

CONSTRUÇÃO DE PERGUNTAS E TESTES DE ESCOLHA MÚLTIPLA

GUIÃO ORIENTADOR PARA REGENTES DO MIM E MIMD DA FMUC

Este documento pretende definir orientações e recomendações aprovadas pelo Conselho Pedagógico da FMUC para a construção de perguntas e testes de escolha múltipla a aplicar em avaliações dos estudantes do MIM e MIMD.

CONSTRUÇÃO DE PERGUNTAS E TESTES DE ESCOLHA MÚLTIPLA

GUIÃO ORIENTADOR PARA REGENTES DO MIM E MIMD DA FMUC

ENQUADRAMENTO

Neste Guião, os docentes encontrarão as principais orientações e recomendações de boas práticas na construção de perguntas e testes de escolha múltipla, baseados na melhor evidência disponível.

A organização da informação tem em consideração a finalidade deste documento, que visa proporcionar recomendações com tradução direta na prática pedagógica e de avaliação das aprendizagens dos estudantes, privilegiando-se uma organização da informação de forma clara, lógica, resumida e orientada para as suas aplicações práticas.

O Guião foi elaborado pelo Gabinete de Educação Médica, sob proposta do Conselho Pedagógico da FMUC. Os regentes de unidades curriculares do MIM e MIMD devem orientar a sua prática avaliativa, e aplicar as recomendações aqui descritas, nos testes cuja finalidade seja a avaliação sumativa dos estudantes destes cursos.

O IMPACTO DOS TESTES NA APRENDIZAGEM

Um já célebre chavão educacional afirma que *“a Avaliação comanda a Aprendizagem”*. Esta ideia é particularmente relevante no caso dos testes escritos, considerando que estes representam uma fatia considerável dos instrumentos de avaliação utilizados nesta Faculdade. Sabe-se que os estudantes são motivados e modelam o seu comportamento em função das exigências da avaliação e tenderão a aprender o que creem que é mais valorizado.

Como a influência dos testes é tão forte, é importante que estes sejam concebidos de modo a irem ao encontro dos objetivos educacionais das unidades curriculares e do curso.

Uma crítica persistente aos testes de escolha múltipla é a de que se limitam a avaliar a capacidade de memorização e de recordação de factos isolados pelos estudantes. Décadas de estudos e investigação neste domínio demonstram que este tipo de testes, desde que adequadamente construídos, também podem avaliar capacidades cognitivas de ordem superior, tais como a interpretação, síntese e aplicação do conhecimento.

I. FAMÍLIAS E TIPOS DE TESTES (ÂMBITO DE APLICAÇÃO)

Quando nos referimos a testes e perguntas de escolha múltipla, tendemos a tomar como referência os itens da denominada escolha múltipla convencional (um enunciado contendo uma pergunta e um conjunto de 3 a 5 opções de resposta, em que apenas uma está correta). Todavia, existe uma série de formatos alternativos e variantes ao modelo convencional, nem todos recomendados de igual forma pelos autores. Na tabela seguinte encontram-se os principais formatos deste tipo de itens, adaptada da classificação proposta por Case e Swanson (2003) bem como as principais características e recomendações associadas a cada um deles:

Tabela 1: Famílias e Formatos de Escolha Múltipla (EM)

