, Steven J. **Constructing test items: multiple-choice, constructed-response, performance and other formats** (Evaluation in education and human services). 2. ed. Boston, MA: Kluwer Academic Publishers, 1998. p. 3.

verbal utilizada pelos estudantes para quem o teste será construído; e d) a habilidade de utilizar as técnicas de escrever itens. Após listar essas desejáveis características, o guia acrescenta o que é reproduzido a seguir.

A excelência na elaboração de itens, contudo, demanda mais do que isso. É preciso imaginação e criatividade na invenção de situações que exijam o conhecimento e as habilidades desejadas. Demanda, principalmente, habilidade e julgamento que só vêm com experiência<sup>5</sup>.

Ou seja, o guia declara explicitamente que, para a produção de itens de múltipla escolha de qualidade, são necessárias certas competências “que só vêm com experiência”.

A pergunta que surge é: **seria possível produzir material didático para produtores de itens de múltipla escolha iniciantes que reduzam a dependência da experiência?** Ou ainda: **cabe pensar em estratégias de treinamento exequíveis que aumentem a qualidade dos itens produzidos quando seus autores ainda não têm muita experiência?**

Para tratar do problema apresentado, conversei pessoalmente ou por correio eletrônico com algumas dezenas de produtores, alguns iniciantes, outros com diversos períodos de experiência na produção e revisão de itens de múltipla escolha. Abordei especialistas da área e participei de meia dúzia de treinamentos. Ao longo de cerca de dezoito meses, escrevi profissionalmente cerca de 350 itens – parte para o projeto “Máquina de Testes” da Secretaria Municipal de Educação do Rio de Janeiro, apoiado pelo Programa Nacional de Tecnologia Educacional (ProInfo) e voltado para o nível fundamental; parte para o “Programa FGV Ensino Médio Digital”, da Fundação Getúlio Vargas,

---

<sup>5</sup> Centro de Políticas Públicas e Avaliação da Educação / Universidade Federal de Juiz de Fora (CAEd/UFJF). **Guia de elaboração de itens**. Juiz de Fora, MG: 2008. p. 18. Disponível em: <[http://www.portalavaliacao.caedufjf.net/wp-content/uploads/2012/02/Guia\\_De\\_Elabora%C3%A7%C3%A3o\\_De\\_Itens\\_MT.pdf](http://www.portalavaliacao.caedufjf.net/wp-content/uploads/2012/02/Guia_De_Elabora%C3%A7%C3%A3o_De_Itens_MT.pdf)>. Acesso em: 29 jan. 2014.

voltado para o nível médio. Durante o trabalho de produção, atentei para dificuldades que encontrei e documentei os erros cometidos, refletindo até mesmo sobre possíveis falhas que poderiam ter sido cometidas, mas foram evitadas. Avaliei alguns guias de elaboração de itens, buscando identificar práticas oriundas da experiência ou relatadas oralmente em treinamentos, mas não integrantes dos mencionados guias. Analisei as referências bibliográficas destes guias, registrando as obras consultadas. Finalmente, no exame destas fontes primárias, basicamente de origem estadunidense, confirmei a incipiência dos referenciais teóricos nacionais.

Neste sentido, um dos meus interlocutores, um professor com 18 anos de experiência na produção de itens<sup>6</sup>, instado a dizer em que medida o manual e a documentação que recebeu foram suficientes para o preparo dos itens, quando do início de seu trabalho de produção de itens em matemática, respondeu: “ajudaram, mas principalmente as conversas com pessoas mais experientes foram fundamentais”. O mesmo professor, interrogado se os manuais e a documentação que ele conhecia, relativos à elaboração e à revisão de itens em matemática, podiam ser considerados adequados para tais tarefas, declarou que “são pouco adequados”, tendo acrescentado: “considero que é um campo fértil para novas publicações”.

Outro autor, que preferiu não se identificar, com experiência de três anos em produção de itens, declarou que, “sem dúvida alguma”, um manual de elaboração e revisão de itens deve ser “longo e apresentar também exemplos e ‘dicas’ práticas para a preparação de itens”. Este professor acrescentou: “são exemplos claros que fazem os futuros elaboradores entenderem o que a habilidade [ou seja, o descritor] deseja aferir”.

Ao término desta pesquisa, conclui que seria possível estabelecer o seguinte objetivo principal: **criar um manual de produção e revisão de itens de**

---

<sup>6</sup> Francisco Linhares, professor e coordenador da Escola Alemã Corcovado. Elaborador de itens para vários concursos públicos e exames vestibulares.

**múltipla escolha que tivesse como objetivo transformar parte do conhecimento tácito – aquele referido no Guia de Elaboração de Itens preparado pelo Centro de Políticas Públicas e Avaliação da Educação da Universidade Federal de Juiz de Fora como oriundo da experiência – em explícito.**

### **3. O MANUAL DE PRODUÇÃO E REVISÃO DE ITENS DE MÚLTIPLA ESCOLHA**

Considerando a natureza do curso de mestrado a que este trabalho de conclusão de curso se associa – ele é profissional; não é acadêmico – considere que seria válido gerar um produto de possível utilização prática como consequência do curso. Concluí que uma boa opção seria construir um guia para se escrever itens de múltipla escolha. Assim, este trabalho de conclusão de curso é composto por esta parte mais curta, que fundamenta e discorre sobre uma parte mais longa, apresentada como anexo, que tem a pretensão de ser um manual de produção e revisão de itens de múltipla escolha voltado para produtores iniciantes.

Um problema que enfrentei no início do trabalho consistiu em constatar que as normas e orientações pedagógicas defendidas nos diversos manuais que examinei não são sempre as mesmas. Em verdade, pode existir uma completa heterogeneidade entre elas. Apresento um par de exemplos a seguir. Encontrei manuais que sugerem a inclusão de imagens de cunho ilustrativo nos itens, com o objetivo de torná-los mais atraentes aos examinados. Entretanto, outros guias exigem que somente figuras necessárias para a solução do problema veiculado pelo item devem ser inseridas, já que imagens ornamentais somente teriam como consequência o desvio de atenção do examinado da atividade que deveria desempenhar durante a busca da opção de resposta correta do item. Outro aspecto polêmico é a contextualização dos itens: alguns manuais exigem que todos os itens produzidos sejam baseados em situações-problema típicas do cotidiano; outros guias prescrevem que isto deve ser perseguido, admitindo que nem sempre será possível atingir tal objetivo; em outros manuais, por fim, sequer se menciona o aspecto da contextualização.

Diante destas disparidades, optei por produzir um manual que não valorizasse certas abordagens em detrimento de outras, mas que alertasse o leitor a

respeito das divergências e dúvidas oriundas das diversas orientações, buscando cobrir uma maior quantidade de cenários, sem se comprometer com escolhas inflexíveis. Uma consequência de minha opção é que, em linhas gerais, foi valorizado mais o conteúdo do que a forma dos itens a serem produzidos. O manual apresentado como anexo, portanto, deve ser usado de maneira complementar a outro manual de caráter mais prescritivo, usualmente fornecido pela instituição que encomenda do autor a produção de itens de múltipla escolha.

Vale ainda dizer que o manual é destinado à produção e à revisão de itens de múltipla escolha, especialmente quando produzidos para a formação de um banco de itens. Mas penso que a leitura do manual pode também ser útil para tarefas menos complexas, até mesmo para a confecção de provas aplicáveis por um professor às suas turmas de estudantes.

O manual apresenta inicialmente, em linhas gerais, as etapas que integram a construção de um banco de itens, acompanhadas de alguns comentários que tratam das avaliações em larga escala e da Teoria da Resposta ao Item (TRI). Ainda que esta parte do manual não descreva detalhadamente as diversas fases de criação e manutenção de um banco de itens, acredito que é necessário ao produtor iniciante conhecer minimamente como seu trabalho é inserido em um cenário mais amplo.

A parte mais importante do manual descreve as diversas partes de um item, com observações detalhadas tratando de cada uma delas. Estes comentários acerca de cada campo do formulário de um item são algumas vezes acompanhados de observações críticas que têm por objetivo tratar as ponderações mais comuns de um produtor iniciante. Sempre que julguei necessário, incluí itens – todos de autoria própria – que exemplificam os assuntos tratados ou, em alguns casos, apresentam erros que devem ser evitados.

O manual tem também o objetivo específico de dar maior profundidade e rapidez às técnicas de geração de distratores, etapa que costuma ser a parte mais difícil e laboriosa quando se deseja criar itens de múltipla escolha<sup>7</sup>. Assim, são apresentados comentários detalhados acerca da natureza e do objetivo dos distratores, bem como são mostradas técnicas para sua geração com o uso de planilhas eletrônicas.

O manual inclui ainda, em forma de tópicos de rápida leitura, as orientações práticas mais importantes para a produção de itens de múltipla escolha. Neste aspecto foram escolhidas as recomendações mais relevantes, cuja inobservância é causa do maior número de falhas na produção.

Finalmente, o manual apresenta comentários à revisão de itens de múltipla escolha.

---

<sup>7</sup> HALADYNA, Thomas M. **Developing and validating multiple-choice test items**. 3. ed. Mahwah, NJ: Lawrence Erlbaum Associates, 2004. p. 120.

## **4. LIMITAÇÕES**

Este trabalho não é um panfleto em defesa das avaliações em larga escala baseada em itens de múltipla escolha. Em nome do contraditório e da atividade crítica que deve permear qualquer atividade pedagógica, apresento a seguir alguns problemas que devem ser considerados.

Em primeiro lugar, seguem considerações acerca das limitações das avaliações em larga escala. Em seguida, são apresentadas limitações deste trabalho.

### **4.1. LIMITAÇÕES DAS AVALIAÇÕES EM LARGA ESCALA**

O processo de ensino e aprendizagem é muito complexo. Qualquer análise que dispensa peso exagerado a algum aspecto em detrimento de outros está provavelmente equivocada. Assim, a utilização de avaliações em larga escala, apesar de trazer incontestáveis benefícios para o processo de diagnóstico da educação, principalmente em um país de elevada população e com muitos problemas, não deve ser supervalorizada. A avaliação em larga escala é inerentemente limitada. Ela se limitará a gerar relatórios, que podem ou não estar adequados. Estes relatórios devem ser produzidos por pessoal que detenha conhecimentos estatísticos e pedagógicos, perfil que não é tão fácil de ser encontrado. Não se pode deixar de considerar que existe ainda o risco de relatórios serem escritos com objetivos ideológicos ou, ainda, econômicos, quando em busca de justificativa para eventuais financiamentos. Ainda se os relatórios estiverem adequados, evidentemente eles não têm

autoexecutoriedade e dependem do estabelecimento de políticas públicas responsáveis e contínuas ao longo dos governos. Destaca-se ainda o risco de se produzir conclusões superficiais, considerando as pontuações calculadas a partir das avaliações em larga escala como seu único resultado, simplificando sobremaneira o processo de avaliação. Um indicador é somente uma medição, que é bem menos que uma avaliação.

Tratando ainda da supervalorização dos indicadores, esta convicção pode levar gestores, professores e estudantes a procurar inflar as pontuações dos testes artificialmente, em busca de resultados maquiados. Estas práticas podem ser: fornecer respostas ou dicas aos estudantes enquanto o teste é aplicado, permitir aos estudantes a mudança das respostas após o teste ter sido terminado, mudar as respostas para eles, fornecer os itens previamente e assim por diante<sup>8</sup>. Esses métodos reprováveis se tornam ainda mais perigosos quando se trata de concursos públicos, que eventualmente selecionam pessoas para cargos bem remunerados, de modo que a segurança se torna um elemento indispensável do processo de teste.

Outra consequência que merece destaque é a supervalorização dos descritores que compõem as matrizes de referência que normalmente são estruturadas para a aplicação de avaliações em larga escala. Esse processo dá origem a novas orientações dos currículos das escolas, especialmente quando se considera a necessidade de ingresso nas instituições de ensino superior mantidas pelo poder público. Dramaticamente, muitas escolas se transformam em academias preparatórias para o Exame Nacional do Ensino Médio (Enem). Mas não são os descritores das avaliações em larga escala que devem ditar os conteúdos dos currículos. São as Diretrizes Curriculares Nacionais para o Ensino Médio que “devem nortear a elaboração da proposta de expectativas de aprendizagem, a formação de professores, os investimentos em materiais didáticos e os sistemas e exames nacionais de avaliação”, conforme o artigo 22

da Resolução CNE/CEB nº 2/2012<sup>9</sup>. Neste sentido, o Guia de Elaboração de Itens preparado pelo Centro de Políticas Públicas e Avaliação da Educação da Universidade Federal de Juiz de Fora realça lucidamente a limitação dos descritores típicos das avaliações em larga escala, como reproduzido a seguir.

É preciso enfatizar que os descritores não podem ser adotados como um conjunto de indicações básicas para as práticas de ensino-aprendizagem nas escolas, uma vez que não contêm a análise do conhecimento da matemática, as orientações didáticas, estratégias e recursos didáticos, as sugestões de como trabalhar os conteúdos, bem como não selecionam a progressão de conteúdos por ano ou ciclos. Esse tipo de orientação cabe às Diretrizes, Parâmetros e Matrizes Curriculares. Aos descritores, cabe, apenas, a referência para a elaboração dos itens que compõem os testes<sup>10</sup>.

Na esteira desta interpretação, acrescenta-se a ponderação feita pelo Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (Inep), tratando do Exame Nacional do Ensino Médio (Enem) e salientando a limitação das matrizes de referência – conjuntos de descritores – em relação aos currículos e dos indicadores e escalas em relação ao processo de ensino e aprendizagem:

Do mesmo modo que as matrizes de referência não se confundem com o currículo, que é muito mais amplo, a interpretação pedagógica das escalas representa um recorte

---

<sup>8</sup> KORETZ, Daniel. **Measuring up**: what educational testing really tells us. Cambridge, MA: Harvard University Press, 2008. p. 252.

<sup>9</sup> BRASIL. Ministério da Educação. Resolução CNE/CEB nº 2, de 30 de janeiro de 2012. Define Diretrizes Curriculares Nacionais para o Ensino Médio. **Diário Oficial da União**, Brasília, DF, 31 jan. 2012. Disponível em: <[http://portal.mec.gov.br/index.php?option=com\\_docman&task=doc\\_download&gid=9864](http://portal.mec.gov.br/index.php?option=com_docman&task=doc_download&gid=9864)>. Acesso em: 15 fev. 2014.

<sup>10</sup> Centro de Políticas Públicas e Avaliação da Educação / Universidade Federal de Juiz de Fora (CAEd/UFJF). **Guia de elaboração de itens**. Juiz de Fora, MG: 2008. p. 14. Disponível em: <[http://www.portalavaliacao.caedufjf.net/wp-content/uploads/2012/02/Guia\\_De\\_Elabora%C3%A7%C3%A3o\\_De\\_Itens\\_MT.pdf](http://www.portalavaliacao.caedufjf.net/wp-content/uploads/2012/02/Guia_De_Elabora%C3%A7%C3%A3o_De_Itens_MT.pdf)>. Acesso em: 29 jan. 2014.

que tomou por base um conjunto de itens aplicados no Enem. Esses itens foram elaborados tendo como ponto de partida as matrizes de referência e não esgotam as possibilidades de abordagens de suas habilidades<sup>11</sup>.

Em se permitindo que os descritores orientem o que é ensinado nas escolas, estar-se-ia criando um círculo vicioso. Vejamos um exemplo. Supondo que o entendimento dos logaritmos seja um ponto crítico na aprendizagem matemática dos estudantes, a consequência esperada é um baixo desempenho em itens que tratem deste conteúdo. Pode ser, então, que muitos itens que abordem logaritmos sejam eliminados após o pré-teste, dada sua elevada dificuldade. Isso provocaria a existência de provas que não versassem sobre o tema. Por fim, os estudantes poderiam deixar de estudar o assunto, diante da justificativa de que não são tratados nos exames.

Ou seja, se o nível de dificuldade dos itens integrantes das provas é definido a partir do resultado dos pré-testes e se os examinados nesta etapa são estudantes que não dominam determinadas habilidades e, em associação a este cenário, os estudantes que efetivamente fazem as avaliações em larga escala são orientados a partir dos conteúdos dos descritores destes mesmos testes, não é razoável esperar algum tipo de incremento no desempenho dos discentes.

É conveniente acrescentar outra análise restritiva quanto à natureza das avaliações em larga escala apresentada pelo Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (Inep), quando tratando da inclusão de ciências no Sistema de Avaliação da Educação Básica (Saeb), que pode ser estendida, sem dúvida, para as demais áreas do conhecimento. O instituto ponderou que

---

<sup>11</sup> BRASIL. Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (Inep). **Interpretação pedagógica das escalas de proficiência – Exame Nacional do Ensino Médio – Enem**. Brasília, DF: MEC/Inep, [2012?]. Disponível em: <<http://mapaitensenem.inep.gov.br/mapaNota/home.seam>>. Acesso em: 02 fev. 2014.

o alcance possível para uma medida por meio de um instrumento do tipo prova a ser aplicado em larga escala é impeditivo de uma pretensão de medir o sentido pleno do uso social do conhecimento que cria a própria realidade quando em uso. Há a necessidade de reconhecer e explicitar que a medida será possível apenas sobre alguns aspectos desse uso social, como os da alfabetização emancipadora ou do letramento em ciências como uso social do conhecimento, pois a característica do instrumento impede a pretensão de medi-lo em sentido *lato*. O instrumento a ser construído para medição desse desempenho no uso da linguagem científica será delimitado sobre aqueles aspectos relativos à alfabetização/letramento escolar em ciências, passíveis de serem medidos por itens/questões objetivas de múltipla escolha, compondo um conjunto de itens que consegue cobrir apenas uma parcela dos componentes curriculares<sup>12</sup>.

Não se deve, portanto, utilizar as avaliações para finalidade diversa daquela para a qual foram talhadas: elas não são orientadoras de currículos. Se esse erro for cometido, não se estaria preparando cidadãos vinculados ao mundo do trabalho e à prática social, como prescreve a lei que estabelece as diretrizes e bases da educação nacional<sup>13</sup>, mas apenas adestrando estudantes para dar respostas corretas para itens de múltipla escolha. O ensino se tornaria muito mais mecânico e os currículos seriam estreitados.

