



Essências EDUCARE

OS TESTES DE ESCOLHA MÚLTIPLA (TEM)

■ Hugo Camilo
■ José António P. da Silva

A avaliação da aprendizagem dos alunos é uma das componentes críticas do processo educativo. Se esta for utilizada de uma forma apropriada, pode ser um factor decisivo para se atingirem os Objectivos da unidade ou curso. Caso contrário, pode pôr em causa quaisquer esforços de inovação e melhoria da qualidade dos métodos e técnicas pedagógicas: os testes são uma forte fonte de motivação e os alunos tenderão a estudar apenas aquilo que crêem lhes será perguntado nos exames.

Os testes de escolha múltipla constituem uma das principais formas de avaliação em todo o mundo, já que aliam facilidade de execução e avaliação, economia de tempo e fiabilidade. **Está provado que podem avaliar processamento cognitivo de alto nível de complexidade, como a interpretação, síntese e aplicação do conhecimento, além de testar a simples recordação de factos isolados.** Contudo, para que tal seja possível, torna-se indispensável respeitar um conjunto de regras básicas que aqui revemos.

Recomendamos vivamente a leitura cuidadosa do livro *"Constructing Written Test Questions For the Basic and Clinical Sciences"* (disponível em: www.nbme.org/PDF/ItemWriting_2003/2003IWGwhole.pdf) e consensualmente reconhecido como a "bíblia" dos testes de escolha múltipla.

1. O que são os TEM?

Consideramos, aqui, apenas as Questões de escolha múltipla, com resposta única: consistem num "tronco" (eg. uma apresentação de um caso clínico) e numa pergunta, seguida por um conjunto de opções, tipicamente uma correcta e quatro incorrectas, designando-se estas por distractores.

2. Porque devo utilizar os TEM?

Economia: Os TEM permitem testar um grande número de alunos em simultâneo com grande economia de tempo e facilidade logística;

Objectividade: Cada resposta é previamente considerada correcta ou incorrecta: não há espaço para a impressão e potencial enviesamento subjectivo na apreciação da resposta dos alunos.

Abrangência: A avaliação envolve sempre um processo de amostragem dos tópicos e objectivos de aprendizagem. Não é possível testar tudo! Os TEM permitem a avaliação de um grande número de itens.

Equidade: As perguntas e as condições são exactamente iguais para todos os alunos.

Boas qualidades psicométricas: Revelam alta fidedignidade e validade porque a inclusão de um número alargado de questões e tópicos aumenta a segurança das inferências acerca das competências dos alunos.

Alguns professores expressam a preocupação de que um aluno pode responder a uma questão de escolha múltipla reconhecendo a opção correcta, ao invés de gerar a resposta espontaneamente (efeito de *"Cueing"*). Embora este efeito seja incontornável, a evidência sugere que não influencia significativamente a natureza dos processos de raciocínio invocados pela pergunta.

Feedback rápido: A correcção rápida contribui para a função formativa da avaliação: os alunos podem saber rapidamente quais são os seus pontos fortes e fraquezas.

3. Que regras devo seguir para escrever boas perguntas de escolha múltipla?

Proteja o tempo necessário para produzir boas perguntas: A elaboração de boas perguntas de escolha múltipla consome bastante tempo. Assuma isso, se quiser garantir qualidade.

Coloque-se na pele dos seus alunos: Imaginar-se na posição dos alunos com o seu grau estimado de conhecimento, maturidade e experiência, poderá ajudá-lo a compreender os processos mentais que estará a solicitar-lhes com cada questão.

Teste a aplicação de conhecimentos. Evite itens que apenas requeiram a recordação de factos isolados: O objectivo da aprendizagem consiste em fazer coisas com o que se aprende. Peça-lhes que apliquem!

Exemplos:

Factos isolados: Qual das seguintes hipóteses caracteriza a desordem de ansiedade de separação?

Aplicação de conhecimento: Um rapaz de 8 anos necessita de ser fortemente estimulado para ir à escola e, uma vez lá, queixa-se frequentemente de dores de cabeça severas e dores de estômago. Por vezes a sua mãe tem que o levar para casa devido aos seus sintomas.

À noite, ele tenta dormir junto dos seus pais. Quando estes insistem que durma no seu próprio quarto, o rapaz diz que existem monstros dentro do armário.

Qual das seguintes hipóteses melhor explica o seu comportamento?