Família	Formato	Descrição e Exemplo	Recomendações
“One-Best-Answer” (Resposta Única)	Tipo A (convencional, 4 ou mais opções, itens simples ou sets)	Ao contrário da família de verdadeiro/falso, este tipo de questões torna claro o número de opções a selecionar.	Esta família de testes é geralmente recomendada pela maioria dos autores, sendo especialmente recomendado que os itens sejam <i>ricos em contexto</i> . Isto é, quando os itens nestes formatos são enquadrados em situações e problemas clínicos a resolver, tendem a suscitar comportamentos cognitivos de maior complexidade, tais como a aplicação de conhecimento, a sua síntese e integração.
	Tipo B (4 ou 5 opções de correspondência em sets de 2-5 itens) Tipo R (Extended-Matching em sets de 2-20 itens)	São compostas por um enunciado, uma questão (Lead-In), seguida por um conjunto de opções de resposta. Exemplo: <i>Um homem de 65 anos tem dificuldade em erguer-se partindo de uma posição sentada e em endireitar o tronco, mas não revela dificuldades em fletir a sua perna. Qual dos seguintes músculos está mais provavelmente lesionado?</i> *A. Glúteo máximo B. Glúteo mínimo C. Femurais D. Iliopsoas E. Obturador interno	
Verdadeiro/Falso	Tipo C (A/B/Ambos/nenhum)	Esta família de testes requer que um aluno selecione todas as opções que são “verdadeiras”. Exemplo: <i>Afirmações verdadeiras sobre a fibrose quística incluem:</i> 1. A incidência de FQ é de 1:2000. 2. As crianças com FQ geralmente morrem na adolescência. 3. Homens com FQ são estéreis. 4. A FQ é uma doença autossómica recessiva.	<u>O National Board of Medical Examiners dos EUA deixou de utilizar esta família de testes nos seus exames de acesso à profissão.</u> A recomendação desta entidade é a de que não se utilize questões de verdadeiro/falso, porquanto tendem a avaliar apenas a recordação de factos isolados, é mais difícil estabelecer a distinção entre o que é absolutamente verdadeiro do que é falso e são muito frequentemente descartadas pelos revisores. São substituídas com vantagens pelas perguntas de “One-Best-Answer”
	Tipo K (verdadeiro-falso complexos)		
	X (itens de verdadeiro-falso simples) Simulações (tais como os Patient Management Problems (PMP))		

II. PRINCÍPIOS GERAIS NA CONSTRUÇÃO DE TESTES DE ESCOLHA MÚLTIPLA (TEM)

A qualidade de um teste de avaliação está deveras relacionada com os conceitos de validade, fidedignidade e justiça. Sem entrar em detalhes sobre cada um destes conceitos de teor mais técnico, importa sintetizar alguns dos corolários lógicos da adoção destes pressupostos de qualidade quando adaptados aos TEM, segundo a revisão da literatura efetuada por McCoubrie (2004):

Um TEM “defensável” ou justo deve cumprir com os seguintes requisitos:

- I. Ser alinhado e coerente com os objetivos de aprendizagem da unidade curricular e do curso;
- II. Ser combinado com a testagem das competências práticas;
- III. Traduzir uma amostra representativa do conteúdo importante;
- IV. Estar isento de itens contendo erros e falhas técnicas;
- V. Utilizar um critério de aprovação/retenção baseado em critérios (absoluto) e não em normas (relativo). Isto é, a aprovação ou retenção dos alunos deve basear-se com um critério externo/referencial de conhecimento ou competência e não na base da comparação com os restantes colegas);
- VI. Ser cuidadoso na reutilização de perguntas de teste;
- VII. Utilizar formatos alternativos que já provaram a sua validade e fidedignidade (p.ex., Extended-Matching)

III. REGRAS E ETAPAS NA CONSTRUÇÃO DAS PERGUNTAS

Escolha Múltipla Convencional

Para a construção de perguntas ou itens de escolha múltipla convencional, procure percorrer os seguintes passos:

1. Construa uma matriz para o **enunciado da pergunta**.

A estrutura geral de um item pode ser representada por uma matriz genérica, que facilita a construção de perguntas em série.

Exemplos:

Um (*descrição do paciente*) apresenta (*tipo de lesão e localização*). Qual das seguintes estruturas está mais provavelmente afetada?

Um (*descrição do paciente*) tem (*achados da história*) e está a tomar (*medicações*). Qual das seguintes medicações é a causa mais provável da sua (*um achado da história, exame físico ou laboratorial*)?