É necessário tratar também da meritocracia e de sua associação com as avaliações em larga escala. Ainda que esteja muito além do escopo deste

---

<sup>12</sup> BRASIL. Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (Inep). **Inclusão de ciências no Saeb**: documento básico. Brasília, DF: MEC/Inep, 2013. p. 25-26. Disponível em: <[http://download.inep.gov.br/educacao\\_basica/prova\\_brasil\\_saeb/menu\\_do\\_professor/matrizes\\_de\\_referencia/livreto\\_saeb\\_ciencias.pdf](http://download.inep.gov.br/educacao_basica/prova_brasil_saeb/menu_do_professor/matrizes_de_referencia/livreto_saeb_ciencias.pdf)>. Acesso em: 15 fev. 2014.

<sup>13</sup> BRASIL. Lei nº 9.394/96, de 20 de dezembro de 1996. Estabelece as diretrizes e bases da educação nacional. **Diário Oficial da União**, Brasília, DF, 23 dez. 1996. Disponível em: <[http://www.planalto.gov.br/ccivil\\_03/leis/l9394.htm](http://www.planalto.gov.br/ccivil_03/leis/l9394.htm)>. Acesso em: 31 jan. 2014.

trabalho discorrer em pormenores sobre isto, a inconveniência residente no uso de avaliações em larga escala para determinar o futuro de milhões de jovens brasileiros é um aspecto que não pode ser aqui ignorado e merece, ao menos, brevíssimos comentários. O Exame Nacional do Ensino Médio (Enem) foi inicialmente pensado como um instrumento para auxiliar o Ministério da Educação na definição de políticas públicas com o objetivo de melhorar a qualidade do Ensino Básico. Entretanto, o exame é utilizado desde 2009 também para selecionar os que ingressam nas universidades públicas. Essa função do exame pode ser reprodutora de uma má distribuição de oportunidades dentre os jovens brasileiros que, se perpetuada, pode limitar sobremaneira a mobilidade social que a educação poderia operar. Assim, em lugar de se usar uma avaliação em larga escala para fins de diagnóstico, tornando possível melhorar a educação para a comunidade, estar-se-ia utilizando o Exame Nacional do Ensino Médio (Enem) principalmente para convalidar a distribuição de vagas de ensino superior para uma elite privilegiada, garantindo seu lugar nas estruturas de poder e prestígio. O problema é de difícil solução e não é restrito à realidade nacional, sendo também objeto de acirradas discussões nos EUA. O professor Nicholas Berthelot Lemann denuncia que a meritocracia estadunidense, em verdade, se caracteriza principalmente por formar uma elite escolhida segundo critérios acadêmicos com o objetivo de governar a nação. Essa minoria privilegiada é educada por meio de recursos públicos e, em tese, deve dar o retorno do investimento por meio de sua atuação no serviço público. Todo o sistema é corroborado por avaliações em larga escala, que orientam o ingresso dos estudantes nas universidades<sup>14</sup>. Não é difícil encontrar semelhança com o caso brasileiro. Na mesma linha de denúncia contra a meritocracia estadunidense, o economista e crítico social Peter Sacks aponta que a ocupação profissional do

---

<sup>14</sup> LEMANN, Nicholas Berthelot. **The big test**: the secret history of the American meritocracy. New York, NY: Farrar, Straus and Giroux, 2000. p. 343.

pai de um estudante é um indicador da sua pontuação no SAT<sup>15</sup> melhor do que praticamente qualquer outro fator<sup>16</sup>. Ele salienta também que os interesses financeiros em torno da indústria de testes estadunidense são muito grandes, tendo esta se tornado um grande negócio baseado em tecnologia que ninguém realmente sabe dizer o tamanho<sup>17</sup>.

Acrescenta-se ainda que o principal suposto teórico da Teoria da Resposta ao Item (TRI) e dos modelos logísticos que têm sido usados amplamente nas avaliações em larga escala é a unidimensionalidade, o que corresponde a dizer que

toda a complexidade intrínseca ao ato de resolução de um problema – de natureza cognitiva ou não – deve ter como causa uma única estrutura latente, denominada  $\theta$ . Assim, existirá uma relação funcional entre  $\theta$  e os padrões das respostas dadas a um problema, sendo essa a ideia fundamental dos (...) modelos logísticos de TRI (...)<sup>18</sup>

Essa estrutura latente  $\theta$  seria a proficiência do avaliado. A crítica que pode ser apresentada a este suposto teórico é que a proficiência – ou a inteligência – não teria uma dimensão única, sendo, em verdade, bem mais complexa. Assim, os modelos estatísticos não teriam o condão de realizar sua tarefa no sentido de medir adequadamente a proficiência. Neste sentido, é bom

---

<sup>15</sup> O *Scholastic Aptitude Test – SAT* é um exame educacional utilizado nos EUA para admissão dos estudantes nas universidades, semelhante ao brasileiro Exame Nacional do Ensino Médio (Enem).

<sup>16</sup> SACKS, Peter. **Standardized minds**: the high price of America's testing culture and what we can do to change it. Cambridge, MA: Perseus Publishing, 1999. p. 2.

<sup>17</sup> SACKS, Peter. **Standardized minds**: the high price of America's testing culture and what we can do to change it. Cambridge, MA: Perseus Publishing, 1999. p. 222.

<sup>18</sup> ANDRIOLA, Wagner Bandeira. Psicometria moderna: características e tendências. **Estudos em avaliação educacional**, São Paulo, v. 20, n. 43, maio/ago. 2009. p. 327. Disponível em: <<http://www.fcc.org.br/pesquisa/publicacoes/eae/arquivos/1496/1496.pdf>>. Acesso em: 06 fev. 2014.

considerar a Teoria das Inteligências Múltiplas de Howard Gardner, psicólogo e educador estadunidense, que pondera o seguinte:

Parece-me, porém, estar cada vez mais difícil negar a convicção de que há pelo menos algumas inteligências, que estas são relativamente independentes umas das outras e que podem ser modeladas e combinadas numa multiplicidade de maneiras adaptativas por indivíduos e culturas<sup>19</sup>.

Por fim, vale ainda mencionar que não existe teoria sólida que explique a cognição humana na profundidade necessária que permita se desenvolver instrumentos de medida adequados. Falta consenso entre pesquisadores no que se refere à maneira que uma pessoa mobiliza seu repertório de competências e conhecimentos para resolver problemas culturalmente relevantes. Assim, é certo que métodos que buscam medir grandezas de algum modo associadas à cognição estão sujeitos a críticas, não podendo se falar em validação perfeita deste processo.

Apesar das limitações apresentadas, é importante salientar que as avaliações em larga escala, principalmente quando acompanhadas de métodos estatísticos contemporâneos e constituídas de itens escritos mediante padrões técnicos e pedagógicos cuidadosamente estabelecidos, são indiscutivelmente melhores do que simplesmente ignorar os problemas educacionais nacionais. A ausência de avaliação corresponderia a navegar sem rumo. Assim, ainda que se rejeite qualquer panaceia em relação às avaliações em larga escala, também não se admite uma perspectiva ideologizada que abomina toda forma de avaliação quantitativa, sem ponderar sobre sua aplicabilidade.

---

<sup>19</sup> GARDNER, Howard. **Estruturas da mente: a teoria das inteligências múltiplas**. Porto Alegre: Artes Médicas Sul, 1994. p. 7.

## 4.2. LIMITAÇÕES DESTE TRABALHO

No que tange especificamente ao produto deste trabalho, ou seja, o manual de produção e revisão de itens de múltipla escolha que consta como anexo, não é possível esperar que o treinamento presencial e interativo, necessário em qualquer processo de formação de bons produtores de itens, seja omitido. Não cabe também desejar que o manual anexo dê conta de todas as dúvidas que podem surgir em um produtor iniciante. Entretanto, como já declarado, é desejável que parte do conhecimento tácito tenha sido transformada em conhecimento explícito e documentada no manual.

Vale ainda enfatizar que a qualidade dos itens produzidos depende de duas atividades complementares indispensáveis: os procedimentos de revisão de itens e os estudos estatísticos das respostas fornecidas pelos avaliados<sup>20</sup>. Este trabalho pretende cobrir a primeira atividade mencionada ao menos parcialmente, já que o manual anexo conta com uma seção dedicada à revisão de itens de múltipla escolha, ainda que foco do aludido guia seja a produção de itens. Entretanto, em função do recorte feito por este trabalho, a segunda atividade mencionada, ou seja, o processo de validação dos itens e, de maneira mais abrangente, dos exames tomados em sua integralidade não será aqui tratado. Esta atividade consiste em avaliar em que medida o teste mede com precisão o que se pretende medir. Isto está profundamente ligado à interpretação e ao uso dos resultados obtidos após a aplicação dos testes, etapa que normalmente é levada a cabo por especialistas. Problemas relevantes podem ocorrer se a etapa de validação for negligenciada. Por exemplo, devido à pressão por melhores resultados, podem ser envidados esforços extremos, por parte das escolas, no sentido de melhorar as

---

<sup>20</sup> HALADYNA, Thomas M. **Developing and validating multiple-choice test items**. 3. ed. Mahwah, NJ: Lawrence Erlbaum Associates, 2004. p. 202.

pontuações obtidas nos testes<sup>21</sup>. Entretanto, pode ser que essas táticas não provoquem necessariamente a melhora da proficiência dos estudantes, ainda que consigam inflar as pontuações. Particularmente, isto se dá quando os estudantes desenvolvem técnicas específicas para achar a opção de resposta correta, mas não necessariamente aumentam seu desempenho no que se refere às habilidades avaliadas pelos itens. Ou seja, se um teste não for adequadamente validado, pode resultar em medições equivocadas que, no caso particular apresentado, poderiam se tornar piores no decorrer do tempo, na medida em que mais esforços artificiais fossem aplicados na melhoria das pontuações. Portanto, o que se enfatiza é que o processo de avaliação em larga escala é incompleto se não houver uma preocupação em se averiguar, detalhada e tecnicamente, a efetividade com a qual os testes medem com acuidade aquilo que se deseja que seja medido. De todo modo, é incontestável que a produção de itens de qualidade contribui para uma melhor adequação na medição pretendida, o que significa dizer que, mesmo sem a preocupação específica com a etapa de validação, o manual anexo contribui para seu sucesso.

Outra limitação que também deve ser mencionada é que o manual apresentado em anexo não está assentado em um corpo teórico abrangente. Não há, de fato, uma preocupação contundente em dar conta das inevitáveis relações entre o processo de ensino e aprendizagem e a produção de itens. Em especial, ainda prevalece, em certos aspectos, uma perspectiva de cunho próximo da abordagem comportamentalista da aprendizagem, com características instrucionais. Mas vale dizer que apresentar algo muito diferente disso é ousado até mesmo no âmbito internacional, pois uma mudança relevante nesta área é apresentada como um desafio para o futuro<sup>22</sup>. Todavia, paradoxalmente, pode ser adequado produzir um manual nos moldes do

---

<sup>21</sup> HALADYNA, Thomas M. **Developing and validating multiple-choice test items**. 3. ed. Mahwah, NJ: Lawrence Erlbaum Associates, 2004. p. 9.

<sup>22</sup> HALADYNA, Thomas M. **Developing and validating multiple-choice test items**. 3. ed. Mahwah, NJ: Lawrence Erlbaum Associates, 2004. p. 266.

anexo, já que o estado da produção de itens no Brasil carece ainda de muito progresso e, assim, talvez fosse temerário avançar demais, se até mesmo nos EUA, país conhecido pelo desenvolvimento na área de testes, ainda não há convergência sobre o que está por vir na produção e revisão de itens. De todo modo, pelo menos evitei apresentar simplesmente uma lista de coisas que devem ser feitas e coisas que devem ser evitadas na produção e revisão de itens de múltipla escolha. É bom observar que alguns dos guias que são usados no Brasil não ficam muito longe disso.

Outro aspecto importante a ser realçado é que o manual apresentado como anexo deste trabalho, ainda que construído a partir de minhas experiências práticas de produção de itens e após entrevistas com outros autores, não foi testado, já que não foi aplicado a qualquer plano de produção e revisão de itens de múltipla escolha. Ou seja, não há qualquer retorno sobre a sua aplicação prática, de modo que questões acerca da extensão do texto, de sua profundidade ou, ainda, do teor de algumas das sugestões apresentadas jamais foram indagadas. Eventualmente, isso pode ser corrigido em trabalhos futuros.

## 5. CONCLUSÃO

O curso de mestrado profissional tem algumas características que o diferenciam do mestrado acadêmico. Uma delas diz respeito ao trabalho de conclusão de curso, que pode ser apresentado em variadas formas<sup>23</sup>.

Neste trabalho, procurei escrever um manual de produção e revisão de itens de múltipla escolha, com o objetivo de gerar um guia com potencial de ser efetivamente lido e aplicado pelos professores que desejam produzir itens de múltipla escolha de qualidade, especialmente aqueles que têm pouca ou nenhuma experiência na área.

Considerando que vários dos conteúdos integrantes do manual anexo não constam dos manuais que examinei, acredito que, ao menos parcialmente, consegui atingir o objetivo proposto, ou seja, produzir material didático para produtores de itens de múltipla escolha iniciantes que minimize a dependência da experiência e torne possível um aumento na qualidade dos itens produzidos. Uma eventual aplicação efetiva do manual ao trabalho de produtores de itens poderia esclarecer em que medida o objetivo foi alcançado.

---

<sup>23</sup> Conforme o § 3º do artigo 7º da Portaria Normativa Nº 17, de 28 de dezembro de 2009 do Ministério da Educação, que dispõe sobre o mestrado profissional no âmbito da Fundação Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – CAPES: “o trabalho de conclusão final do curso poderá ser apresentado em diferentes formatos, tais como dissertação, revisão sistemática e aprofundada da literatura, artigo, patente, registros de propriedade intelectual, projetos técnicos, publicações tecnológicas; **desenvolvimento de aplicativos, de materiais didáticos e instrucionais e de produtos, processos e técnicas**; produção de programas de mídia, editoria, composições, concertos, relatórios finais de pesquisa, softwares, estudos de caso, relatório técnico com regras de sigilo, manual de operação técnica, protocolo experimental ou de aplicação em serviços, proposta de intervenção em procedimentos clínicos ou de serviço pertinente, projeto de aplicação ou adequação tecnológica, protótipos para desenvolvimento ou produção de instrumentos, equipamentos e kits, projetos de inovação tecnológica, produção artística, sem prejuízo de outros formatos, de acordo com a natureza da área e a finalidade do curso, desde que previamente propostos e aprovados pela CAPES” (grifei).

## 6. REFERÊNCIAS BIBLIOGRÁFICAS

ANDRIOLA, Wagner Bandeira. Psicometria moderna: características e tendências. **Estudos em avaliação educacional**, São Paulo, v. 20, n. 43, maio/ago. 2009. p. 319-340. Disponível em: <<http://www.fcc.org.br/pesquisa/publicacoes/eae/arquivos/1496/1496.pdf>>. Acesso em: 06 fev. 2014.

BRASIL. Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (Inep). **Entenda a sua nota no Enem – Guia do participante**. Brasília, DF: MEC/Inep, 2012. Disponível em: <<http://mapaitensenem.inep.gov.br/mapaNota/home.seam>>. Acesso em: 02 fev. 2014.

\_\_\_\_\_. Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (Inep). **Inclusão de ciências no Saeb**: documento básico. Brasília, DF: MEC/Inep, 2013. Disponível em: <[http://download.inep.gov.br/educacao\\_basica/prova\\_brasil\\_saeb/menu\\_do\\_professor/matrizes\\_de\\_referencia/livreto\\_saeb\\_ciencias.pdf](http://download.inep.gov.br/educacao_basica/prova_brasil_saeb/menu_do_professor/matrizes_de_referencia/livreto_saeb_ciencias.pdf)>. Acesso em: 15 fev. 2014.

\_\_\_\_\_. Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (Inep). **Interpretação pedagógica das escalas de proficiência – Exame Nacional do Ensino Médio – Enem**. Brasília, DF: MEC/Inep, [2012?]. Disponível em: <<http://mapaitensenem.inep.gov.br/mapaNota/home.seam>>. Acesso em: 02 fev. 2014.

\_\_\_\_\_. Lei nº 9.394/96, de 20 de dezembro de 1996. Estabelece as diretrizes e bases da educação nacional. **Diário Oficial da União**, Brasília, DF, 23 dez. 1996. Disponível em: <[http://www.planalto.gov.br/ccivil\\_03/leis/l9394.htm](http://www.planalto.gov.br/ccivil_03/leis/l9394.htm)>. Acesso em: 31 jan. 2014.

\_\_\_\_\_. Lei Nº 10.741, de 1º de outubro de 2003. Dispõe sobre o Estatuto do Idoso e dá outras providências. **Diário Oficial da União**, Brasília, DF, 03 out. 2003. Disponível em:<[http://www.planalto.gov.br/ccivil\\_03/leis/2003/l10.741.htm](http://www.planalto.gov.br/ccivil_03/leis/2003/l10.741.htm)>. Acesso em: 01 fev. 2014.

\_\_\_\_\_. Ministério da Educação. Portaria Normativa Nº 17, de 28 de dezembro de 2009. Dispõe sobre o mestrado profissional no âmbito da Fundação Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – CAPES. **Diário Oficial da União**, Brasília, DF, 29 dez. 2009. Disponível em:<[http://www.profmat-sbm.org.br/docs/Portaria\\_Normativa\\_MEC17\\_28\\_12\\_2009.pdf](http://www.profmat-sbm.org.br/docs/Portaria_Normativa_MEC17_28_12_2009.pdf)>. Acesso em: 31 jan. 2014.

\_\_\_\_\_. Ministério da Educação. Resolução CNE/CEB nº 2, de 30 de janeiro de 2012. Define Diretrizes Curriculares Nacionais para o Ensino Médio. **Diário Oficial da União**, Brasília, DF, 31 jan. 2012. Disponível em: <[http://portal.mec.gov.br/index.php?option=com\\_docman&task=doc\\_download&gid=9864](http://portal.mec.gov.br/index.php?option=com_docman&task=doc_download&gid=9864)>. Acesso em: 15 fev. 2014.

Centro de Políticas Públicas e Avaliação da Educação / Universidade Federal de Juiz de Fora (CAEd/UFJF). **Guia de elaboração de itens**. Juiz de Fora, MG: 2008. Disponível em: <[http://www.portalavaliacao.caedufjf.net/wp-content/uploads/2012/02/Guia\\_De\\_Elabora%C3%A7%C3%A3o\\_De\\_Itens\\_MT.pdf](http://www.portalavaliacao.caedufjf.net/wp-content/uploads/2012/02/Guia_De_Elabora%C3%A7%C3%A3o_De_Itens_MT.pdf)>. Acesso em: 29 jan. 2014.