- A. Comportamento próprio da idade
- B. Esquizofrenia infantil
- C. Desordem de ansiedade de separação
- D. Fobia social
- E. Desordem psicótica partilhada

Foque os itens em conceitos-chave: Aborde temas e princípios que re-presentem informação essencial para a futura prática profissional do médico/médico dentista. Não avalie factos triviais.

Os tópicos e objectivos mais importantes devem pesar mais e estar mais representados no teste do que tópicos menos importantes, sendo que o tempo de testagem também deve reflectir a importância relativa de cada tópico.

Sempre que possível, baseie a questão num caso clínico ou num problema relevante: Esta regra permite garantir simultaneamente a relevância e a aplicação de conhecimentos!

A grande maioria das questões pode ser redigida com base na apresentação de um caso clínico. O "tronco" deve começar pode incluir parte ou a totalidade dos dados listados abaixo. Procure o equilíbrio adequado entre os dados necessários à resposta e "os distractores"

- Idade, Género (eg. um homem de 45 anos)
- Local de atendimento (eg. desloca-se ao centro de saúde)
- Queixas (eg., por causa de uma dor de cabeça)
- Duração (eg. que se prolonga há dois dias)
- História do doente
- Achados físicos

Resultados dos testes diagnósticos (se relevante)
Tratamento inicial, achados subsequentes, etc. (se relevante)

O tronco deve ser explícito. A pergunta clara. As opções sucintas: Inclua toda a informação necessária no “tronco” da questão, que deve consistir num problema formulado de forma explícita. Formule uma pergunta clara. Ofereça opções curtas e objectivas. **Note bem:** Deve ser possível responder à pergunta formulada sem ter acesso às opções de resposta.

Selecione bons distractores: plausíveis, homogêneos e com uma extensão semelhante à da opção correcta.

Considere o uso de “templates”: Uma forma de facilitar a elaboração dos itens de escolha múltipla consiste em utilizar *templates* que definem a estrutura global dos itens. Pode gerar muitas questões utilizando o mesmo *template*.

Exemplo: **Um (incluir descrição do doente) é incapaz de (Incapacidade funcional). Quais dos seguintes (órgão, tecido) mais provavelmente terá sido lesionado?**

Vários *templates* adicionais podem ser consultados em “Constructing Written Test Questions For the Basic and Clinical Sciences”, disponível em: www.nbme.org/PDF/ItemWriting_2003/2003IWGwhole.pdf

Considere o recurso a “bancos de perguntas”: Existem vários “bancos” internacionais que permitem o acesso a perguntas validadas e testadas em aplicações prévias. Em alguns casos, as perguntas são pagas. (eg.: *National Board of Medical Examiners*, disponível em: www.nbme.org/programs-services/medical-schools/customized-assessment.html).

Bem mais interessante e produtiva é a participação em consórcios cooperativos em que os docentes têm oportunidade de propor questões e as ver apreciadas por outros peritos (eg.: *Universities Medical Assessment Partnership* e *IDEAL Consortium*, disponíveis em: www.umap.org.uk e www.hkwebmed.org/idealweb/homeindex.html, respectivamente).

4. Que erros e falhas técnicas devo evitar na construção dos TEM?

As perguntas só serão eficazes se forem bem estruturadas e isentas de erros que comprometam a justiça na interpretação dos resultados e das consequências que deles são extraídas. Seguem-se alguns exemplos de erros e falhas técnicas mais comuns:

Relacionados com a “esperteza” (“testwiseness”): As perguntas oferecem pistas ou indícios de que o “esperto” ou “hábil” retira proveito, inferindo qual a resposta correcta sem que detenha o conhecimento relevante necessário à resposta.

Exemplos:

Pistas gramaticais (um ou mais distractores não decorrem gramaticalmente do “tronco” do item – por princípio, são falsos); **pistas lógicas** (um sub-conjunto de opções esgotam as alternativas de resposta); **uso de termos absolutos** (termos como “sempre” ou “nunca” são usados nas opções – em princípio, são falsas); **resposta correcta mais longa** (a resposta correcta é mais longa, mais completa ou mais específica que as outras opções); **repetição de palavras** (as mesmas palavras são repetidas no “tronco” e na opção correcta); **estratégia convergente** (a resposta correcta inclui a maioria dos elementos em comum com as restantes opções).

Relacionados com o conteúdo: Evite a todo o custo questões baseadas em simples recordação de dados: elas não apelam ao raciocínio dos seus alunos, não demonstram que eles digerido a informação, não garantem que eles farão uso adequado dos conhecimentos que demonstram, elas estimulam a que no futuro os seus alunos de limitem a memorizar, sem compreender.