Um (*descrição do paciente*) tem (*sintomas, sinais ou doença específica*) e está a ser tratado com (*fármaco ou classe de fármacos*). O fármaco atua ao inibir qual dos seguintes (*funções, processos*)?

Na sequência de (*um procedimento*), um (*descrição do paciente*) desenvolve (*sintomas e sinais*). Os achados laboratoriais mostram (*achados*). Qual das seguintes é a causa mais provável?

Enquadre, sempre que possível, o enunciado no contexto de uma **Vinheta Clínica**, que pode incluir toda ou alguma da informação seguinte:

Mesmo que não recorra a uma *vinheta clínica*, procure que o enunciado da

Idade/Género (p.ex.: um homem de 45 anos)
Local (p.ex.: dá entrada nas urgências)
Queixa principal (p.ex.: queixando-se de uma cefaleia)
Duração (p.ex.: que persiste há 2 dias)
História do paciente (com história familiar?)
Achados Físicos
+/- Resultados de exames diagnósticos
+/- Tratamento inicial, achados subsequentes, etc.

pergunta seja *rico em contexto*. Isto é, que esteja enquadrado num cenário-problema relevante e que promova a aplicação do conhecimento e/ou a resolução de problemas.

Deve incluir todos os factos relevantes; não deve ser dada informação adicional nas opções;

Não deve conter "ratoeiras", nem ser demasiado complexo.

2. **Redija o Lead-in** (pergunta ou afirmação)

- O *lead-in* indica a relação entre os enunciados (casos clínicos) e as opções, clarificando as questões colocadas aos examinandos. É um elemento essencial;
- Verifique que a pergunta pode ser respondida sem olhar para as opções de resposta;
- Não deve ser redigido na negativa (p.ex., evitar usar "exceto" ou "não" no *lead-in*). Se o fizer, assinalo devidamente com letra maiúscula e/ou sublinhado.
- Podem ser redigidos mais *lead-ins* e opções se se pretender construir um *Set* com mais do que um item para o mesmo caso clínico.

3. **Prepare uma lista de Opções**

- A lista de opções deve conter palavras isoladas ou frases muito curtas. Não devem ser usados verbos, pois estes devem estar no *Lead-in*;
- Liste as opções por ordem alfabética, ou lógica, se esta existir;
- As opções devem ser homogéneas (só diagnósticos, só investigação, só causas de doença, ou só mecanismos fisiopatológicos, etc.);
- As opções devem ser relevantes e plausíveis.

Exemplos de lead-ins e listas de opções:

Qual dos seguintes é (anormal)?

As listas de opções podem incluir sítios de lesões; lista de nervos; lista de músculos; lista de enzimas; lista de hormonas; tipos de células; lista de neurotransmissores; lista de toxinas, moléculas, vasos, etc.

Qual dos seguintes achados é mais provável?

As listas de Opções podem incluir lista de resultados laboratoriais; lista de sinais físicos adicionais; resultados de autópsia; resultados de análise de ADN, etc.

Qual dos seguintes é a causa mais provável?

A lista de Opções pode incluir lista de mecanismos subjacentes à doença; medicações que podem causar efeitos secundários; fármacos ou classes de fármacos; agentes tóxicos, defeitos metabólicos, vírus, etc.

Qual dos seguintes deve ser administrado?

As Opções podem incluir fármacos, vitaminas, aminoácidos, enzimas, hormonas, etc.

Extended-Matching

As regras gerais para a construção dos itens de escolha múltipla convencional aplicam-se às perguntas no formato de *Extended-Matching*. As principais diferenças encontram-se na sequência das etapas na construção dos itens.

1. Identifique um **Tema**

Ao contrário da escolha múltipla tradicional, aqui, o tema, as opções e o Lead-In escrevem-se antes do(s) caso(s) clínico(s);

O tema pode ser uma queixa principal (eg. Dor no peito, fadiga), uma situação circunstancial (eg. Admissão ou alta médica num departamento de urgências), ou uma classe de fármacos (eg. Agentes hipertensivos, antibióticos).