FONTANIVE, Nilma Santos; KLEIN, Ruben. Avaliação em larga escala: uma proposta inovadora. **Em aberto**, Brasília, v. 15, n. 66, abr./jun. 1995. p. 29-34. Disponível em: <<http://emaberto.inep.gov.br/index.php/emaberto/article/viewFile/995/899>>. Acesso em: 05 fev. 2014.

GARBI, Gilberto Geraldo. **A rainha das ciências**: um passeio histórico pelo maravilhoso mundo da matemática. 3. ed. rev. e ampl. São Paulo, SP: Livraria da Física, 2009.

GARDNER, Howard. **Estruturas da mente**: a teoria das inteligências múltiplas. Porto Alegre: Artes Médicas Sul, 1994.

HALADYNA, Thomas M. **Developing and validating multiple-choice test items**. 3. ed. Mahwah, NJ: Lawrence Erlbaum Associates, 2004.

HONORATO, Renata. **Alunos do Colégio Christus terão de refazer o Enem 2011**. 26 out. 2011. Disponível em: <<http://veja.abril.com.br/noticia/educacao/alunos-do-colegio-christus-terao-de-refazer-o-enem-2011>>. Acesso em: 29 jan. 2014.

KLEIN, Ruben. Testes de rendimento escolar. In: MELLO E SOUZA, Alberto de (org.) **Dimensões da avaliação educacional**. 3. ed. Petrópolis, RJ: Vozes, 2011. p. 110-138.

KORETZ, Daniel. **Measuring up**: what educational testing really tells us. Cambridge, MA: Harvard University Press, 2008.

LEMANN, Nicholas Berthelot. **The big test**: the secret history of the American meritocracy. New York, NY: Farrar, Straus and Giroux, 2000.

LIMA, Elon Lages. **Matemática e ensino**. 3. ed. Rio de Janeiro, RJ: Sociedade Brasileira de Matemática, 2007.

MACHADO, Ledo Vaccaro. **Avaliação de larga escala e proficiência matemática**. 2010. 172 f. Dissertação (Mestrado em Ensino de Matemática) – Programa de Pós-graduação em Ensino de Matemática, Instituto de Matemática, Universidade Federal do Rio de Janeiro, Rio de Janeiro, RJ.

Disponível em: <<http://www.pg.im.ufrj.br/pemat/26%20Ledo%20Vaccaro.pdf>>. Acesso em: 29 jan. 2014.

NUNES, Carlos Henrique Sancineto da Silva; PRIMI, Ricardo. Impacto do tamanho da amostra na calibração de itens e estimativa de escores por teoria de resposta ao item. **Avaliação psicológica**. Porto Alegre, v. 4, n. 2, nov. 2005. p. 141-153. Disponível em: <[http://pepsic.bvsalud.org/scielo.php?script=sci\\_arttext&pid=S1677-04712005000200006&lng=pt&nrm=iso](http://pepsic.bvsalud.org/scielo.php?script=sci_arttext&pid=S1677-04712005000200006&lng=pt&nrm=iso)>. Acesso em 15 fev. 2014.

ORGANISATION FOR ECONOMIC CO-OPERATION AND DEVELOPMENT (OECD). **BRAZIL – Country note – Results from PISA 2012**. [2013?]. Disponível em: <<http://www.oecd.org/pisa/keyfindings/PISA-2012-results-brazil.pdf>>. Acesso em 15 fev. 2014.

\_\_\_\_\_. **PISA 2009 results: executive summary**. 2010. Disponível em: <<http://www.oecd.org/pisa/46643496.pdf>>. Acesso em 15 fev. 2014.

OSTERLIND, Steven J. **Constructing test items: multiple-choice, constructed-response, performance and other formats** (Evaluation in education and human services). 2. ed. Boston, MA: Kluwer Academic Publishers, 1998.

SACKS, Peter. **Standardized minds: the high price of America's testing culture and what we can do to change it**. Cambridge, MA: Perseus Publishing, 1999.

# ANEXO I – MANUAL DE PRODUÇÃO E REVISÃO DE ITENS DE MÚLTIPLA ESCOLHA

## SUMÁRIO

I. APRESENTAÇÃO .....	2
II. CONSTRUÇÃO DE UM BANCO DE ITENS .....	4
III. ESTRUTURA DE UM ITEM .....	10
III. 1. IDENTIFICAÇÃO .....	10
III. 2. DISCIPLINA.....	11
III. 3. NÍVEL DE DIFICULDADE.....	11
III. 4. GABARITO .....	14
III. 5. DESCRITOR.....	14
III. 6. TEXTO-BASE .....	18
III. 7. COMANDO .....	34
III. 8. OPÇÕES DE RESPOSTA.....	37
III. 9. COMENTÁRIOS .....	48
IV. ORIENTAÇÕES PRÁTICAS PARA A PRODUÇÃO DE ITENS.....	53
V. REVISÃO DE ITENS .....	55

## I. APRESENTAÇÃO

Este manual é voltado principalmente para a produção e a revisão de itens de múltipla escolha, sobretudo quando produzidos para a formação de um banco de itens. Secundariamente, pode ser usado como fonte de treinamento para produção de itens de múltipla escolha avulsos, para um teste específico e de amplitude menor, como uma prova a ser aplicada a uma turma de estudantes.

Em geral, os comentários e os itens usados a título de exemplo neste manual estão voltados para a formação de um banco de itens que se aplique a avaliações em larga escala que tenham por objetivo diagnosticar a aprendizagem sob uma perspectiva externa. Entretanto, são poucas as alterações necessárias para que as instruções sejam usadas em avaliações puramente classificatórias, ou seja, para fins de seleção, como em um concurso público. Também são reduzidas as adaptações requeridas para aplicações deste manual a processos de ensino e aprendizagem em um tipo de avaliação conhecida como “formativa” ou “processual”, com a função de se verificar se os estudantes estão atingindo os objetivos previstos no decorrer de sua trajetória e que possibilitem correções no processo antes de finda uma etapa.

Em caso de formação de um banco de itens, este manual tem um caráter complementar e deve ser utilizado em associação com o manual de produção e revisão de itens de múltipla escolha disponibilizado pela instituição responsável pela formação do banco. As instruções veiculadas neste manual têm caráter mais generalista que aquelas normalmente encontradas em um manual principal para produção e revisão de itens de múltipla escolha. Essa perspectiva foi escolhida para que seja possível utilizar este manual associado ao maior número possível de linhas pedagógicas e técnicas eleitas pelas diversas instituições que buscam construir bancos de itens. Assim, em vários momentos, serão encontradas aqui orientações de natureza múltipla, sendo

necessário atentar ao que é efetivamente prescrito no manual da instituição. No caso de eventual divergência entre o aqui advogado e o disposto no manual da instituição, é certo que as normas deste último devem prevalecer.

Este manual está voltado para a construção de itens associados à disciplina matemática, mas, com as devidas adaptações na interpretação, pode ser utilizado como referência para a produção e revisão de itens de múltipla escolha em outras áreas de conhecimento.

A próxima parte deste manual, seção II, apresenta informações úteis ao produtor de itens acerca da construção de um banco de itens, com brevíssimas observações sobre avaliação em larga escala e sobre a Teoria da Resposta ao Item (TRI).

Na parte posterior, seção III, é apresentada uma proposta de estrutura de um item, com observações detalhadas tratando de cada uma de suas partes. Há também vários exemplos de itens que podem ajudar a compreender os pontos mais complicados.

Em seguida, na seção IV, são apresentadas orientações práticas que podem ser úteis durante a produção e revisão de itens de múltipla escolha.

Finalmente, a seção V é dedicada exclusivamente à revisão de itens de múltipla escolha.

## II. CONSTRUÇÃO DE UM BANCO DE ITENS

O conteúdo deste manual de produção e revisão de itens de múltipla escolha pode, sem dúvida, ajudar ao professor que deseja, por exemplo, apenas preparar provas formadas de questões de múltipla escolha para uso com suas turmas de estudantes. Entretanto, tarefas muito complexas envolvem o treinamento em produção e revisão de itens de múltipla escolha, como as avaliações em larga escala ou a construção de portais na internet para apoio a estudantes. Neste sentido, é útil aos produtores, especialmente aos iniciantes, conhecer, ainda que superficialmente, o processo de formação de um banco de itens para uma avaliação em larga escala. É o que será apresentado a seguir.

As avaliações em larga escala podem ser compreendidas como um processo útil à definição de políticas públicas para a educação. Elas são externas às unidades escolares, têm caráter sistêmico e não se confundem com o processo de avaliação controlado pelos professores junto às suas turmas de estudantes.

Como são aplicadas a um elevado número de pessoas, as avaliações em larga escala quase que invariavelmente são testes compostos de itens de múltipla escolha, já que os esforços para obtenção das pontuações associadas aos examinados podem ser feitos por máquinas, redundando em economia de tempo e de recursos financeiros. Além disso, o uso de itens objetivos elimina o risco de falha na isonomia durante a atribuição de pontuações aos avaliados. Item, vale esclarecer, é a unidade básica de observação em qualquer teste<sup>24</sup>, que é instrumento de medida da avaliação. O que se deseja medir é o desempenho de um grupo de estudantes ou, ainda, a sua proficiência. Na

---

<sup>24</sup> HALADYNA, Thomas M. **Developing and validating multiple-choice test items**. 3. ed. Mahwah, NJ: Lawrence Erlbaum Associates, 2004. p. 3.

prática escolar, fala-se que uma prova, por exemplo, tem 40 questões; na terminologia mais técnica, diz-se que um teste é constituído por 40 itens.

Além dos itens de múltipla escolha normalmente integrantes das avaliações em larga escala, existem outros tipos, como aqueles conhecidos como itens de “resposta livre” ou de “resposta construída”, nos quais, por exemplo, o examinado preenche lacunas. Estes tipos de teste não são tratados neste manual, que é restrito aos itens de múltipla escolha.

As avaliações em larga escala servem também para acompanhar, ao longo do tempo, as alterações de desempenho de diversas populações. Com o uso adequado de tecnologias como a Teoria da Resposta ao Item (TRI) é possível garantir esta desejada comparação. De acordo com Klein,

a TRI, associada a outros procedimentos estatísticos, permite comparar alunos, estimar a distribuição de proficiências da população e subpopulações e a ainda monitorar os progressos de um sistema educacional<sup>25</sup>.

Ou seja, mesmo usando testes diferentes em populações diferentes, garante-se, por exemplo, que os estudantes que concluíram o ensino médio em certo ano melhoraram ou pioraram em relação ao ano anterior. Isso ocorre porque é criada uma escala comum a todos os avaliados. É possível até mesmo aplicar testes diferentes a dois grupos de pessoas e construir comparações entre eles, ordenando as pontuações atribuídas aos examinados, tudo em um mesmo exame, sem perda de isonomia. De fato, foi o que ocorreu em um episódio no qual vazou parte dos conteúdos sigilosos do Exame Nacional do Ensino Médio (Enem) no ano de 2011. Alguns itens foram utilizados indevidamente em um estabelecimento de ensino em Fortaleza, no Ceará, durante simulados usados na preparação de estudantes para o exame e, incontestavelmente, isto conferiu

---

<sup>25</sup> KLEIN, Ruben. Testes de rendimento escolar. In: Souza, Alberto de Mello e (org.) **Dimensões da avaliação educacional**. 3. ed. Petrópolis, RJ: Vozes, 2001. p. 121.

vantagem a estes examinados. Após ter sido delimitado o conjunto de examinados que poderiam ter se beneficiado do vazamento, somente em relação a estes foram cancelados 14 itens, sem que fosse necessário repetir o exame em todo o território nacional, o que seria imenso transtorno para os participantes e representaria vultosa perda de recursos financeiros<sup>26</sup>. A utilização de exames diferentes associada à manutenção da isonomia entre os candidatos contraria o senso comum, mas é suportada cientificamente pela TRI agregada a procedimentos estatísticos.

Vai além do escopo deste manual discorrer mais sobre a TRI e seus fundamentos matemáticos, que são complexos. O leitor interessado pode se remeter ao documento intitulado “Entenda a sua nota no Enem – Guia do participante”, publicado pelo Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (Inep)<sup>27</sup>, um guia bastante didático que pode servir de introdução à TRI. Por ora, basta acrescentar que, para que seja possível fazer as comparações entre grupos e resultados ao longo do tempo, é necessário que os itens passem por um procedimento denominado pré-teste. Nesta etapa, um grupo de examinados recebe e resolve os itens e, por meio de procedimentos estatísticos posteriores, cada item é associado a um conjunto de parâmetros, que usualmente são três: (a) a capacidade do item de discriminar aqueles que dominam a habilidade nele tratada daqueles que não a possuem; (b) o seu nível de dificuldade; e; (c) a probabilidade de acerto do item ao acaso. O número de participantes no processo de pré-teste normalmente fica entre 200 e 500 sujeitos<sup>28</sup>. Novamente, entrar em mais detalhes

---

<sup>26</sup> HONORATO, Renata. **Alunos do Colégio Christus terão de refazer o Enem 2011**. 26 out. 2011. Disponível em: <<http://veja.abril.com.br/noticia/educacao/alunos-do-colegio-christus-terao-de-refazer-o-enem-2011>>. Acesso em: 29 jan. 2014.

<sup>27</sup> BRASIL. Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (Inep). **Entenda a sua nota no Enem – Guia do participante**. Brasília, DF: MEC/Inep, 2012. Disponível em: <<http://mapaitensenem.inep.gov.br/mapaNota/home.seam>>. Acesso em: 02 fev. 2014.

<sup>28</sup> É razoável esperar bons resultados quanto aos parâmetros (b) e (c) quando o número de examinados que participam do procedimento de pré-teste é maior que 200. Em especial, para garantir a estabilidade do parâmetro (a), “amostras com 500 sujeitos trazem resultados muito

ultrapassaria o recorte pretendido neste manual e o leitor que deseja aprofundar a compreensão deste modelo de três parâmetros pode recorrer a trabalhos específicos, disponíveis em língua portuguesa<sup>29</sup>.

O processo de avaliação em larga escala é longo e dividido em etapas. Em primeiro lugar, deve ser enumerado um conjunto de habilidades que se deseja medir em uma população. Esta lista de habilidades, conhecida tecnicamente como “matriz de referência”, deve ser compatível com a população examinada. Por exemplo, a matriz de referência do Exame Nacional do Ensino Médio (Enem) deve conter as habilidades esperadas dos estudantes que terminam o ensino médio. Cada uma das habilidades é apresentada sob a forma de um descritor, conceito que será esclarecido em detalhe na subseção III.5 deste manual. A formação da matriz de referência é feita por especialistas, que também decidem a quantidade de itens necessários para o banco que se deseja que seja produzido, bem como estabelecem um cronograma e um orçamento para todo o processo. Em seguida, um manual de produção e revisão de itens de múltipla escolha deve ser preparado. Depois, pessoal com experiência docente deve ser recrutado e treinado. São especificadas as quantidades de itens que devem ser produzidos por cada autor. Inicia-se a produção dos itens, etapa que será profundamente detalhada na próxima seção. Após os itens serem produzidos, eles são revisados pedagógica e tecnicamente, preferencialmente por outros produtores mais experientes. Podem ser previstos outros tipos de revisão, especialmente relacionadas à conformidade à língua portuguesa e à questão dos direitos autorais dos

---

próximos aos estimados com amostras maiores”, conforme NUNES, Carlos Henrique Sancineto da Silva; PRIMI, Ricardo. Impacto do tamanho da amostra na calibração de itens e estimativa de escores por teoria de resposta ao item. **Avaliação psicológica**. Porto Alegre, v. 4, n. 2, nov. 2005. p. 149. Disponível em: <[http://pepsic.bvsalud.org/scielo.php?script=sci\\_arttext&pid=S1677-04712005000200006&lng=pt&nrm=iso](http://pepsic.bvsalud.org/scielo.php?script=sci_arttext&pid=S1677-04712005000200006&lng=pt&nrm=iso)>. Acesso em 15 fev. 2014.

<sup>29</sup> Por exemplo, no elucidativo Anexo IV – Descrição do Modelo Logístico de Três Parâmetros – encontrado em MACHADO, Ledo Vaccaro. **Avaliação de larga escala e proficiência matemática**. 2010. 172 f. Dissertação (Mestrado em Ensino de Matemática) – Programa de Pós-graduação em Ensino de Matemática, Instituto de Matemática, Universidade Federal do Rio de Janeiro, Rio de Janeiro, RJ. Disponível em: <<http://www.pg.im.ufrj.br/pemat/26%20Ledo%20Vaccaro.pdf>>. Acesso em: 29 jan. 2014.

elementos de apoio incluídos no texto dos itens. A cada processo de revisão, o item pode ser aprovado sem ressalvas, pode receber pequenos ajustes, pode retornar para o produtor para alterações ou, ainda, pode ser rejeitado, se estiver em profundo desacordo com as exigências estabelecidas no manual. Após as revisões, o item pode ser avaliado por meio de programas de computador especiais com o objetivo de garantir seu ineditismo, pois, em geral, os produtores se comprometem formalmente no sentido de que estarão efetivamente criando material novo. Uma vez formado um conjunto de itens, eles são pré-testados com um grupo de examinados. A partir da análise estatística detalhada dos resultados do pré-teste, sob a ótica técnica e pedagógica, os itens podem ser enviados para o banco de itens ou serem rejeitados. Aproximadamente metade dos itens sobrevive após todas as atividades relacionadas à produção de itens, aos procedimentos de revisão e verificações<sup>30</sup>. Admitindo que se utilize a Teoria da Resposta ao Item (TRI), a análise estatística rotula cada item com os três parâmetros citados anteriormente: capacidade de discriminação, dificuldade e possibilidade de acerto ao acaso<sup>31</sup>. Qualificado com estes parâmetros, o item passa a compor um banco, de onde os itens com diversos níveis de dificuldade podem ser selecionados para integrar um teste. Este exame pode ser aplicado aos avaliados por meio da impressão em papel ou usando meios digitais. Após a aplicação do teste, especialistas elaboram relatórios pedagógicos que podem servir de suporte para a tomada de decisões por parte dos gestores locais e para a formação de políticas públicas amplas na área da educação.

---

<sup>30</sup> HOLTZMAN, K; CASE, S. M. e RIPKEY, D *apud* HALADYNA, Thomas M. **Developing and validating multiple-choice test items**. 3. ed. Mahwah, NJ: Lawrence Erlbaum Associates, 2004. p. 97.