Relacionados com fraseologia das questões: A construção complexa, confusa, imprecisa ou inadequada das questões introduz dificuldades artificiais que nada têm a ver com as competências que deveriam estar a ser avaliadas.

Exemplos:

A pergunta é ambígua; O tronco é demasiado complexo ou longo; Opções longas, complicadas ou duplas (pode estar-se a avaliar a velocidade de leitura mais do que o próprio conhecimento.); **Dados numéricos apresentados de forma inconsistente** (formato das opções numéricas não é uniforme); **Termos “quantitativos” são vagos** (eg. raramente, habitualmente – têm diferentes significados para diferentes pessoas); **a linguagem e/ou terminologia usada nas diferentes opções não é coerente; As opções não são apresentadas numa ordem lógica; Pergunta e opções resultam em duplas negações. Evite itens formulados pela negativa** (eg. os que incluem os termos “excepto”, “não”, ou “nem” na pergunta).

Evite itens traiçoeiros ou desnecessariamente complexos.

5. O que fazer com os resultados dos testes de escolha múltipla?

Muitas escolas proporcionam aos seus docentes informação analítica sobre a qualidade dos itens de um teste, após a sua aplicação. Assim, antes ainda publicar as classificações dos alunos, é possível anular itens mal concebidos, identificar crenças erróneas dos alunos e seleccionar itens particularmente bons do ponto de vista estatístico, para eventuais utilizações futuras para identificar os que precisam de correcção.

Entre outros estudos analíticos possíveis, destacam-se os seguintes, pelo seu uso disseminado e impacto potencial na melhoria da qualidade dos testes: **Percentagem de “bons e maus” alunos que selecciona cada opção:** A análise da distribuição das respostas entre os grupos que obtêm os melhores (e 50% melhores) e os piores resultados (eg. 50% piores) permite extrair conclusões importantes quanto: a) à correcção da chave de resposta, b) à existência de distractores que podem ser confundidos com a resposta certa, c) à existência de mais do que uma opção correcta para a mesma pergunta, d) à existência de distractores demasiado fáceis, que devem ser revistos, e) à capacidade discriminativa da pergunta e seu grau de dificuldade.

Índice de Dificuldade. ($p = \sum x/n$) ($\sum x$ é a soma de todos os resultados correctos obtidos no item; n é o nº de respondentes). Dá-nos uma medida da proporção de alunos que responderam correctamente. Este índice varia entre 0 e 1, sendo que a escolha entre itens com maior ou menor grau de dificuldade dependerá das finalidades visadas. Em geral, um índice de dificuldade moderado (próximo de 0,5) é desejável.

Índice de Discriminação. ($D = p+ - p-$) ($p+$ é o índice de dificuldade do grupo mais forte e $p-$ é o do grupo mais fraco) É uma forma – entre outras possíveis – de averiguar se um item discrimina entre os melhores e piores alunos. Este índice varia entre 0 e 1, sendo necessária a revisão dos itens com valores muito próximos de 0, por não discriminarem entre os menos e mais competentes, logo, sendo passíveis de conterem alguma deficiência na sua construção.

Em suma, a aplicação cuidadosa de técnicas apropriadas de construção e validação de TEM e dos seus itens permite uma qualidade acrescida dos testes aplicados aos alunos e conduzirá a uma maior justiça, equidade e consequência construtiva do processo avaliativo.

Para saber mais:

- Brady, A. (2005). *Assessment of learning with multiple-choice questions*. In: *Nurse Education in Practice*, 5: 238-242, 2005.
- Case, S. M., Swanson, D. B. (2001) (3rd Edition). *Constructing Written Test Questions For the Basic and Clinical Sciences*. National Board of Medical Examiners.
- Haladyna, T.M. (1989). *A Taxonomy of Multiple-Choice Item-Writing Rules*. In: *Applied Measurement in Education*, Vol. 2, No 1: 37-50, 1989.
- Haladyna, T.M., Downing, S. M., Rodriguez, M.C. (2002). *A Review of Multiple-Choice Item-Writing Guidelines for Classroom Assessment*. In: *Applied Measurement in Education*, Vol. 15, Nº 3: 309-334, 2002.
- McCoubrie, P. (2004). *Improving the fairness of multiple-choice questions: a literature review*. In: *Medical Teacher*, Vol. 26, No 8: 709-712, 2004.