2. Escreva o **Lead-in** (pergunta ou afirmação)

O lead-In indica a relação entre os enunciados (casos clínicos) e as opções, clarificando as questões colocadas aos examinandos. É um elemento essencial;

Exemplos:

Para cada um dos seguintes pacientes, selecione (p.ex., nervo) que está mais provavelmente (afetado, lesionado, deficiente, não-funcional).

Os conjuntos de opções poderiam ser listas de nervos, listas de músculos, listas de enzimas, listas de hormonas, etc.

Para cada um dos seguintes pacientes, selecione a (p.e.x., causa) mais provável.

Os conjuntos de opções poderiam ser listas de mecanismos de doença subjacentes, medicamentos com efeitos secundários, agentes tóxicos, mecanismos hemodinâmicos, etc.

3. Prepare uma **lista de Opções**

Exemplos de Tópicos para listas de Opções

artérias	enzimas	estruturas endócrinas	fármacos/classes de fármacos
nervos	tipos de células	distúrbios metabólicos	anomalias congénitas
músculos	vírus	processos patológicos	estruturas cardíacas
hormonas	toxinas	testes diagnósticos	diagnósticos

4. Escreva pelo menos dois **enunciados de casos clínicos**

Exemplo:

1. Um homem de 72 anos, dextro, tem fraqueza e hiperreflexividade do membro direito inferior, uma resposta plantar extensora à direita, força normal no braço direito e movimentos faciais normais. Resp.: A

2. um homem de 62 anos, dextro, apresenta uma hemiparesia espástica à direita, uma resposta plantar extensora à direita. O seu discurso é fluente, e tem uma compreensão normal das instruções verbais e escritas. Resp.: G

IV. ORIENTAÇÕES PARA A CONSTRUÇÃO DAS PERGUNTAS

As orientações para a elaboração de perguntas aqui apresentadas são adaptadas das propostas de vários autores (Haladyna et al (2002), Dubins et al (2016), Case & Swanson (2003), etc.), e baseiam-se quer em revisões sistemáticas de estudos empíricos, quer em recomendações de boas práticas, com um grau considerável de consenso, apresentada nos manuais e literatura especializada.

Formatação e Conteúdo	Enunciado	Opções
Cada item deve incidir sobre conteúdos e objetivos de aprendizagem importantes. Evite conteúdo trivial	As instruções no enunciado devem ser muito claras.	Elabore o maior número de opções plausíveis que puder. A investigação sugere que 3 opções plausíveis são suficientes.
Os itens devem ser construídos para um nível de dificuldade adequado. Evite introduzir dificuldade irrelevante	A ideia principal deve estar contida no enunciado e não nas opções	Certifique-se que apenas uma das opções é a resposta <i>mais correta</i> .
Os itens devem testar a aplicação do conhecimento ao invés da mera recordação de informação e factos isolados	Evite a excessiva verbosidade (enunciados excessivamente longos ou com detalhes irrelevantes)	Varie a localização da opção correta
Use material novo para testar níveis de complexidade cognitiva mais elevada (e não apenas a recordação)	Redija o enunciado na afirmativa. Evite negativas tais como NÃO ou EXCETO. Se tiver que usar, garanta que estão sublinhadas ou a negrito	Ordene as opções de forma lógica ou numérica
Evite conteúdo demasiado específico ou demasiado genérico	Deve ser possível responder à pergunta sem olhar para as opções	As opções devem ser independentes, sem sobreposições

Formate os itens verticalmente ao invés de horizontalmente		As opções devem ser homogêneas em conteúdo e unidimensionais (p. ex., todas diagnósticos, todas exames laboratoriais ou tratamentos)
Minimize a quantidade de leitura em cada item.		As opções devem ser curtas e com extensão semelhante.
Edite e teste os itens.		Evite ou use cuidadosamente “Nenhum dos anteriores” ou “Todos os anteriores”
		Escreva as opções na afirmativa; evite negativas tais como NÃO
		Certifique-se que as opções distratoras são todas plausíveis
		Aproveite os erros típicos dos alunos para escrever as suas opções distratoras