<sup>31</sup> É interessante observar o desconhecimento sobre a TRI e que alguns de seus aspectos contrariam o senso comum. Em um exemplo, ainda tratando do episódio ocorrido em Fortaleza, em 2011, alguns setores da imprensa nacional divulgaram, com alarme, que os itens usados no Exame Nacional do Ensino Médio (Enem) foram os mesmos que haviam sido aplicados em pré-testes. Ora, a metodologia, em verdade, exige que todos os itens integrantes do exame que é aplicado em nível nacional tenham passado pelo pré-teste, etapa do processo indispensável para a “calibração” de cada item no que se refere aos três parâmetros estatísticos de capacidade de discriminação, dificuldade e possibilidade de acerto ao acaso.

Vale ainda acrescentar que etapas deste processo podem não ser realizadas, caso o objetivo seja menos ambicioso do que uma avaliação em larga escala, como, por exemplo, na hipótese de simples formação de um banco de itens para treinamento de estudantes.

O processo descrito em apertada síntese é intrincado e exige a participação de muitos profissionais. De modo geral, um educador inaugura sua participação em um processo deste tipo por meio da elaboração de itens, já que esta atividade requer a participação de um numeroso corpo docente experiente e, apesar de inerentemente complexa e trabalhosa, é a ação que serve de fundamento para todo o processo. Além do mais, em um bom banco de itens há necessidade de reposição. Ele deve ser constantemente alimentado com itens novos, procedimento necessário para garantir ineditismo na experiência no avaliado. Assim, a demanda por produção de itens de qualidade, apesar de maior na formação inicial do banco, permanece durante seu uso. Se se considerar a possibilidade de exclusão de itens do banco a partir da sua divulgação em testes em larga escala, essa necessidade de reposição é ainda mais manifesta.

### **III. ESTRUTURA DE UM ITEM**

Nesta seção, dividida em nove subseções, é apresentada uma proposta de estrutura de um item, cujas partes são: identificação, disciplina, nível de dificuldade, gabarito, descritor, texto-base, comando, opções de resposta e comentários. Cada uma destas partes é tratada em uma subseção.

Dependendo do manual de produção e revisão de itens de múltipla escolha preparado pela instituição responsável pela construção do banco de itens, outras partes podem estar presentes na estrutura do item. Em especial, podem existir campos nesta estrutura – também conhecida como “formulário” ou “máscara” – destinados ao preenchimento por parte dos revisores técnicos, pedagógicos, linguísticos e de direitos autorais. Normalmente cada um destes revisores pode separar os itens em classes como as seguintes: aprovado, aprovado com alterações, devolver ao autor para alterações, rejeitado. De acordo com o manual de produção e revisão de itens de múltipla escolha, estas classes podem ser mais ou menos numerosas. É possível, ainda, que uma ou mais das classes mencionadas não se apliquem a certos procedimentos de revisão.

#### **III. 1. IDENTIFICAÇÃO**

Cada item deve ser identificado de modo preciso, tornando-o único e facilitando a indexação em um banco de dados. A identificação facilita também as referências a um item, no caso de retorno do mesmo ao produtor após alguma das etapas de revisão. O manual de produção e revisão de itens de múltipla

escolha pode informar aos produtores como a identificação deve ser feita. Os responsáveis pela distribuição do trabalho entre os produtores devem fornecer informações claras a estes sobre a identificação. Por exemplo, um produtor deve saber a numeração inicial e a final de um lote de itens que deve criar.

É conveniente que a identificação de um item não revele quem foi o produtor, pois se recomenda que os revisores façam seu trabalho da maneira mais isenta possível. Assim, é aconselhável que o código de cada item não tenha qualquer informação a respeito do produtor, como seu nome ou as iniciais de seu nome completo. O mesmo se aplica aos nomes dos arquivos que contêm os itens.

### **III. 2. DISCIPLINA**

Este campo apresenta o nome da disciplina ou área de conhecimento do item.

### **III. 3. NÍVEL DE DIFICULDADE**

Durante a produção de um item, o autor deve indicar sua dificuldade segundo o definido no manual de produção e revisão de itens de múltipla escolha da instituição que os encomendou. De forma geral, são três classes, como fácil, médio e difícil. É razoável ainda entender que os itens fáceis podem ser simplesmente fáceis ou muito fáceis, o mesmo valendo, analogamente, para os itens difíceis, que poderiam ser considerados simplesmente difíceis ou muito

difíceis. É possível ainda associar um número – de um até três, por exemplo – ao nível de dificuldade do item. Se os itens forem utilizados em um processo de pré-testagem, a classificação dada pelo produtor tem caráter provisório, já que cada item terá um nível de dificuldade próprio que será determinado pelo resultado do pré-teste. Em verdade, se for descoberto, a partir do procedimento de pré-teste, que um item é extremamente fácil ou extremamente difícil, ele provavelmente não será aproveitado, dada sua inutilidade para discriminar os participantes do teste. De todo modo, a classificação feita pelo autor é importante por exigir que ele construa itens de dificuldades diferentes. Para tanto, a experiência docente e uso da linguagem adequada aos avaliados são dois fatores indispensáveis ao produtor, que deve ser capaz de antever a dificuldade que os avaliados irão experimentar ao buscar resolver o item.

O manual de produção e revisão de itens de múltipla escolha da instituição pode ainda exigir que o autor indique qual o nível de escolaridade relacionado ao item, se os descritores da matriz de referência forem divididos por ano letivo. Em um exemplo, a matriz pode ser dividida entre os quatro últimos anos do ensino fundamental e o produtor deve indicar se o item é destinado para estudantes do sexto, sétimo, oitavo ou nono ano. Neste caso, é admissível que alguns descritores apareçam na matriz de referência em mais de um ano, situação em que os itens deverão ter sua dificuldade avaliada também em função disso. Por outro lado, nada impede que a matriz seja voltada para um período maior que um ano, como no caso de itens produzidos para aplicação ao término do ensino médio, como ocorre com o Exame Nacional do Ensino Médio (Enem).

É mais ou menos usual se preparar itens segundo uma distribuição aproximadamente semelhante à apresentada a seguir: a) itens muito fáceis: 10%; b) itens fáceis: 20%; c) itens médios: 40%; d) itens difíceis: 20%; e) itens muito difíceis: 10%. Vale dizer que o produtor não deve se esquivar a produzir os itens muito fáceis, temendo por uma eventual facilidade excessiva, já que todos os níveis de dificuldade apresentados serão necessários para um

juízo adequado dos avaliados. O mesmo se aplica aos itens muito difíceis.

No entanto, especialmente no caso de itens preparados para concursos públicos muito concorridos que buscam selecionar candidatos para cargos públicos, é recomendável uma concentração maior dos itens nas classes de itens médios, difíceis e muito difíceis. Isso se dá pela preocupação em evitar que muitos participantes consigam acertar todos os itens do teste. Se isto ocorrer, de maneira que o número de participantes com nota máxima seja maior que o número de vagas oferecidas, os destinados aos cargos públicos disputados terão todos a mesma nota máxima, havendo ainda outros que, também com nota máxima, não terão direito aos cargos, o que enfraquece sobremaneira o princípio da seletividade, um dos que devem orientar os concursos públicos<sup>32</sup>.

Em especial, o autor de itens deve investigar, por meio da leitura do manual de produção e revisão de itens de múltipla escolha fornecido pela instituição responsável pela formação do banco, qual a abordagem técnica adotada para que cada item seja classificado segundo seu nível de dificuldade.

Em certos casos, o que depende da instituição, uma complexa taxonomia é estabelecida, a qual deve ser plenamente assimilada pelo produtor, com o objetivo de evitar a rejeição de itens por não conformidade quanto ao nível de dificuldade. Recomenda-se também que, se for necessário criar um conjunto de itens relacionados a determinado descritor – por exemplo: um fácil, um médio e outro difícil –, o produtor planeje estes itens de maneira conjunta, garantindo uma melhor adequação dos itens aos níveis de dificuldade pretendidos.

---

<sup>32</sup> A legislação federal, por meio do parágrafo único do artigo 27 da Lei Nº 10.741, de 1º de outubro de 2003 (Estatuto do Idoso), estabelece que “o primeiro critério de desempate em concurso público será a idade, dando-se preferência ao de idade mais elevada”. Critérios de desempate adicionais, desde que previstos em edital, podem ser a preferência dada aos que participaram como jurados em Tribunais do Júri ou como mesários em eleições.

### **III. 4. GABARITO**

Deve haver indicação clara da única opção de resposta correta. As demais opções de resposta devem ser necessariamente erradas. Há consenso no sentido de que não é cabível construir itens em que se exija do avaliado apontar a mais correta dentre as opções de resposta.

### **III. 5. DESCRITOR**

Dependendo do manual de produção e revisão de itens de múltipla escolha e da instituição responsável por formar o banco de itens, o descritor pode ser referido como “competência” ou “habilidade”. Pode ser ainda chamado de “objetivo”. Não faz parte do escopo deste trabalho discutir estas definições, defendendo esta ou aquela nomenclatura. Neste manual, dentre todos estes termos – descritor, competência, habilidade, objetivo – será usado o primeiro.

Um descritor sempre está associado a uma ação a ser efetuada pelo avaliado. Há, portanto, necessidade de um verbo, como: enunciar, identificar, aplicar, calcular, provar, formular, estimar ou julgar. Aqui, estes verbos – que são exemplos dentre muitos outros – estão apresentados do mais simples ao mais complexo no que se refere à mobilização das capacidades cognitivas do avaliado.

É muito importante entender com abrangência um descritor, atentando para a ação que ele realmente está associado dentre todas aquelas que podem ser efetuadas pelo avaliado. É útil também refletir sobre o que o descritor não é, ou

seja, é importante ter cuidado na elaboração do item no sentido de não ultrapassar os limites propostos pelo descritor. Por exemplo, em um item associado a um descritor como “operar com números inteiros envolvendo as operações de adição e/ou subtração” não deve existir uma situação-problema, mas apenas uma proposta de cálculo. Esta somente deve surgir em um descritor como “resolver situações-problema com números inteiros, envolvendo diferentes significados das operações (adição, subtração, multiplicação, divisão, potenciação e radiciação)”. Um produtor que, inadvertidamente, considerar apenas o primeiro destes dois descritores, sem atentar para a existência do segundo, pode produzir um item, em verdade, voltado para o segundo descritor, o que, em tese, tornaria o item rejeitado em sua primeira revisão.

Assim, para que seja claramente delimitado o que trata cada descritor, é imprescindível ao produtor examiná-los em conjunto, não sendo recomendável iniciar a produção sem uma lista completa dos descritores, que é a matriz de referência definida antes do início da produção de itens. Os responsáveis pela distribuição do trabalho aos produtores devem, portanto, fornecer a estes a lista dos descritores. Alternativamente, a matriz de referência pode estar presente no manual de produção e revisão de itens de múltipla escolha.

Eventualmente, uma lista de descritores pode apresentar, para cada um deles, uma ou mais habilidades entendidas como pré-requisitos da habilidade central do item que deve ser produzido, o que ajudaria a delimitar o âmbito de seu descritor. Se o manual for detalhado, cada descritor pode vir acompanhado ainda de um pequeno texto que particulariza alguns aspectos do descritor e eventualmente o compara com outros descritores com os quais poderia ser confundido. Pode existir ainda um item finalizado como exemplo de cada descritor, o que pode ser especialmente útil para produtores iniciantes.

Há instituições que organizam os descritores de maneira muito próxima aos conteúdos, em uma orientação mais estruturalista e, portanto, clássica. De fato, os professores estão acostumados, de maneira geral, a ensinar e usar os

materiais didáticos segundo os conteúdos. Assim, espera-se que um descritor como “resolver equações do segundo grau completas e incompletas” seja inteligível para qualquer professor de matemática. Entretanto, existe uma linha pedagógica que privilegia uma visão mais pragmática da aprendizagem, valorizando a aplicação dos conhecimentos na vida, em situações possíveis e, mais que isso, típicas do cotidiano. Descritores alinhados com essa linha de pensamento são como “avaliar propostas de intervenção na realidade envolvendo variação de grandezas” e, muitas vezes, são inacessíveis aos professores produtores de itens em um primeiro momento. É por isso que o docente, de modo geral, deve dedicar um tempo suficiente para comparar os descritores de uma matriz de referência para que seja possível produzir adequadamente um item, considerando um descritor dado. Afinal, a formação universitária e a estrutura dos livros didáticos, dentre outros fatores, tornam mais difícil interpretar descritores como este último apresentado, que se afasta de uma proposta mais estruturalista e próxima dos conteúdos. Enfim, em certos descritores integrados a um corpo mais pragmático, uma interpretação equivocada pode resultar na produção de um item que, de fato, refere-se a outro descritor da matriz de referência. Neste caso, se for possível apenas mudar o descritor e manter o restante do item, o trabalho está salvo. Entretanto, se o item deveria ter sido produzido exatamente para o descritor dado, o trabalho pode ser perdido, no caso de rejeição do item durante o processo de revisão.

Outro aspecto que deve ser observado é que cada item deve estar associado a apenas um descritor, não sendo admissível um item que mobilize habilidades diferentes. Evidentemente, trata-se de evitar a conjunção de habilidades com aproximadamente mesma complexidade. Nos casos em que se trata de habilidades tidas como pré-requisito, é até mesmo necessário que um item exija duas ou mais habilidades associadas para que o avaliado encontre a opção de resposta correta. Por exemplo, um descritor do tipo “reconhecer e aplicar o Teorema de Pitágoras em situações-problema” certamente envolverá habilidades como “efetuar cálculos que envolvam potenciação e radiciação”.

Ou seja, este descritor é pré-requisito daquele, de modo que ambos tratam de competências necessárias ao avaliado para apresentar uma resposta satisfatória. Por outro lado, um item cujo descritor é “calcular medidas de tendência central” não pode apresentar um gráfico complexo e detalhado em seu texto-base, sob pena de ser também associado a um descritor como “resolver problema com dados apresentados em tabelas ou gráficos”. Uma crítica que pode ser apresentada a esta exigência de não usar mais de um descritor em um item é que, na vida fora da escola, ou seja, no mundo da ciência e do trabalho, as situações jamais surgem de maneira tão compartimentalizada. Uma resposta razoável a esta crítica é a que afirma que, sendo o item o elemento mais simples de um teste, aquele deve ser, por natureza, simples, sendo este, o teste, necessariamente mais complexo, o verdadeiro instrumento capaz de permitir uma análise mais completa do avaliado, a partir das interações entre as respostas apresentadas a vários itens. Ainda tratando da limitação de uma habilidade por item, nada impede que um professor, em sua atividade individual, produza também itens menos afetos a esta regra, admitindo a associação de dois ou mais descritores, até mesmo no sentido de proporcionar aos estudantes o contato com situações mais complexas. Vale dizer que as competições mais sofisticadas de matemática, como as olimpíadas, são completamente incompatíveis com a limitação aqui enfatizada. Entretanto, é necessário ter em mente que os manuais de produção e revisão de itens de múltipla escolha invariavelmente rejeitam itens associados a duas ou mais habilidades de complexidade semelhante.

Tratando da maneira que o item é construído pelo produtor, é importante ter em mente que o descritor é o ponto de partida, ou seja, o indicado é que todo o item, em todas as suas partes constitutivas, deve ser gerado a partir da habilidade veiculada no descritor. Assim, a armadilha fácil, comum aos produtores iniciantes, que consiste em ter uma boa ideia, construir o item e, depois disso, procurar o descritor mais apropriado, deve ser evitada. De modo geral, o item deve ser criado a partir do descritor. Se muitos itens forem criados a partir de boas ideias, pode haver uma concentração indesejável de itens em

determinados descritores, o que pode comprometer o equilíbrio da produção de um autor e, por conseguinte, a constituição do banco de itens. Este tipo de problema é mais comum quando se encomenda um lote composto de muitos itens a cada produtor, sem que sejam especificados quantos itens devem ser produzidos para cada descritor. Essa falha é mais rara quando os pedidos são feitos ao longo do tempo, com intervalos periódicos, sendo os descritores claramente definidos para cada período de produção.

### **III. 6. TEXTO-BASE**

O texto-base, que pode ser considerado a porção mais fundamental do item, apresenta o problema que será analisado pelo avaliado.

O texto-base deve ser de fácil compreensão e o autor do item deve ter em mente os avaliados que terão menor preparo, criando um texto que seja facilmente compreendido por estes examinados. O importante é não gerar situações nas quais não seja possível afirmar com certa segurança se um erro é fruto da falta de domínio da habilidade tratada no descritor ou se decorre de alguma dificuldade na compreensão do texto-base. Entretanto, a clareza deve se coadunar com o rigor e não podem existir erros conceituais no item. O equilíbrio entre a clareza e o rigor pode ser uma fonte de dificuldade quando da produção de itens.

O texto-base não deve ser desnecessariamente longo, pois o tempo destinado à compreensão e à escolha da opção de resposta para cada item, por parte do avaliado, é geralmente reduzido, limitado a pouquíssimos minutos. No caso do Exame Nacional do Ensino Médio (Enem), por exemplo, o tempo médio para tratar cada item é de três minutos. Entretanto, se se descontar o tempo de

preenchimento do formulário de respostas, ele é ainda menor. Assim, o autor de itens deve ter em mente a limitação de tempo do avaliado durante seu trabalho de produção. No entanto, é razoável esperar itens com texto-base mais longo, se existirem, concomitantemente, itens com texto-base mais curto feitos pelo mesmo produtor. A obrigatoriedade de produzir somente itens curtos limitaria sobremaneira a criatividade na produção e suprimiria, de pronto, a inclusão de textos maiores, eventualmente necessários em abordagens mais complexas e criativas, principalmente nas questões inovadoras que devem apresentar certas informações imprescindíveis ao avaliado para a plena compreensão da situação-problema. No caso de itens que envolvam cálculos, é fortemente recomendado que o produtor os faça do mesmo modo que o avaliado os fará, usando os mesmos meios – normalmente papel e caneta –, com objetivo de estimar o tempo que será gasto no procedimento por um avaliado.

No texto-base podem ser incluídos diversos elementos de apoio, eventualmente conhecidos como “suporte”. Eles podem ser trechos de obras ou notícias de jornais, de revistas ou ainda veiculadas via internet. Os elementos de apoio podem ainda ser imagens, quadrinhos, charges, mapas, tabelas, gráficos, comunicados publicitários<sup>33</sup> e outros. Os elementos de apoio podem ser ligados a fatos reais ou oriundos de situações hipotéticas imaginadas pelo produtor.

Há divergência quanto à imprescindibilidade dos elementos de apoio incluídos no texto-base: em alguns manuais de produção e revisão de itens de múltipla escolha está disposto que somente os elementos de apoio cruciais para a compreensão do item por parte do avaliado devem ser integrantes do texto-base; em outros é admitida a inclusão de elementos de apoio de valor meramente ilustrativo ou estético como motivadores do avaliado. Assim, se um

---

<sup>33</sup> Uma campanha pela redução de acidentes de trânsito, por exemplo, pode ser usada como elemento de apoio do texto-base. Entretanto, como se verá mais a frente, uma propaganda de determinada marca ou produto deve ser evitada.

item se refere a um personagem histórico, por exemplo, a inclusão de sua fotografia é rejeitada ou aprovada, dependendo do manual utilizado. A imagem de um atleta em um item que trata da velocidade de um competidor em uma pista de atletismo tem a mesma natureza dúbia quanto à validade de sua inclusão em um item. Assim, o produtor deve estar atento a este aspecto, sob pena de ter itens devolvidos pelo revisor, se houver proibição de inclusão de elementos de apoio desnecessários e eles forem colocados no item. No caso de incerteza sobre a permissão de se incluir imagens de cunho exclusivamente ilustrativo, a melhor atitude do produtor é evitar estas imagens, incluindo no item somente aquelas necessárias para que o examinado encontre a opção de resposta verdadeira. Vale ainda dizer que a qualidade gráfica dos elementos de apoio incluídos no item deve ser observada: imagens de baixa resolução ou mal digitalizadas são indesejadas.

É comum se exigir que um texto-base não seja relacionado a mais de um item, principalmente se eles integrarem uma plataforma eletrônica de testes, ou seja, se eles não forem impressos em papel. De fato, se a ordem dos itens for alterada ou se algum deles for excluído, poderia haver prejuízo no entendimento do teste. Apesar disso, existem algumas instituições que permitem o uso de um elemento de apoio em mais de um item. O produtor, portanto, deve estar informado se é permitido usar os mesmos elementos de apoio em itens diferentes. Por exemplo, um mesmo gráfico de colunas pode ser usado em três itens: no item fácil, o autor poderia estabelecer ao avaliado que identificasse o valor relacionado à maior coluna; no item médio, o examinado poderia ser instado a colocar os valores relacionados às colunas em ordem; no mais difícil, poderia ser solicitado que ele indicasse a diferença entre o maior e o menor valores. Em todos os casos, é bom salientar, o elemento de apoio em comum deve estar repetido nos itens em que for necessário, de modo que o avaliado sempre consiga entender perfeitamente o item usando apenas as informações dele integrantes, jamais necessitando de outros itens para chegar à opção de resposta correta. Afinal, repise-se, os itens devem ser independentes entre si.

Um aspecto muito importante no que se refere aos elementos de apoio incluídos pelo produtor nos itens é a questão dos direitos autorais. O produtor deve receber instruções precisas a respeito deste assunto. Há instituições que contratam fornecedores de imagens. Estes provedores oferecem milhões de figuras e fotografias que podem ser localizadas a partir de uma pesquisa usando palavras-chave. Outras instituições permitem, com maiores ou menores restrições, que o produtor procure por imagens em qualquer sítio, usando uma ferramenta de buscas de imagens na internet. Pode existir uma etapa de revisão do item na qual especialistas em direito autoral verificam se o elemento de apoio incluído pelo autor do item é de livre publicação, é sujeito a pagamento pelo uso ou se tem uso vetado ou proibitivo do ponto de vista econômico. É importante que os produtores e os revisores sejam detalhadamente informados a respeito de eventuais limitações quanto aos direitos autorais, para que não sejam produzidos bons itens posteriormente rejeitados em função de violações às políticas de direito autoral da instituição organizadora do banco de itens.

Ainda no que tange aos elementos de apoio incluídos nos itens, é necessário que o produtor registre quando os elementos de apoio nele incluídos são de autoria própria ou, no caso de elementos de apoio obtidos de terceiros, que seja feita a referência bibliográfica precisa para o correto registro dos créditos por parte dos responsáveis pela publicação ou impressão dos itens. A referência bibliográfica deve ser apresentada em harmonia com o prescrito no manual de produção e revisão de itens de múltipla escolha disponibilizado pela instituição responsável pela formação do banco.

O autor dos itens deve ainda estar informado se é permitido efetuar alterações nos elementos de apoio incluídos no texto-base. Há manuais de produção e revisão de itens de múltipla escolha que proíbem modificações, de modo que os elementos de apoio devem ser incluídos integralmente ou, no máximo, é admitida apenas a supressão de partes do elemento de apoio. Isso se justifica pelo fato ser mais difícil obter licença de uso dos elementos de apoio junto aos que detêm os direitos autorais quando estes elementos são alterados. Assim,

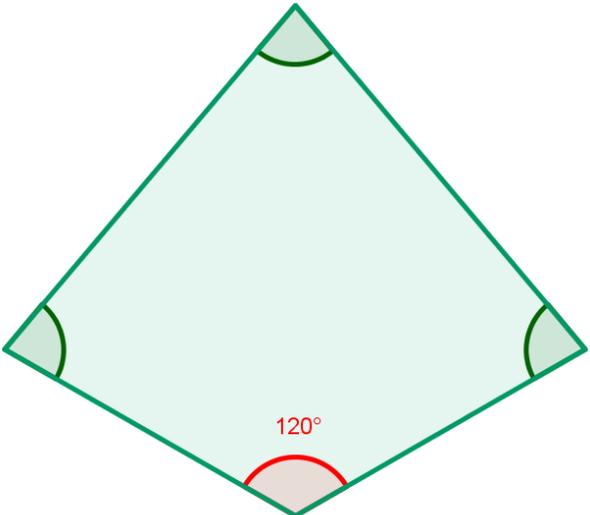
segundo esta orientação, restaria ao produtor somente criar um elemento de apoio – imagem ou gráfico, por exemplo – se não encontrar algo que atenda perfeitamente ao texto-base. Há manuais, entretanto, que não apresentam restrições às adaptações porventura feitas pelo produtor em elementos de apoio produzidos por terceiros ou retirados de um banco de elementos de apoio colocado à sua disposição.

Vários manuais de produção e revisão de itens de múltipla escolha vetam a inclusão de elementos de apoio de certas fontes. Normalmente os livros didáticos estão incluídos nestas fontes proibidas. Sítios de baixa confiança, como blogs pessoais, por exemplo, também podem ser proibidos. O produtor deve estar atento a eventuais restrições, sob pena de ter seu item rejeitado.

Deve ser dada especial atenção ao uso de cores. Se os itens forem utilizados exclusivamente em testes aplicados usando meios digitais, normalmente não há problemas. É até mesmo indicado que sejam usadas cores: é muito mais útil fazer referência, por exemplo, ao “polígono azul” do que ao “polígono situado no canto inferior direito” ou ao “polígono ABCD”. Entretanto, caso os itens sejam impressos, é importante saber se haverá cores ou apenas tons de cinza. Um gráfico produzido em cores e impresso em tons de cinza pode ficar ininteligível, de modo que o examinado pode confundir as cores das linhas, barras ou colunas. Neste caso, algumas soluções cabíveis são usar apenas as cores preta, cinza e branca ou utilizar linhas de espessura diferentes. O exemplo veiculado no item X1 exhibe como o uso de cores pode simplificar o texto-base<sup>34</sup>. Entretanto, este item não poderia ser impresso em tons de cinza.

---

<sup>34</sup> Os itens apresentados neste manual como exemplos seguem um formato (ou uma “máscara”) padronizado e simplificado, cujos campos são aqueles considerados indispensáveis. Assim, não foram incluídos campos possivelmente presentes, como “nível de dificuldade” e “ano de escolaridade”. Os campos de comentários surgem apenas na subseção III.9, que trata especialmente do assunto. Também não há campos destinados aos revisores técnicos, pedagógicos, linguísticos e de direitos autorais, que eventualmente estariam presentes. Os itens ora são aplicáveis ao ensino fundamental, ora são relacionados ao ensino médio. Os descritores apresentados nos itens não pertencem a uma matriz de referência específica e são usados apenas como exemplos.

<b>Identificação</b>	X1	<b>Disciplina</b>	Matemática	<b>Gabarito</b>	B
<b>Descritor</b>	Resolver problema utilizando propriedades dos polígonos (soma de seus ângulos internos e externos, número de diagonais, cálculo da medida de cada ângulo interno nos polígonos regulares).				
<b>Texto-base</b>					
O quadrilátero abaixo tem ângulos internos de mesma medida indicados em verde e um ângulo interno indicado em vermelho, que mede $120^\circ$ .					
					
<b>Comando</b>					
A medida de cada ângulo indicado em verde é					
<b>Opções de resposta</b>					
A. $60^\circ$ .					
B. $80^\circ$ .					
C. $90^\circ$ .					
D. $120^\circ$ .					

O autor do item deve estar atento à proibição de incluir mensagens indesejáveis no texto-base, como aquelas de apologia política, de propaganda de marcas e produtos, atentatórias a quaisquer grupos sociais ou que revelem preferências pessoais do autor, como torcer por um time de futebol. O texto deve ainda apresentar conteúdo compatível com a faixa etária dos examinados, especialmente no que se refere a conteúdos sexuais ou relacionados à violência. O autor deve verificar ainda se há questões adicionais de natureza ética junto à instituição que encomendou os itens. Deve, também, não inserir qualquer elemento que desvie a atenção do examinado da habilidade veiculada no descritor. É importante verificar ainda se o conteúdo

veiculado pelo texto-base e as expressões escolhidas têm abrangência mais ou menos homogênea dentre os participantes. Por exemplo, se um teste será aplicado em âmbito nacional, não seria adequado tratar de uma situação que é de maior conhecimento em uma região ou cidade do país. O mesmo se aplica para o vocabulário usado no item. Outro aspecto relevante é evitar escolher, para o texto-base, algum tema polêmico dentro da própria área de conhecimento a que o item em produção pertence, evitando conteúdos que podem ter abordagens diferentes entre os docentes e nos materiais pedagógicos. É o caso, por exemplo, da dúvida acerca da inclusão do zero dentre os números naturais, ou, ainda, sobre a definição do trapézio: alguns consideram que basta um par de lados paralelos para que um quadrilátero seja considerado trapézio – o que incluiria os paralelogramos (e, por extensão, os retângulos, os losangos e os quadrados) no conjunto dos trapézios – enquanto outros preferem exigir exatamente um par de lados paralelos – e isso excluiria os paralelogramos do conjunto dos trapézios.

Existem dois aspectos que devem ser privilegiados pelo autor de itens e que representam, em conjunto, uma oposição às abordagens livrescas, ou seja, semelhantes às encontradas facilmente nos materiais didáticos tradicionais. Os itens devem ser: (a) criativos; e (b) orientados ao conhecimento, em lugar da simples memorização. Quanto à criatividade, há quem diga que construir bons itens de teste é, provavelmente, o mais exigente tipo imaginável de produção textual<sup>35</sup>. É recomendável que o autor se proponha a examinar diversas atividades do cotidiano, notícias de jornais e revistas, bem como tenha em mente conteúdos de outras áreas de conhecimento. Por exemplo, uma notícia que versa sobre o crescimento do produto interno bruto (PIB) pode servir como tema de um item que trate de variação de grandezas ou de interpretação de gráficos. Vários itens de matemática podem ser ainda desenvolvidos a partir de problemas físicos ou químicos, gerando uma desejável relação entre as

---

<sup>35</sup> THORNDIKE, R. L. *apud* HALADYNA, Thomas M. **Developing and validating multiple-choice test items**. 3. ed. Mahwah, NJ: Lawrence Erlbaum Associates, 2004. p. 65.

disciplinas. Estas preocupações estão intimamente relacionadas com a necessidade de criar itens que valorizem o raciocínio diante de situações interessantes, deixando menos espaço para a memorização de fórmulas e nomenclaturas. Nesta linha, alguns manuais sugerem ao elaborador que sopesse a conveniência de incluir fórmulas e definições no texto-base<sup>36</sup>. O elaborador deve tentar ao máximo, portanto, estabelecer conexões com outras áreas de conhecimento. Por isso, buscar inspiração em livros didáticos não é uma boa medida para gerar itens criativos e que provoquem interesse no examinado. Vale ainda dizer que, uma vez que se saiba que os itens livrescos e a adoção dos descritores como orientadores dos currículos são temeridades importantes associadas às avaliações em larga escala, não se pode deixar de oferecer um contraponto interessante à indesejada influência apontada das avaliações sobre os currículos. Trata-se de considerar que, se avaliações em larga escala, ao longo do tempo, exigirem do examinado certas práticas diferentes do tradicionalmente encontrado na escola e nos materiais didáticos, pode ser que os currículos sejam afetados positivamente, ou seja, pode ser que novas formas de ensino comecem a ser privilegiadas, em nome do bom resultado que se deseja obter na avaliação em larga escala. É o que registraram Ruben Klein e Nilma Santos Fontanive:

Não se pode perder de vista que, ainda que a avaliação deva refletir o que é ensinado nas escolas, ela deve também indicar caminhos de renovação da prática escolar, não se restringindo, portanto, ao que se costuma definir como “currículo mínimo”<sup>37</sup>.

---

<sup>36</sup> Contra esta recomendação é bom observar que um dos supostos teóricos da Teoria da Resposta ao Item (TRI) é a chamada “independência local dos itens”. Assim, um item de um teste não poderia, em tese, conter informações, como fórmulas e definições, que ajudassem os examinados a escolher a opção de resposta verdadeira em outros itens do mesmo teste. Acerca da independência local dos itens, o leitor pode encontrar mais informações em ANDRIOLA, Wagner Bandeira. *Psicometria moderna: características e tendências. Estudos em avaliação educacional*, São Paulo, v. 20, n. 43, maio/ago. 2009. p. 327. Disponível em: <<http://www.fcc.org.br/pesquisa/publicacoes/eae/arquivos/1496/1496.pdf>>. Acesso em: 06 fev. 2014.

<sup>37</sup> FONTANIVE, Nilma Santos; KLEIN, Ruben. Avaliação em larga escala: uma proposta inovadora. *Em aberto*, Brasília, v. 15, n. 66, abr./jun. 1995. p. 29-34. Disponível em:

O item X2, apesar de tratar de um descritor bastante convencional, foge da abordagem normalmente utilizada pelos livros didáticos e é orientado ao conhecimento. Ele pode ser considerado um exemplo de abordagem criativa que tem completa relação com a habilidade veiculada pelo descritor.

<b>Identificação</b>	X2	<b>Disciplina</b>	Matemática	<b>Gabarito</b>	C
<b>Descritor</b>	Identificar e utilizar noções de juros compostos em situações-problema.				
<b>Texto-base</b>					
Muitos clientes que recebem chamadas telefônicas do setor de cobranças de uma empresa querem saber qual o valor atualizado de um pagamento atrasado. Para fazer este cálculo, deve ser aplicada uma taxa de juros mensal, em regime de juros compostos. O funcionário que efetua as chamadas telefônicas tem, para auxiliar seu trabalho, a tabela mostrada a seguir, que cobre até seis meses de atraso.					
QUANTIDADE DE MESES		FATOR DE MULTIPLICAÇÃO			
1		1,0350			
2		1,0712			
3		1,1087			
4		1,1475			
5		1,1877			
6		1,2293			
Para calcular o valor devido pelo cliente, ele multiplica, usando uma calculadora comum, o valor original da dívida do cliente pelo “fator de multiplicação” da tabela, de acordo com o número de meses de atraso.					
<b>Comando</b>					
A taxa de juros mensal praticada pela empresa é					
<b>Opções de resposta</b>					
A. 1,0350%.					
B. 1,2293%.					
C. 3,5000%.					
D. 7,1200%.					

Há de se ter ainda especial preocupação quanto à contextualização. Em uma primeira aceção, para alguns, ela é indispensável, de modo que todos os itens devem ter uma situação-problema baseada em fatos do cotidiano, conforme

<<http://emaberto.inep.gov.br/index.php/emaberto/article/viewFile/995/899>>. Acesso em: 05 fev. 2014.

ocorreu no item X2. Dentre os que se alinham à necessidade imperiosa da contextualização, há os que buscam diferenciar o contexto do pretexto: não se trataria de simplesmente adicionar algumas pobres informações que tornariam o item um pouco mais parecido com uma situação do dia a dia, mas apresentar ao avaliado situações de fato oriundas de seu contexto cultural. Nesta acepção pedagógica, um item que se inicia com “um professor pediu aos seus alunos que calculassem...” é mero pretexto, não é verdadeira contextualização. Ainda de acordo com a ótica dos que defendem uma contextualização baseada em situações do cotidiano, são também rejeitados certos contextos fantásticos ou artificiais, que apresentam situações exageradas ou atípicas. O item X3 é um exemplo deste tipo, pois a atividade determinada pelo comando é estranha demais para encontrar abrigo na realidade.

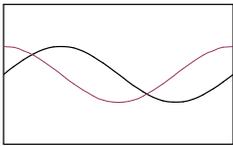
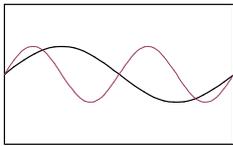
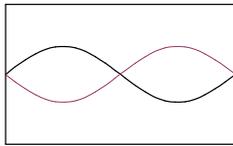
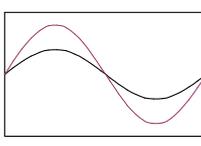
<b>Identificação</b>	X3	<b>Disciplina</b>	Matemática	<b>Gabarito</b>	C
<b>Descritor</b>	Identificar padrões numéricos ou princípios de contagem.				
<b>Texto-base</b>					
Em um berçário, há sete leitos para os bebês, que são três meninos e quatro meninas. Os leitos dos bebês ficam alinhados, um ao lado do outro, como na figura abaixo.					
					
A enfermeira responsável pelo berçário deseja organizar os leitos de maneira que exista um menino em cada extremidade.					
<b>Comando</b>					
O número de formas que a enfermeira pode organizar os leitos do berçário é					
<b>Opções de resposta</b>					
A. 5.					
B. 30.					
C. 720.					
D. 5.040.					

Ainda tratando da contextualização, há uma segunda acepção segundo a qual pode haver uma compreensão mais amena de sua natureza, sendo admissível que itens tenham uma complexidade interna, mas não estejam realmente inspirados em situações sociais. Ou seja, admite-se um item relativamente

inovador cujo contexto é restrito à sua área de conhecimento; no entanto são rejeitados itens mais simples e tidos como livrescos, que se restringem a apresentar diretamente um problema tipicamente escolar.