V. ERROS E FALHAS TÉCNICAS A EVITAR

Uma percentagem variável de alunos é capaz de tirar vantagem de erros e falhas técnicas na construção dos itens, relacionados com a sua maior ou menor *testwiseness*, bem como com algumas pistas que o docente inadvertidamente proporciona sobre a resposta correta ao formular as questões de forma menos cuidada. Em sentido inverso, muitos alunos podem ser prejudicados quando se deparam com itens mal construídos, contendo falhas que induzem a uma resposta errada. As tabelas seguintes sintetizam os modos de evitar os principais erros identificados, em função dos benefícios injustamente concedidos (Tabela 3) ou dos prejuízos causados (Tabela 4).

Tabela 3: Evitar erros e falhas técnicas que beneficiam (injustamente) os Alunos

Enunciado	Opções
Certifique-se que as questões são independentes – evite dar a resposta a um item do teste no enunciado de outro item.	Evite o uso de termo absolutos, tais como “sempre”, ou “nunca”, que podem dar pistas sobre a resposta correta
Evite que o conhecimento da resposta a um enunciado seja necessária à resposta a um enunciado subsequente	As opções distratoras devem ser alternativas plausíveis para um aluno, independentemente do seu nível de competência, e seguir uma lógica de raciocínio. Nenhuma deve ser obviamente errada.
Evite pistas gramaticais ou sintáticas (p.e.x. uma palavra ou frase é incluída no enunciado e na resposta correta)	Mantenha as opções curtas e com uma extensão semelhante. As opções mais longas e elaboradas são mais frequentemente a resposta correta.
	Evite opções opostas ou mutuamente exclusivas, porque podem sinalizar que uma delas está correta.
	Não use linguagem utilizada nos manuais de referência para a resposta correta, e não para os distratores
	Evite elementos repetidos nas opções, porque assim a pergunta pode prestar-se a uma estratégia de convergência (a resposta correta geralmente inclui a maioria dos elementos em comum com as outras opções)
	A posição da resposta correta deve variar (p.ex., não colocar todas as respostas corretas na alínea c))
	Evite ou use cuidadosamente o formato de Verdadeiro/Falso, considerando que são geralmente

menos discriminatórias, e possibilitam 50% de probabilidade de acerto ao acaso.
Procure reduzir o efeito de *cueing* (pistas inadvertidas) ao utilizar o formato de Extended-Matching (até 20 opções de resposta)

Tabela 4: Evitar erros e falhas técnicas que prejudicam os alunos

Enunciado	Opções
O enunciado deve ser claro, conciso e conter apenas informação relevante	A resposta correta deve ser melhor que as restantes opções e claramente correta (<i>não significa que as opções distratoras tenham que ser incorretas, apenas menos corretas que a resposta-chave</i>)
Evite enunciados contendo “ratoeiras” ou desnecessariamente complicados	As opções devem ser tão breves (curtas) quanto possível. As palavras repetidas nas opções devem ser movidas para o enunciado
Evite as negativas (p.e.x., “exceto”) e duplas negativas. Se necessário, enfatizar com maiúsculas e negrito e usar opções apenas com uma palavra	As opções não devem ser complicadas ou duplas
No <i>lead-in</i> , deve instruir os alunos para selecionarem a melhor opção (que pode ser melhor do que as restantes, mas não tem que necessariamente ser a única correta)	Evite termos vagos ou imprecisos, tais como “frequentemente”, “geralmente”, “raramente”, “algumas vezes”, etc.
Evite a verbosidade (excesso de palavras), minudências irrelevantes ou conteúdo destinado a induzir em erro o aluno	Evite usar “ <i>nenhum dos anteriores</i> ” como opção Evite sobreposições entre as opções. Estas devem ser independentes
Escreva o enunciado como uma pergunta (um <i>lead-in</i> claro) ao invés de uma afirmação incompleta	Quando são usadas opções numéricas, devem ser listadas por ordem numérica (ou cronológica), num único formato A linguagem nas opções deve ser coerente e estas devem estar colocadas por uma ordem lógica