Por fim, em uma terceira acepção, podem ser admitidos itens que tratem diretamente do descritor, sem preocupações maiores com a contextualização, que é tida somente como desejável, sendo elegível de acordo com a opinião e a criatividade dos autores dos itens, como o item X9, que surgirá mais a frente.

No sentido de elucidar um pouco mais estas linhas de pensamento, o item X4 pode ser usado como apoio.

<b>Identificação</b>	X4	<b>Disciplina</b>	Matemática	<b>Gabarito</b>	B
<b>Descritor</b>	Compreender a natureza cíclica das funções trigonométricas.				
<b>Texto-base</b>					
<p>Considere as funções reais de variáveis reais <math>f(x) = \text{sen}(x)</math> e <math>g(x) = \text{sen}(2x)</math>, com <math>0 \leq x \leq 2\pi</math>, e observe cada um dos esboços de gráfico a seguir.</p>					
					
I		II		III	
					
				IV	
<b>Comando</b>					
O esboço que mostra as representações gráficas sobrepostas das funções $f(x)$ e $g(x)$ é					
<b>Opções de resposta</b>					
A. I.					
B. II.					
C. III.					
D. IV.					

Na primeira acepção, o item X4 deve ser rejeitado de pronto, por não apresentar qualquer referência ao contexto social. De acordo com a segunda acepção, o item X4 pode ser aceito, já que existe um cenário – ainda que puramente matemático – em que o avaliado deve se referenciar para apresentar a opção de resposta correta. Por fim, na terceira acepção, o item X4 está perfeitamente aceitável, já que atende completamente o descritor.

Por meio do exemplo apresentado no item X4, é possível se verificar que o tema da contextualização tem um nível de subjetividade que exige do produtor de itens mais judicioso que obtenha informações precisas acerca do assunto antes de iniciar seu trabalho, seja por meio da leitura atenta do manual de produção e revisão de itens de múltipla escolha e da observação dos exemplos eventualmente nele incluídos, seja por meio do esclarecimento do assunto no treinamento que deve preceder a primeira etapa de produção. A questão da contextualização é sensível por remeter à discussão acerca da natureza abstrata ou concreta da matemática, da sua orientação pura ou aplicada. A favor dos que julgam exagerada a exigência de contextualização, vale lembrar que foram vários os momentos históricos em que um estudo matemático apenas teórico mostrou sua utilidade somente séculos depois, como a aplicações das cônicas de Apolônio feitas por Newton<sup>38</sup>. Por outro lado, em defesa da preocupação com a contextualização, é bom considerar que a chamada “Matemática Moderna”, fulgurante nas décadas de 1960 e 1970, acabou exagerando na conceituação, em detrimento da manipulação e a das aplicações, o que afastava a matemática do dia a dia<sup>39</sup>. Ao largo da discussão, o autor de itens deve, se produzindo para alguma instituição, abandonar suas convicções pessoais que se referem à dicotomia apresentada e se alinhar à proposta técnica e pedagógica escolhida, que preferencialmente deve estar explicitada no manual de produção de itens de múltipla escolha.

Ainda no que tange à contextualização, existem descritores que são perfeitamente compatíveis com essa exigência. Por exemplo, no caso do descritor “identificar e utilizar noções de juros compostos em situações-problema” há inúmeras situações do cotidiano que podem servir para o texto-base de um item, como aquelas relacionadas a investimentos, empréstimos ou financiamentos. Entretanto, há descritores que praticamente impedem uma boa

---

<sup>38</sup> GARBI, Gilberto Geraldo. **A rainha das ciências**: um passeio histórico pelo maravilhoso mundo da matemática. 3. ed. rev. e ampl. São Paulo, SP: Livraria da Física, 2009. p. 109.

<sup>39</sup> LIMA, Elon Lages. **Matemática e ensino**. 3. ed. Rio de Janeiro, RJ: Sociedade Brasileira de Matemática, 2007. p. 141.

contextualização. São aqueles que tratam de operações mentais mais simples, que se referem a habilidades mais voltadas para o aspecto operacional do que para análises mais sofisticadas. É o caso de “efetuar cálculos com números reais”, ainda mais quando se realça que é necessário que o âmbito da atividade cognitiva do examinado não ultrapasse o previsto neste descritor, ou seja, o item não pode se tornar associado ao descritor “analisar, interpretar e resolver situações-problema com números reais, envolvendo diferentes significados das operações (adição, subtração, multiplicação, divisão, potenciação e radiciação)”. Se a preocupação com a contextualização for extrema, uma solução não muito elegante é fazer uso de uma situação do cotidiano apenas como introdutória, já que não parece ser possível, como seria indicado, fazer com que a contextualização permeasse todo o item. É o que alguns, ironicamente, chamam de *pretextualização*, já que a situação do cotidiano terá sido usada apenas como um pretexto para apresentar o cálculo que deve ser efetuado pelo examinado, como ocorre no item X5.

<b>Identificação</b>	X5	<b>Disciplina</b>	Matemática	<b>Gabarito</b>	A
<b>Descritor</b>	Efetuar cálculos com números reais.				
<b>Texto-base</b>					
Para calcular a resistência total ( $R_T$ ) em um circuito elétrico, um técnico precisa efetuar o seguinte cálculo:					
$R_T = \frac{1}{\frac{1}{20} + \frac{1}{30}}$					
<b>Comando</b>					
O valor numérico encontrado para a resistência total é					
<b>Opções de resposta</b>					
A. 12.					
B. 25.					
C. 50.					
D. 60.					

Há discussão acerca dos nomes dados aos personagens que aparecem nos itens. Para alguns, podem ser usados nomes próprios e, neste caso, pode ser uma boa ideia usar nomes em ordem alfabética, como Ana, Bianca, Carla e

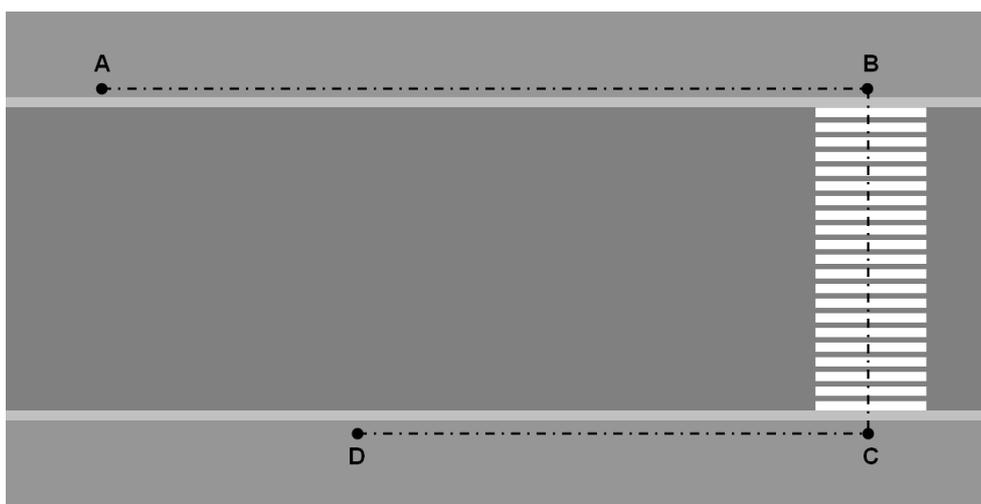
Daniele, como ocorre no item X7, que será apresentado mais a frente. Para outros, nomes próprios devem ser evitados e os personagens devem ser como “um trabalhador rural”, “um maratonista” ou “dois jogadores”. O produtor deve ser informado por meio do manual de produção e revisão de itens de múltipla escolha ou no treinamento precedente ao início do trabalho acerca de eventuais limitações quanto aos nomes que podem ser usados no texto-base de cada item.

Como comentado anteriormente, em muitos itens é necessário incluir figuras, como formas geométricas ou gráficos cartesianos. Em vários casos, a situação é bastante específica, de modo que não é possível se encontrar a figura desejada em um banco de imagens ou por meio de uma página de busca na internet. Nestes casos, o produtor deve criar a imagem. Para tanto, pode usar as ferramentas incluídas em um editor de textos ou, no caso de gráficos, usar uma planilha eletrônica, como ocorreu no item X4. Pode ainda fazer uso de um software específico de geometria dinâmica, como o GeoGebra.

Foge ao escopo deste texto discorrer em detalhe sobre o uso destas possibilidades. Entretanto, o que importa realçar é que o autor de itens verá seu trabalho ficar muito mais produtivo e rico se puder usar as possibilidades gráficas de uma planilha eletrônica. Em especial, o uso de um software como o GeoGebra pode tornar o trabalho muito mais rápido e provocar uma aparência muito profissional, como ocorre no item X14, que será apresentado posteriormente. O investimento de tempo no aprendizado do GeoGebra tem um enorme retorno na rapidez com que novos itens são produzidos a partir do uso deste software.

Apenas a título de ilustração, a figura apresentada a seguir mostra uma estrada e uma recomendável trajetória – indicada por uma linha pontilhada – de uma pessoa que está no ponto A e deseja atravessar a estrada, indo até o ponto D. Dadas as medidas AB, BC e CD, o avaliado deve calcular a medida AD, que seria a trajetória mais curta, porém insegura. O avaliado, é claro, deve usar o Teorema de Pitágoras para solucionar o problema. O que importa aqui é que a

imagem foi criada usando uma planilha eletrônica, apenas por meio da escolha adequada do tamanho das células e dos tons de cinza de cada uma delas. A planilha teve então desabilitada a exibição das linhas de grade. A escolha da ferramenta adequada para o desenho, portanto, tem influência decisiva na qualidade da imagem produzida e na rapidez do trabalho.



De maneira universal, os manuais de produção e revisão de itens de múltipla escolha defendem que o texto-base deve conter exatamente as informações necessárias para que o avaliado encontre a opção de resposta correta. Isso significa dizer que não devem faltar informações – o que é óbvio –, mas também não devem existir informações excedentes. Essa posição poderia ser criticada, já que, no cotidiano, faz parte das competências necessárias ao mundo do trabalho e da ciência selecionar as informações úteis para tratar um problema, uma vez que nunca há garantia de que não existirá falta ou excesso de dados. Nessa perspectiva, a apresentação de problemas que sempre têm exatamente os dados necessários se afasta da realidade social para a qual se busca preparar os estudantes. Respostas possíveis a esta crítica incluem o reduzido tempo destinado ao trato de cada item por parte do avaliado e a dificuldade de se verificar se a causa de erros na escolha da opção de resposta certa se deve à falta de domínio da habilidade veiculada pelo descritor ou simplesmente por confusão a respeito dos dados do problema. Em se aceitando a exigência de incluir exatamente os dados necessários – sem mais

nem menos – no texto-base, não seria cabível, por exemplo, solicitar ao avaliador que calculasse a área de um retângulo, fornecendo as medidas da base, da altura e da diagonal. A informação desnecessária – neste exemplo, a medida da diagonal – jamais poderia ser apresentada. Abaixo é apresentado o item X6, outro exemplo com informação desnecessária: o salário mensal do cliente é totalmente imprestável para a solução do problema.

<b>Identificação</b>	X6	<b>Disciplina</b>	Matemática	<b>Gabarito</b>	C
<b>Descritor</b>	Identificar e utilizar noções de juros compostos em situações-problema.				
<b>Texto-base</b>					
Um cliente de uma instituição financeira, <b>que recebia um salário mensal de R\$ 3.000,00 [INFORMAÇÃO DESNECESSÁRIA]</b> , contratou um empréstimo de R\$ 10.000,00 sob o regime de juros compostos e com taxa de juros de 10% ao mês.					
Dois meses após receber o dinheiro do empréstimo, o cliente conseguiu fazer um primeiro pagamento de R\$ 1.000,00. Dois meses após este pagamento, pagou o que restava do empréstimo com um segundo pagamento.					
<b>Comando</b>					
O valor do segundo pagamento feito pelo cliente foi					
<b>Opções de resposta</b>					
A. R\$ 13.000,00.					
B. R\$ 13.200,00.					
C. R\$ 13.431,00.					
D. R\$ 13.641,00.					

No texto-base em particular, mas também em todo o item, o produtor deve dar atenção às prescrições da instituição responsável pela formação do banco quanto à formatação, que normalmente são encontradas no manual de produção e revisão de itens de múltipla escolha. Isto se aplica, dentre outros aspectos, ao tipo de fonte dos caracteres usados no editor de textos, ao uso de negrito e itálico, à pontuação, ao uso de aspas duplas e simples e aos recuos, espaçamentos e alinhamentos. No caso de disciplinas como a matemática, que exigem o uso de equações, fórmulas, frações e sinais especiais, o manual de produção e revisão de itens de múltipla escolha deve apresentar instruções que padronizem a inclusão destes elementos no texto-base e nas demais partes integrantes do item. Este manual não apresenta prescrições voltadas para a

formatação dos itens a serem produzidos pelos autores, já que elas têm caráter exclusivamente voltado à forma, sem relação com o conteúdo. Por isso, é imprescindível aos produtores que verifiquem o manual de produção e revisão de itens de múltipla escolha fornecido pela instituição responsável pela construção do banco de itens, se for o caso, para que sejam escritos itens compatíveis com os rígidos padrões necessários, que podem variar de acordo com a instituição.

### **III. 7. COMANDO**

O comando é também referido por alguns manuais como “enunciado”. Ele deve apresentar, com absoluta clareza, o que é esperado do avaliado. Ele deve ser preferencialmente curto e direto.

Alguns manuais de produção e revisão de itens de múltipla escolha permitem o uso de uma pergunta no comando, enquanto outros vetam.

De maneira geral, os manuais proíbem o uso de negativas no comando, como “não” ou “exceto”.

É proibido também que se inste o avaliado a escolher a opção de resposta mais correta, ou seja, dentre as opções de resposta deve haver necessariamente uma resposta verdadeira e as restantes devem ser indiscutivelmente falsas.

Alguns manuais rejeitam opções de resposta que exijam do avaliado que estabeleça relações entre proposições apresentadas no texto-base e opções de resposta. Assim, de acordo com esta perspectiva, não seria adequado, por exemplo, apresentar três proposições ao avaliado no texto-base seguidas de

opções de resposta como “apenas as proposições I e II estão corretas”. Estes itens são eventualmente referidos como “itens de múltipla escolha composta”.

Existe divergência acerca dos itens de múltipla escolha comumente chamados de “verdadeiro ou falso”. Alguns manuais os rejeitam, outros os aceitam, desde que não muito numerosos. Estes itens exigem do avaliado maior dispêndio de tempo, uma vez que se faz necessário analisar todas as opções de resposta, conforme no item X7.

<b>Identificação</b>	X7	<b>Disciplina</b>	Matemática	<b>Gabarito</b>	D												
<b>Descritor</b>	Resolver problema com dados apresentados em tabelas ou gráficos.																
<b>Texto-base</b>																	
Cinco participantes de um concurso público tiveram pontuações conforme o gráfico de barras a seguir.																	
<table border="1"> <caption>Dados do Gráfico de Barras</caption> <thead> <tr> <th>Participante</th> <th>Pontuação</th> </tr> </thead> <tbody> <tr> <td>Ana</td> <td>25</td> </tr> <tr> <td>Bianca</td> <td>26</td> </tr> <tr> <td>Carla</td> <td>27</td> </tr> <tr> <td>Daniele</td> <td>28</td> </tr> <tr> <td>Eliane</td> <td>30</td> </tr> </tbody> </table>						Participante	Pontuação	Ana	25	Bianca	26	Carla	27	Daniele	28	Eliane	30
Participante	Pontuação																
Ana	25																
Bianca	26																
Carla	27																
Daniele	28																
Eliane	30																
Eliane foi a participante com melhor resultado, tendo obtido 30 pontos.																	
<b>Comando</b>																	
Do gráfico de barras se infere que																	
<b>Opções de resposta</b>																	
A. Ana obteve 25 pontos.																	
B. Bianca obteve 26 pontos.																	
C. Carla obteve 27 pontos.																	
D. Daniele obteve 28 pontos.																	

No item X7, o avaliado deve testar a validade das quatro opções de resposta para decidir qual é a verdadeira. O produtor deve estar informado se itens

como este são permitidos pelo manual de produção e revisão de itens de múltipla escolha. Na dúvida, deve evitar esta estratégia. Uma maneira fácil de testar se o item de múltipla escolha é do tipo “verdadeiro ou falso” é simplesmente observar se ele pode ser resolvido sem olhar as suas opções de resposta. Se isto não for possível, provavelmente o item pode ser classificado como “verdadeiro ou falso”. Em oposição, se é possível resolver o item sem sequer ler as opções de resposta, ele jamais será tomado como “verdadeiro ou falso”.

Uma maneira de contornar esse problema no item X7 seria reduzir o número de participantes no texto-base para três, como Ana, Bianca e Carla. Esta última seria a participante de pontuação máxima e o comando seria alterado para algo do tipo: “as pontuações de Ana e Bianca, respectivamente, são:”.

Ainda tratando de itens de múltipla escolha que são conhecidos como “verdadeiro ou falso”, mesmo que eles sejam aceitos por alguns manuais de produção e revisão de itens de múltipla escolha, ainda que limitadamente, eles não se confundem com itens nos quais as opções de respostas apresentam proposições totalmente independentes, como se fossem, cada uma delas, um item separado, configuração tida como inaceitável. Um indicador de item como este é um comando indesejado do tipo “é correto afirmar que”.

Alguns manuais de produção e revisão de itens de múltipla escolha enfatizam que todas as informações estejam presentes no texto-base, de modo que não exista qualquer dado no comando. Nesta ótica, é rejeitado um comando do tipo “se a produção da fábrica aumentar em 10%, o número de automóveis produzidos será”. Para que item seja considerado adequado, a informação a respeito do aumento de 10% na produção deve ser transportada para o texto-base e o comando deve ficar restrito a: “o número de automóveis produzidos será”.

### III. 8. OPÇÕES DE RESPOSTA

Dentre as opções de resposta<sup>40</sup>, repise-se, haverá somente uma verdadeira, sendo as demais falsas. Ainda que falsas, nenhuma das respostas erradas pode ser absurda. As respostas falsas são denominadas distratores.

Haladyna afirma a construção de distratores é trabalhosa e se constitui a parte mais difícil da produção de itens de múltipla escolha<sup>41</sup>. Portanto, o produtor, principalmente o iniciante, deve manter toda sua atenção e zelo nesta parte do processo de construção de itens.