VI. ORIENTAÇÕES PARA A PLANIFICAÇÃO E APLICAÇÃO DOS TESTES

Para além das orientações anteriores sobre a construção dos itens de teste considerados individualmente, há também alguns aspetos que importa considerar, relativos aos procedimentos de planificação do teste, à definição de papéis dos diversos intervenientes, e aos processos da sua aplicação junto do alunos.

Em suma, procuraremos agrupar nas tabelas seguintes as recomendações em função da etapa processual em que ocorrem.

I. Planificação/preparação prévia à aplicação

1. Envolve, sempre que possível, todos os interlocutores necessários, o que significa um esforço concertado de docentes das unidades curriculares, coordenadores curriculares a diferentes níveis e outros peritos nos domínios científicos específicos e de educação médica no desenvolvimento do *blueprint* do teste e nas decisões sobre os itens individuais.
2. Desenvolva um *blueprint* ou tabela de especificações do teste, que garanta que os conceitos que são considerados essenciais e importantes são tratados de forma justa e equilibrada numa determinada prova (Vd exemplo de *blueprint*, adiante)
3. Preveja questões que testem os múltiplos níveis da taxonomia de conhecimento (p.ex., taxonomia de *Bloom*, que inclui os níveis de recordação, compreensão, aplicação, análise e síntese/avaliação)
4. Planeie a criação de um *banco de questões* que permita caracterizar e classificar as questões em função dos seus atributos psicométricos (dificuldade, discriminação) e dos objetivos e conteúdos pedagógicos que visam atingir
5. Preveja formação específica para os docentes envolvidos na construção de testes e perguntas de escolha múltipla
6. Defina um processo de *standard setting* para o teste. Isto é, como será tomada a decisão de aprovação/reprovação dos alunos

reduzir injustiças numa determinada prova e tomar decisões informadas sobre a reutilização das perguntas em testes futuros.

Após a aplicação de um TEM, este é objeto de leitura ótica em equipamento dedicado, a informação é carregada numa base de dados e, de imediato, convertida num relatório de resultados do exame que contém informação de diferentes tipos: classificações atribuídas aos alunos que realizaram o teste (em diferentes escalas), estatísticas descritivas sobre o teste considerado globalmente, bem como informação estatística sobre cada um dos itens do teste considerado isoladamente.

Classificações atribuídas aos alunos

Para além das classificações brutas, que são a simples soma das pontuações atribuídas a cada resposta correta, é ainda fornecida a classificação final na escala de 0-20 e percentual. São também calculados os denominados Z-scores, que mostram quão acima ou abaixo da média se encontra um determinado aluno em unidades de Desvio Padrão. Os Z-scores permitem que os docentes comparem os resultados dos alunos em testes diferentes com diferentes notas totais.

Exemplo:

N.º aluno	Nome	Classificação	Classificação final	Classificação por %	Z-Scores
		32,0	13,61702128	68%	-0,35
		24,0	10,21276596	51%	-1,50
		33,0	14,04255319	70%	-0,21
		31,0	13,19148936	66%	-0,50
		22,0	9,361702128	47%	-1,79
		42,0	17,87234043	89%	1,09

Informação estatística sobre os itens do teste

Dentre os indicadores disponíveis que permitem avaliar as perguntas do teste, encontram-se:

a) *Análise da distribuição de respostas pelas opções*. O objetivo aqui é perceber de que modo as respostas dos alunos se distribuíram pela opções disponíveis, o que permite detetar opções distratoras disfuncionais (que não são selecionadas pelos alunos e deverão ser revistas ou excluídas) ou eventuais erros na atribuição da resposta-chave (correta).

b) *Índice de Facilidade (IF)*. Trata-se da percentagem do número total de alunos que responderam à pergunta corretamente. Varia entre 0 (ninguém responde corretamente) e 1 (todos respondem corretamente). Este índice permite compor testes futuros com um grau de dificuldade desejado (Vd capítulo anterior).

c) *Índice de Discriminação (ID)*. Permite aferir quão bem uma pergunta consegue discriminar entre os “melhores” e “piores” alunos, naquele teste. Varia entre -1 e 1. Se uma questão tem um *ID* elevado é provavelmente bastante discriminativa e consegue captar melhor as diferenças de conhecimento entre estes dois grupos de alunos. Quando os valores deste índice são negativos isso significa que os alunos “melhores” responderam a pergunta incorretamente e

os “piores” corretamente, o que deverá conduzir a que esta seja revista ou descartada do teste.

No exemplo seguinte, é possível verificar-se que, no item P1, apenas 3% dos alunos respondem corretamente (IF=0,03), sendo que a pergunta em causa não contribui em nada para discriminar os alunos neste teste (ID=0,00). Podemos ainda confirmar que 75,8% dos alunos escolhem a opção “E” como resposta. Isto leva a suspeitar que esta possa ser uma pergunta cuja resposta-chave está incorretamente atribuída.

Exemplo:

	Distribuição de frequências das opções de resposta																Dicotomização			
	A		B		C		D		E		F		G		H		Índice de Facilidade	Grupo Superior (27%)	Grupo Inferior (27%)	Índice de Discriminação
	N	%	N	%	N	%	N	%	N	%	N	%	N	%	N	%				
P1	1	3,0%	1	3,0%	1	3,0%	5	15,2%	25	75,8%	0	0	0	0	0	0	0,03	0,00	0,00	0,00
P2	0		10	30,3%	0		23	69,7%	0		0		0		0		0,30	0,56	0,11	0,44
P3	0		1	3,0%	2	6,1%	0		30	90,9%	0		0		0		0,91	1,00	0,78	0,22

Estatísticas sobre o teste

Para além da informação sobre os itens individuais, há todo um conjunto de estatísticas produzidas sobre o teste, considerado globalmente, e que permitem ter uma noção aproximada do grau de “normalidade” do teste, ou da sua maior ou menor dificuldade. São produzidas *medidas de tendência central* (média, mediana, moda, e medidas de dispersão (desvio padrão, mínimo, máximo).

É ainda útil a análise da informação sobre o *Erro Padrão da Medida*, que é uma medida de quanto os resultados de um teste estão dispersos em torno de um resultado “verdadeiro”. Permite derivar *intervalos de confiança* para os resultados obtidos. Esta estatística está diretamente relacionada com a *fidedignidade* ou *consistência* do teste: quanto maior o *Erro Padrão da Medida*, menor a *consistência*.

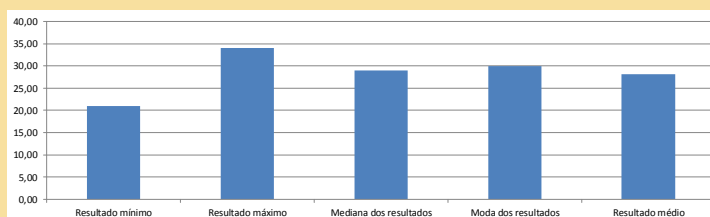
Por último, importa considerar um dos principais indicadores de qualidade de um teste, que é a sua *fidedignidade* ou *consistência*. Esta indica-nos a estabilidade e precisão do teste e é uma medida da capacidade do teste gerar resultados semelhantes quando aplicado em diferentes ocasiões.

Existem diferentes métodos e fórmulas para calcular a consistência de um teste. Neste caso, é usado o *alpha de Cronbach*,, que estima a consistência interna ou o grau de homogeneidade dos itens do teste. Varia entre 0 e 1 e deve situar-se, idealmente, acima de 0,70, mas não muito mais que 0,90.