Haladyna afirma ainda que a chave para o desenvolvimento de distratores é a plausibilidade<sup>42</sup>. Um item deve ser corretamente respondido por aqueles que dominam a habilidade veiculada pelo descritor e respondido equivocadamente por aqueles que não a dominam. Ou seja, um distrator plausível parece uma resposta correta para aqueles não conseguem atingir o objetivo do descritor.

É importante realçar, no entanto, que afirmar que as opções de resposta devem ser parecidas não significa dizer que devem ser esteticamente aproximadas. A similaridade está muito mais afeita à dificuldade do examinado que não domina a habilidade veiculada no descritor em diferenciar o certo do

---

<sup>40</sup> Há quem afirme que “opções de resposta” não podem ser chamadas de “alternativas”, pois a expressão “alternativa” estaria ligada à ideia de apenas duas possibilidades. Entretanto, o dicionário Houaiss apresenta um dos significados para “alternativa” como “uma de duas ou mais possibilidades pelas quais se pode optar”. O mesmo dicionário apresenta “opção” como “uma de duas ou mais possibilidades pelas quais se pode optar; alternativa”. Evitando a polêmica, este manual usará a expressão “opções de resposta” para se referir ao conjunto de escolhas possíveis do avaliado.

<sup>41</sup> HALADYNA, Thomas M. **Developing and validating multiple-choice test items**. 3. ed. Mahwah, NJ: Lawrence Erlbaum Associates, 2004. p. 120.

<sup>42</sup> HALADYNA, Thomas M. **Developing and validating multiple-choice test items**. 3. ed. Mahwah, NJ: Lawrence Erlbaum Associates, 2004. p. 120.

errado do que à aparência gráfica das opções de resposta. Ou seja, plausibilidade não é sinônimo de semelhança.

Por isso, os distratores devem conter os erros comuns cometidos pelos avaliados. O produtor deve tentar ao máximo se colocar no lugar dos examinados, buscando reconstruir o possível raciocínio do avaliado e tentar identificar seus caminhos falhos. Neste sentido, não há qualquer dúvida em afirmar que a experiência docente é indispensável para que se forme um referencial dos erros mais comuns cometidos pelos estudantes. Os professores com grande prática sequer chegam a se surpreender com as falhas dos estudantes: elas se repetem de maneira mais ou menos uniforme enquanto os objetivos de aprendizagem não são atingidos. Por outro lado, aquele que domina a disciplina e não está acostumado com os erros dos estudantes não consegue intuir com facilidade quais as falhas mais frequentes. Um par de exemplos: o professor experiente sabe que muitos estudantes afirmarão que  $82 - 29 = 67$  ou que a raiz quadrada de 64 é 32. Entretanto, um item não deve ser construído com o objetivo de provocar o erro, ou seja, não deve ser capcioso ou dúbio, como quem torce para que o confuso avaliado seja vencido pela desatenção. As dificuldades artificiais são intoleráveis.

As opções de resposta devem formar um conjunto homogêneo, sem que uma se destaque das demais. Se as opções forem textos, devem ter seus tamanhos aproximadamente iguais. Se forem números, não deve haver discrepância relevante entre eles. Por exemplo, o conjunto de opções de resposta formado por 0,2; 5; 7 e 9 não é bem escolhido, uma vez que o primeiro número se destaca dos demais e o número de casas decimais não é o mesmo. Em outro exemplo, o item X8 apresenta um conjunto de opções de resposta não homogêneo, uma vez que uma das opções – o gabarito – apresenta duas respostas possíveis, sendo a única opção diferente das demais. Esta discrepância pode acabar atraindo alguns avaliados que não dominam a habilidade veiculada no descritor e escolhem a opção mais por sua diferença em relação às demais. Uma maneira de contornar o problema é simplesmente manter na opção D apenas “a média” ou “a mediana”.

<b>Identificação</b>	X8	<b>Disciplina</b>	Matemática	<b>Gabarito</b>	D
<b>Descritor</b>	Avaliar propostas de intervenção na realidade utilizando conhecimentos de estatística e probabilidade.				
<b>Texto-base</b>					
A maior parte dos conjuntos de dados numéricos mostra uma tendência em torno de um ponto central, como na situação abaixo descrita. Há um conjunto de várias anotações a respeito dos tempos, em minutos, que uma loja de departamentos demora no tratamento das reclamações telefônicas de seus clientes. Pode ser conveniente procurar a quantidade de minutos para a qual tendem os diversos tempos de atendimento telefônico da loja de departamentos.					
<b>Comando</b>					
Uma medida que pode apresentar a tendência central procurada é					
<b>Opções de resposta</b>					
A. a variância.					
B. a amplitude.					
C. o desvio-padrão.					
D. a média ou a mediana. <b>[ÚNICA OPÇÃO COM DUAS MEDIDAS]</b>					

É importante ainda estabelecer uma ordem na apresentação das opções de resposta, que pode ser numérica crescente ou decrescente ou, ainda, alfabética, dependendo do item. No caso de textos, pode ser crescente no que se refere ao número de caracteres de cada opção, com a mostrada no item X8. Entretanto, existem situações nas quais os examinados são testados por meios digitais que apresentam as opções de resposta randomicamente. Neste caso, os autores devem ser informados sobre o fato, para que tomem as medidas adequadas durante a construção das opções de resposta, uma vez que há situações em que a exibição em determinada ordem pode ser indesejada. Vale acrescentar ainda que, em geral, uma opção de resposta, seja verdadeira ou falsa, não deve ter palavras como “nada”, “tudo”, “sempre” ou “nunca”, as quais podem destacá-la em relação às demais e que venham a atrair a atenção do avaliado.

É bom considerar que a apresentação das opções de resposta pode despertar no avaliado a busca por um atalho, quando este não dominar plenamente a habilidade veiculada por meio do descritor. Em alguns casos, isso pode até ser

mais fácil do que tentar resolver o item da maneira que o seu autor pretendia que fosse escolhida pelo examinado. Isso ocorre, com destaque, quando é possível testar todas as opções de resposta até se encontrar a que é correta. Por exemplo, se o objetivo é resolver uma equação e são apresentadas possíveis raízes nas opções de resposta, o avaliado pode substituir a incógnita pelos números apresentados, até encontrar a opção verdadeira. Isso pode ser contornado como no exemplo apresentado no item X9, no qual é necessário resolver as equações para encontrar a opção de resposta verdadeira. Apesar de a solução do item envolver mais uma etapa operacional, a saber, a contagem de números inteiros no conjunto  $S$ , o examinado fica impedido de resolver o item apenas por meio de testes.

<b>Identificação</b>	X9	<b>Disciplina</b>	Matemática	<b>Gabarito</b>	B
<b>Descritor</b>	Resolver equações do segundo grau completas e incompletas.				
<b>Texto-base</b>					
As raízes da equação $2x^2 - x - 1 = 0$ são $x_A$ e $x_B$ e as raízes da equação $3y^2 - 11y - 4 = 0$ são $y_A$ e $y_B$ .					
No conjunto $S = \{x_A, x_B, y_A, y_B\}$ há quatro números reais diferentes.					
<b>Comando</b>					
Quantos números inteiros há no conjunto $S$ ?					
<b>Opções de resposta</b>					
A. 1.					
B. 2.					
C. 3.					
D. 4.					

Há certo padrão no sentido de que o conjunto de opções de resposta seja formado por quatro elementos quando o item é aplicável ao nível fundamental e por cinco elementos quando os avaliados são integrantes do nível médio ou superior. Entretanto, em testes integrantes de olimpíadas, por exemplo, é comum um conjunto de cinco opções de resposta para cada item mesmo no nível fundamental, já que se deseja que os itens tenham um nível de dificuldade desafiador. Apesar disso, alguns estudos apontados em maior

detalhe por Haladyna<sup>43</sup> concluíram que a ótima configuração parece ser um conjunto de três opções – uma resposta correta e dois distratores. Esta quantidade de opções seria um limite natural para a maior parte dos itens de múltipla escolha. Isto se sustentaria no fato de que se revela, após análise estatística das respostas dadas pelos examinados aos itens, que alguns distratores simplesmente não servem para separar aqueles que dominam a habilidade do descritor daqueles que não a possuem. Conforme revela a prática e Haladyna corrobora, muitas vezes os produtores têm grande dificuldade em achar uma quarta ou quinta opção de resposta, porque tipicamente ela não existe. Sobretudo no que se refere à desejada homogeneidade das opções de resposta, problemas adicionais serão enfrentados a respeito da simetria ou do paralelismo das estruturas quando são necessárias cinco opções de resposta. É por isso que deve haver alguma tolerância por parte dos revisores no sentido de que nem todos os distratores, em todos os itens, são, de fato, distratores propriamente ditos, mas algumas das opções de resposta falsas podem surgir apenas como uma falha de raciocínio já tratada em outro distrator ou, ainda, como um complemento estético das outras opções de resposta.

Uma armadilha fácil que deve ser evitada pelos produtores consiste na inclusão de uma informação extra ou de uma complicação adicional no texto-base, de modo que, considerando um erro do avaliado em função da utilização destes artifícios incluídos quando da criação dos distratores, surge um equívoco que, em verdade, não discrimina se o avaliado domina ou não domina a habilidade do descritor.

O produtor deve estar cauteloso para não incluir uma opção de resposta que contenha o conteúdo de outra. Se este erro for cometido, o avaliado atento eliminará facilmente as duas opções de resposta, se estiver certo que a menos restritiva entre as duas é falsa. Afinal, por implicação lógica, a opção mais

---

<sup>43</sup> HALADYNA, Thomas M. **Developing and validating multiple-choice test items**. 3. ed.

restritiva será falsa também. Por exemplo, se a opção A afirma que um número procurado é maior que 10 e a opção B dispõe que tal número é maior que 20, o avaliado, percebendo que, se é falso que um número é maior que 10, então é falso que ele é maior que 20, concluirá que se a opção A é falsa, então a opção B também é. A falha pode ser ainda mais grave: o item pode ter duas respostas possíveis. O exemplo apresentado no item X10 é uma ocorrência deste tipo, no qual as opções de resposta A e B estão ambas corretas, em um contexto de definições no qual se considera que toda elipse é uma circunferência.

<b>Identificação</b>	X10	<b>Disciplina</b>	Matemática	<b>Gabarito</b>	A <b>[TAMBÉM B!]</b>
<b>Descritor</b>	Reconhecer as seções cônicas e seus elementos.				
<b>Texto-base</b>					
Em um plano foram marcados todos os pontos que distam 10 cm do ponto P.					
<b>Comando</b>					
A figura que surgiu no plano foi uma					
<b>Opções de resposta</b>					
A. circunferência.					
B. elipse.					
C. hipérbole.					
D. parábola.					

Os manuais de produção e revisão de itens de múltipla escolha raramente comentam sobre isso, mas é conveniente buscar, em um lote de itens, que as respostas verdadeiras sejam distribuídas de maneira mais ou menos uniforme em relação às letras que o avaliado deve ser escolher. Assim, se há quatro opções de resposta – A, B, C e D –, é bom que o número de gabaritos seja de aproximadamente 25% para cada letra.

O uso de planilhas eletrônicas pode ser extremamente útil para tornar a produção de distratores de um item mais rápida. Além disso, pode servir para a criação de um conjunto de opções de resposta que tornem um item mais elegante. Esta característica, a elegância, não é essencial na produção, mas é desejável. O uso de planilhas eletrônicas pode também servir para tornar o

---

Mahwah, NJ: Lawrence Erlbaum Associates, 2004. p. 112-113.

conjunto de opções de resposta mais homogêneo. Um par de exemplos pode ser útil para esclarecer a vantagem do uso de planilhas eletrônicas. O primeiro deles é apresentado no item X11.

<b>Identificação</b>	X11	<b>Disciplina</b>	Matemática	<b>Gabarito</b>	C
<b>Descritor</b>	Calcular medidas de tendência central.				
<b>Texto-base</b>					
Oito analistas financeiros apresentaram previsões para o valor futuro de uma ação de uma empresa, conforme a tabela abaixo.					
		Analista 1	R\$ 33,00		
		Analista 2	R\$ 33,00		
		Analista 3	R\$ 33,00		
		Analista 4	R\$ 31,00		
		Analista 5	R\$ 37,00		
		Analista 6	R\$ 40,00		
		Analista 7	R\$ 41,00		
		Analista 8	R\$ 40,00		
<b>Comando</b>					
A mediana dos valores apresentados é					
<b>Opções de resposta</b>					
A. R\$ 33,00.					
B. R\$ 34,00.					
C. R\$ 35,00.					
D. R\$ 36,00.					

No item X11, as opções de resposta são quantias que não apresentam parte fracionária e estão em ordem crescente. Além disso, há diferença de exatamente um real entre duas opções adjacentes. Enfim, há elegância. Além disso, a apresentação das opções de resposta não dá qualquer pista que permita ao avaliado escolher ao acaso qualquer delas.

Para calcular a mediana é mandatório, em primeiro lugar, ordenar os valores. Como o número de valores é par, é necessário, em segundo lugar, calcular a média entre os dois elementos centrais da lista ordenada, isto é, deve-se calcular a média entre o quarto e o quinto valores financeiros do item, considerando as quantias já em ordem.

A opção A é a moda (o valor mais frequente) dos números da lista, a opção B representaria uma tentativa falha de obter a mediana sem colocar os números em ordem (ou seja, é a média entre o quarto e o quinto valores da lista apresentada, ainda não ordenada), a opção C é a mediana dos valores apresentados e a opção D é a média dos números da lista.

Mas como chegar a um conjunto de números a serem apresentados elegantemente nas opções de resposta, como aqueles que constam no item X11? É certo que isto pode ser conseguido por meio de tentativas, usando papel e caneta, por exemplo. Pode-se também tentar criar um método matemático para isto. Entretanto, o caminho mais curto é simplesmente digitar uma lista de números em oito células de uma planilha eletrônica e inserir as funções apropriadas em outras quatro células da planilha. Tais funções devem refletir, além do gabarito, os erros mais comuns dos avaliados, conforme descrito acima. Em seguida, por meio de ajustes nos números da lista, basta observar o resultado apresentado nas células onde estão as funções até que se consiga um conjunto de números satisfatório para as opções de resposta. A figura a seguir apresenta uma sugestão deste procedimento. As células de A1 até A8 apresentam os valores da tabela e as células de A10 até A14 apresentam as funções que geram as opções de resposta. A planilha eletrônica foi configurada para que as funções fossem apresentadas em lugar dos valores calculados.

	A
1	33
2	33
3	33
4	31
5	37
6	40
7	41
8	40
9	
10	=MODO(A1:A8)
11	=MÉDIA(A4:A5)
12	=MED(A1:A8)
13	=MÉDIA(A1:A8)

Um segundo exemplo é apresentado no item X12, ainda no sentido de realçar a utilidade do uso de planilhas eletrônicas.

<b>Identificação</b>	X12	<b>Disciplina</b>	Matemática	<b>Gabarito</b>	B
<b>Descritor</b>	Resolver problema que envolva variações proporcionais, diretas ou inversas entre grandezas.				
<b>Texto-base</b>					
Um supercomputador é usualmente formado por vários processadores que, ao mesmo tempo, executam tarefas muito complexas. Quanto maior o número de processadores em um supercomputador, mais rápida é a execução de uma tarefa.					
Em uma universidade, um supercomputador tinha 16 processadores iguais e conseguia fazer um cálculo complexo em 12 minutos.					
O supercomputador foi modificado e passou a ter 24 processadores idênticos aos que existiam anteriormente.					
<b>Comando</b>					
O tempo para fazer o mesmo cálculo, em minutos, passou a ser					
<b>Opções de resposta</b>					
A. 6.					
B. 8.					
C. 18.					
D. 32.					

Na opção A, o avaliado pode erroneamente concluir que, se houve acréscimo de 50% no número de processadores, deve ter ocorrido uma redução de 50% no tempo de execução da tarefa. A opção B apresenta o resultado do cálculo correto, no qual o avaliado identifica que são variações proporcionais inversas entre as grandezas. Ao escolher a opção C, o avaliado pode, em um engano, ter considerado que são variações proporcionais diretas entre as grandezas. Por fim, a opção D é a escolhida por um avaliado que, sem entender bem as variações das grandezas, tenha calculado o produto entre 16 e 24 e dividido o resultado por 12.

Na elaboração de um item como o apresentado no exemplo X12, o produtor pode encontrar dificuldades em escolher um terno de números para ser usado no texto-base que permita resultados inteiros nas quatro situações descritas acima, relativas ao gabarito e aos três distratores. No caso do item X12, os

números usados foram 16, 24 e 12. Como se pode chegar rapidamente a um terno de números que permita os resultados inteiros, como aqueles apresentados nas opções de resposta? Novamente, o uso de tentativas, em papel e caneta, pode ser uma saída. Ainda seria cabível pensar em critérios de divisibilidade para buscar números adequados. No entanto, mais uma vez, como ocorreu no exemplo anterior, o uso de uma planilha eletrônica pode ser um caminho mais curto. Basta digitar três números em células da planilha e criar as funções adequadas em outras quatro células, uma relacionada ao gabarito e três que devem se referir aos erros mais comuns, conforme descrito anteriormente. A figura apresentada em seguida é uma proposta deste método, na qual as células de A1 até A3 apresentam o terno de números usado no texto-base e as células de A5 até A8 apresentam as funções. Como no exemplo anterior, a planilha eletrônica foi configurada para que as funções fossem apresentadas em lugar dos valores.

	A
1	16
2	12
3	24
4	
5	$= (1 - (A3/A1 - 1)) * A2$
6	$= A1 * A2 / A3$
7	$= A2 * A3 / A1$
8	$= A3 * A1 / A2$

Tendo sido apresentados dois exemplos, é conveniente discorrer mais um pouco sobre o uso de planilhas eletrônicas, observando que, em ambos os exemplos, é possível repetir rapidamente os conteúdos da coluna A em outras colunas e, por meio da digitação de valores diferentes, observar e comparar diversas possibilidades de opções de resposta para criar o item.

A figura a seguir mostra o procedimento descrito para o primeiro exemplo. São apresentadas 16 possibilidades, nas quais apenas o último número da lista varia, na oitava linha. É fácil observar que somente a coluna M apresenta números inteiros diferentes em todas as opções de resposta.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
1	33,00	33,00	33,00	33,00	33,00	33,00	33,00	33,00	33,00	33,00	33,00	33,00	33,00	33,00	33,00	33,00
2	33,00	33,00	33,00	33,00	33,00	33,00	33,00	33,00	33,00	33,00	33,00	33,00	33,00	33,00	33,00	33,00
3	33,00	33,00	33,00	33,00	33,00	33,00	33,00	33,00	33,00	33,00	33,00	33,00	33,00	33,00	33,00	33,00
4	31,00	31,00	31,00	31,00	31,00	31,00	31,00	31,00	31,00	31,00	31,00	31,00	31,00	31,00	31,00	31,00
5	37,00	37,00	37,00	37,00	37,00	37,00	37,00	37,00	37,00	37,00	37,00	37,00	37,00	37,00	37,00	37,00
6	40,00	40,00	40,00	40,00	40,00	40,00	40,00	40,00	40,00	40,00	40,00	40,00	40,00	40,00	40,00	40,00
7	41,00	41,00	41,00	41,00	41,00	41,00	41,00	41,00	41,00	41,00	41,00	41,00	41,00	41,00	41,00	41,00
8	28,00	29,00	30,00	31,00	32,00	33,00	34,00	35,00	36,00	37,00	38,00	39,00	40,00	41,00	42,00	43,00
9																
10	33,00	33,00	33,00	33,00	33,00	33,00	33,00	33,00	33,00	33,00	33,00	33,00	33,00	33,00	33,00	33,00
11	34,00	34,00	34,00	34,00	34,00	34,00	34,00	34,00	34,00	34,00	34,00	34,00	34,00	34,00	34,00	34,00
12	33,00	33,00	33,00	33,00	33,00	33,00	33,50	34,00	34,50	35,00	35,00	35,00	35,00	35,00	35,00	35,00
13	34,50	34,63	34,75	34,88	35,00	35,13	35,25	35,38	35,50	35,63	35,75	35,88	36,00	36,13	36,25	36,38

A próxima figura trata do segundo exemplo apresentado. Novamente há 16 possibilidades, nas quais se verifica que vários ternos são inadequados para o item, havendo muitos casos que resultam em números não inteiros e, eventualmente, há números repetidos. É bom notar também que, em lugar de se escolher o terno formado pelos números 16, 12 e 24, poderia ser usado o terno formado pelos números 20, 30 e 24.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
1	16	16	16	16	18	18	18	18	20	20	20	20	24	24	24	24
2	12	14	16	18	20	22	24	26	28	30	32	34	36	38	40	42
3	24	24	24	24	24	24	24	24	24	24	24	24	24	24	24	24
4																
5	6	7	8	9	13,33	14,67	16	17,33	22,4	24	25,6	27,2	36	38	40	42
6	8	9,333	10,67	12	15	16,5	18	19,5	23,33	25	26,67	28,33	36	38	40	42
7	18	21	24	27	26,67	29,33	32	34,67	33,6	36	38,4	40,8	36	38	40	42
8	32	27,43	24	21,33	21,6	19,64	18	16,62	17,14	16	15	14,12	16	15,16	14,4	13,71

A quantidade de conjuntos de opções de resposta poderia ser muito maior, sem grandes dificuldades. Neste caso, seria mais conveniente organizar os dados de modo que os conjuntos de opções de resposta estivessem dispostos em linhas, ao invés de colunas.

Em certas situações, o uso adequado da função ALEATÓRIO() associada à função INT pode ser muito útil, tornando possível poupar o trabalho de digitação de vários conjuntos de opções de resposta. A função ALEATÓRIO() retorna um número aleatório maior ou igual que zero e menor que um. A função INT, por sua vez, arredonda um número para baixo até o número inteiro mais próximo. Por exemplo, a função composta  $INT(ALEATÓRIO()*10)$  gera um número inteiro entre 0 (inclusive) e 9 (inclusive). Se muitas células foram

adequadamente preenchidas com esta função, é possível criar um conjunto de números aleatórios que substitui um grande conjunto de números que seriam digitados pelo produtor do item. Pressionando a tecla F9, os valores são novamente calculados, ou seja, novos números aleatórios são criados, de modo que novas possibilidades podem surgir para as opções de resposta.

Por fim, é importante realçar que a produção de distratores é, como já antes mencionado, a parte mais difícil e laboriosa quando se deseja criar itens de múltipla escolha. É, ainda, parte integrante do item, tendo importância capital quando ocorre sua validação. Por estes motivos, a atenção aos distratores não deve ser negligenciada. A importância dos distratores poderia ser ainda maior se fossem construídos itens em que a escolha de respostas falsas por parte do examinado gerasse pontuações intermediárias, o que se afastaria da lógica binária, na qual um item vale um se a opção escolhida é a verdadeira e vale zero se for escolhido um dos distratores. Ainda que não se verifiquem, nos dias atuais, aplicações significativas desta modalidade de pontuação gradual quanto aos distratores, pode ser que, no futuro, seu emprego se desenvolva<sup>44</sup>.

### III. 9. COMENTÁRIOS

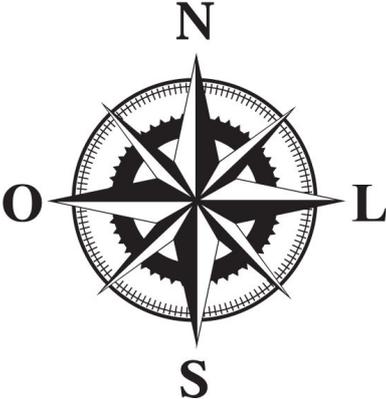
Os manuais de produção e revisão de itens de múltipla escolha costumam exigir do produtor alguns comentários em cada item. Não existe um padrão para tais comentários e, a seguir, serão apresentadas três hipóteses.

---

<sup>44</sup> HALADYNA, Thomas M. **Developing and validating multiple-choice test items**. 3. ed. Mahwah, NJ: Lawrence Erlbaum Associates, 2004. p. 275.

Uma possibilidade consiste em exigir do produtor que escreva um comentário único para cada item com explicações gerais sobre o assunto tratado no descritor. O produtor deve elucidar como se chega à opção de resposta correta, podendo ainda comentar sobre os distratores e os erros que cada um deles representa. Se adequado, pode haver uma pequena introdução, tratando, em linhas gerais, do objetivo veiculado pelo distrator. Em seguida, deve ser apresentado o raciocínio necessário para resolver o item. Se existirem cálculos, eles devem ser apresentados. No caso de plataforma eletrônica, o comentário surge para o examinado após este fornecer a resposta ao item ou, ainda, após o término do teste composto de uma quantidade fixa de itens. Este modelo é apropriado também para itens apresentados impressos, quando integram um teste, situação em que os comentários podem ser divulgados publicamente após o exame ter sido realizado pelos avaliados. Outra possibilidade consiste em uma coletânea de itens apresentados em único volume, sendo que os comentários gerais dirigidos aos examinados podem constar apartados dos itens, geralmente na parte final da publicação. Quando os comentários são dirigidos ao examinado, o produtor de itens deve manter a linguagem usada na produção do texto-base e dos demais elementos do item, usando clareza e precisão, já que, tanto no caso de acerto como no caso de erro por parte do examinado, o comentário terá uma função didática importante. Um erro comum que deve ser evitado pelo autor do item é simplesmente repetir ou parafrasear a opção verdadeira – um comentário com este, que alguns manuais chamam de justificativa tautológica, é praticamente inútil e deixa de exercer sua função didática.

Uma segunda forma de apresentar comentários, típica de plataformas eletrônicas, é proporcionar um pequeno texto ao examinado para cada uma das opções de resposta, conforme se verifica no item X13, que apresenta comentários independentes endereçados ao examinado.

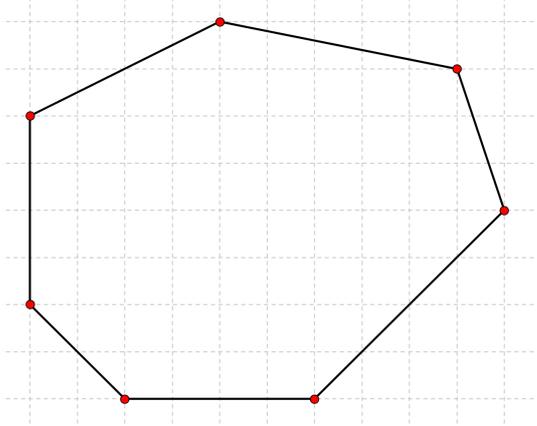
<b>Identificação</b>	X13	<b>Disciplina</b>	Matemática	<b>Gabarito</b>	B
<b>Descritor</b>	Reconhecer e aplicar o Teorema de Pitágoras em situações problema.				
<b>Texto-base</b>					
<p>Um entregador de encomendas saiu com sua motocicleta do depósito onde trabalha, viajou três quilômetros para o sul e fez uma entrega. Em seguida, viajou oito quilômetros para o leste e fez outra entrega. Finalmente, viajou nove quilômetros para o norte e fez sua última entrega. Há uma estrada em linha reta que leva o entregador de volta ao depósito onde trabalha.</p> <p>Utilize a figura abaixo para entender melhor a situação, se for necessário.</p>					
					
<b>Comando</b>					
Quantos quilômetros o entregador deve viajar em linha reta para retornar ao depósito onde trabalha?					
<b>Opções de resposta</b>					
A. 8.	Será que você considerou apenas o deslocamento no sentido leste-oeste? Tente novamente, preferencialmente fazendo um desenho.				
B. 10.	Muito bem! Você deve ter percebido que o problema pode ser resolvido pelo Teorema de Pitágoras, tomando como catetos a medida leste-oeste de oito quilômetros e a medida norte-sul de seis quilômetros (que vem da diferença entre nove e três). Para encontrar a medida da hipotenusa o cálculo é: $\sqrt{6^2 + 8^2} = \sqrt{36 + 64} = \sqrt{100} = 10.$				
C. 12.	Talvez seja melhor reler o problema e fazer um desenho. Tente de novo!				
D. 20.	Será que você simplesmente somou as distâncias? Tente novamente, preferencialmente fazendo um desenho.				

Como se observa no item X13, no caso de escolha da opção verdadeira, em geral o avaliado é parabenizado pelo acerto e o texto do comentário explica como o item é resolvido. Isto se justifica por não haver certeza se o examinado

domina, de fato, a habilidade tratada no descritor e eventualmente ter acertado o gabarito sem muita segurança ou até mesmo por puro acaso. No caso de o examinado errar, indicando uma opção de resposta falsa, o comentário pode apresentar mais informações acerca do descritor e, de acordo com as especificações da plataforma eletrônica, o avaliado pode ter uma ou mais chances de tentar novamente acertar, agora contando com a ajuda oferecida pelo comentário. O comentário para cada uma das opções de resposta falsa pode ser único ou, se cabível, diferente para cada tipo de falha cometida pelo avaliado, ocasião em que o possível erro pode ser apresentado ao examinado, para que ele perceba sua falha.

Esta maneira de apresentar comentários é típica de processos de avaliações formativas ou processuais, nas quais os estudantes são ajudados pelos testes, por meio dos quais se verifica se eles estão atingindo os objetivos intermediários previstos no decorrer de sua trajetória, de modo que sejam possíveis ações, por parte dos professores, antes do fim de uma etapa. Da mesma forma que na hipótese tratada no parágrafo anterior, o produtor de itens deve manter a linguagem clara e precisa em seus textos, considerando a função didática dos comentários dirigidos ao avaliado.

Por fim, uma terceira maneira de se exigir a produção de comentários do autor de itens consiste em requerer textos não mais direcionados ao examinado, mas que têm o revisor dos itens como destinatário. Neste caso, uma possibilidade é o produtor apresentar a solução do item junto ao gabarito e explicações que justificam a escolha dos distratores associadas a cada um deles. Neste caso, a preocupação com a adequação da linguagem pode ser um pouco reduzida, já que o leitor dos comentários domina o assunto e estes textos não serão publicados, servindo apenas para munir o revisor dos raciocínios usados pelo produtor para construir os distratores. O item X14 é um exemplo destes tipos de comentários.

<b>Identificação</b>	X14	<b>Disciplina</b>	Matemática	<b>Gabarito</b>	B
<b>Descritor</b>	Resolver problema utilizando propriedades dos polígonos (soma de seus ângulos internos e externos, número de diagonais, cálculo da medida de cada ângulo interno nos polígonos regulares).				
<b>Texto-base</b>					
<p>Uma empresa tem sete prédios que são ligados por cabos de comunicação de dados de forma semelhante a um anel, conforme mostrado na figura abaixo, na qual os pontos vermelhos simbolizam os prédios e os segmentos pretos representam os cabos. A empresa pretende instalar mais cabos para estabelecer ligações diretas entre todos os prédios.</p>					
					
<b>Comando</b>					
O número de novos cabos a serem instalados é					
<b>Opções de resposta</b>					
A. 7.	Errado. O examinado pode considerar que a resposta é sete simplesmente por existirem sete cidades.				
B. 14.	Certo. O examinado deve contar o número de diagonais do polígono, possivelmente utilizando a fórmula $\frac{n(n-3)}{2}$ .				
C. 21.	Errado. O examinado pode estabelecer um método de contagem que inclua erroneamente os lados do polígono.				
D. 28.	Errado. O examinado pode tentar utilizar um método de contagem que considere cada ligação duplamente, sem perceber que o cabo que liga o prédio X ao prédio Y é o mesmo que liga o prédio Y ao prédio X.				

#### IV. ORIENTAÇÕES PRÁTICAS PARA A PRODUÇÃO DE ITENS

A seguir são apresentadas orientações gerais para a produção de itens, que apresentam uma síntese dos pontos críticos expostos anteriormente.

- Lembre-se que cada item deve tratar de somente um descritor.
- Conheça toda a matriz de referência e não apenas o descritor relacionado ao item em produção.
- Conheça detalhadamente as regras de formatação e as restrições do manual de produção que orienta o trabalho.
- Evite iniciar itens por ideias criativas. Comece pelo descritor e gere a ideia depois.
- Seja probo: cada item deve ser necessariamente inédito. Falhas neste aspecto comprometem a sua reputação individual e podem resultar no seu afastamento.
- Seja preferencialmente criativo, afastando-se das abordagens tipicamente livrescas.
- Privilegie o conhecimento do avaliado em lugar de valorizar sua capacidade de memorização.
- Considere que o tamanho médio dos itens produzidos deve ser compatível com o tempo médio destinado a cada um, tipicamente de três minutos.
- Qualquer que seja a disciplina do item, tenha a adequação à norma culta da língua portuguesa e o acatamento à lógica como indispensáveis.

- Lembre-se que seu objetivo não é derrotar o examinado: não devem existir itens com induções a erros ou com “pegadinhas”.
- Em todas as partes do item, persiga a clareza sem abrir mão do rigor conceitual.
- Não subestime o trabalho destinado à criação de distratores. Eles têm importante função pedagógica e são a parte mais difícil da tarefa.
- Siga as regras do manual de produção, mesmo que não concorde integralmente com elas.
- Use uma lista de verificações quando terminar o item para confirmar que ele está conforme as regras definidas. Se possível, tenha acesso à lista de verificações que o revisor técnico-pedagógico vai utilizar. Se não for possível, faça sua própria lista de verificações.
- Escreva o item e aguarde ao menos um dia. Depois faça uma leitura crítica em busca de eventuais correções ou melhorias, antes de enviá-lo para revisão.
- Observe os prazos e o modo de envio dos itens, atentando para confirmações de recebimento, se disponíveis. O atraso ou falha no recebimento podem comprometer sua reputação individual e o trabalho coletivo.

## V. REVISÃO DE ITENS

A revisão técnica e pedagógica de itens de múltipla escolha, chamada também em alguns manuais de “leitura crítica”, é feita preferencialmente por profissionais que já têm experiência em produção de itens. Uma vez tendo escrito muitos itens, um docente certamente terá um olhar mais atento e detectará as falhas mais comuns, principalmente aquelas típicas de produtores iniciantes. Melhor ainda se os profissionais responsáveis pelas revisões tiverem experiência em validação, ou seja, se já tiverem participado das análises pedagógicas e estatísticas de itens e testes que usualmente ocorrem após a aplicação dos exames. Neste caso, será possível deduzir possíveis falhas que eventualmente condenariam o item e evitar que isso ocorra no item que está em desenvolvimento.

Dependendo do processo definido pela instituição responsável pela formação do banco de itens, pode haver um ou dois procedimentos sucessivos de revisão técnica e pedagógica. É comum também que existam revisões relacionadas à conformidade à língua portuguesa e à questão dos direitos autorais dos elementos de apoio incluídos no texto dos itens.

Geralmente, o manual de produção e revisão de itens da instituição responsável dispõe de listas de verificações detalhadas que devem ser completamente observada pelos respectivos revisores. Eles devem verificar cuidadosamente se todos os aspectos dispostos nestas listas foram atendidos, em um trabalho semelhante ao de um inspetor de qualidade. Em caso de não conformidade, devem observar se podem fazer pequenas alterações no item, se devem retornar o item para que o produtor faça as modificações necessárias ou, ainda, se devem simplesmente rejeitar o item.

É certo que o objetivo principal do revisor em seu trabalho é lançar um novo olhar aos itens produzidos, buscando verificar a adequação técnica e

pedagógica e a conformidade com as regras estabelecidas para a produção. Entretanto, o revisor de itens tem também a função de tutorar os produtores. Afinal, é cediço que o conhecimento não advém exclusivamente do treinamento inicial dos produtores, mas também da prática. Assim, quando um revisor devolve um item ao seu autor, seja com anotações a respeito de pequenas modificações realizadas, seja com pedido de alterações ao produtor, deve perceber também a dimensão pedagógica deste ato, isto é, deve compreender que esta é uma oportunidade de esclarecer ao produtor alguma norma que não estava sendo atendida. Se assim proceder, os participantes da equipe de produção compartilharão seus conhecimentos e a produção tenderá a ser mais rápida e ter melhor qualidade.

Durante a troca de mensagens entre revisor e produtor, que, em alguns casos, é feita de maneira que ambos não sejam identificados, cada um deles deve, em um exercício de empatia, tentar se colocar no lugar do outro. O revisor deve ter o cuidado de registrar as inadequações de forma polida. Reciprocamente, o autor deve receber as críticas respeitosamente e, aceitando-as ou não, responder de maneira educada. A prática mostra que podem surgir conflitos entre produtores e revisores quando estas formalidades são abandonadas, pois algumas pessoas podem interpretar discordâncias como se fossem ofensas pessoais, se entenderem que as críticas foram construídas com o intuito de minimizar a sua capacidade intelectual.