Exemplo:

Informação estatística sobre o Exame (Resultados obtidos após o mapeamento de todas as versões, originando uma única versão)

Número de Testes	33
Número de Perguntas	38
Resultado mínimo	21,00
Resultado máximo	34,00
Mediana dos resultados	29,00
Moda dos resultados	30,00
Resultado médio	28,15
Desvio padrão dos resultados	3,38
Erro Padrão da medida	2,03
Consistência interna do teste	0,64



VIII. BIBLIOGRAFIA RECOMENDADA

Tavakol, M., Dennick, R., 2011. Post-examination analysis of objective tests. *Medical teacher* 33, 447-458.

Case, S.M., Swanson, D.B., 2003. *Constructing Written test Questions for the Basic and Clinical Sciences* (3rd ed). National board of Medical Examiners, Philadelphia, PA.

McCoubrie, P., 2004. Improving the fairness of multiple-choice questions: a literature review. *Medical Teacher*, vol. 26, nº8, 709-712.

Haladyna, T., Downing, S., Rodriguez, M.C., 2002. A review of multiple-Choice Item-Writing Guidelines for Classroom Assessment. *Applied Measurement in Education*, 15 (3), 309-334.

Dubins, D. N., Gregory, M., Lalitha,, R., 2016. When passing fails: Designing multiple choice assessments to control for false positives. *Currents in Pharmacy Teaching and Learning*, 8, 598-608.

VIII. ANEXOS: FORMULÁRIOS PARA A CONSTRUÇÃO DE PERGUNTAS DE ESCOLHA MÚLTIPLA:

A.1. Formulário para construção de itens de escolha múltipla (convencional);

A.2. Formulário para construção de itens do tipo *Extended – Matching*;

A.3. *Checklist* para revisão de perguntas e testes de escolha múltipla

A.1. Formulário para construção de itens de escolha múltipla (tipo A)

1. Enunciado do Caso/Problema

--

2. Lead-In (afirmação ou pergunta)

#1

#2 (preencher este campo se pretender construir um Set com mais do que um item por caso clínico)

3. Lista de Opções

A.
B.
C.
D.
E.
F.
G.

A.2. Formulário de construção de questões Extended-Matching (tipo R)

1. Identificação do Tema

2. Lead-In (afirmação ou pergunta)

3. Lista de Opções

A	N
B	O
C	P
D	Q
E	R
F	S
G	T
H	U
I	V
J	W
K	X
L	Y
M	Z

4. Enunciado dos casos clínicos

Caso nº 1
Caso nº 2
Caso nº 3

A.3. Checklist para revisão de perguntas e testes de escolha múltipla

	Recomendação	Marque com um V se cumprir a recomendação
Enunciado	O enunciado deve ser claro, conciso e conter apenas informação relevante	
	A ideia principal deve estar contida no enunciado e não nas opções	
	Enquadre, sempre que possível, o enunciado no contexto de uma Vinheta Clínica. Mesmo que não recorra a uma vinheta clínica, procure que o enunciado da pergunta seja rico em contexto	
	Evite a excessiva verbosidade (enunciados excessivamente longos ou com detalhes irrelevantes)	
	Redija o enunciado na afirmativa. Evite negativas tais como NAO ou EXCETO. Se tiver que as usar, garanta que estão sublinhadas ou a negrito	
	Deve ser possível responder à pergunta sem olhar para as opções	
	Certifique-se que as questões são independentes – evite dar a resposta a um item do teste no enunciado de outro item.	
	Evite que o conhecimento da resposta a um enunciado seja necessária à resposta a um enunciado subsequente	
	Evite pistas gramaticais ou sintáticas (p.e.x. uma palavra ou frase é incluída no enunciado e na resposta correta)	
	Evite enunciados contendo “ratoeiras” ou desnecessariamente complicados	
	No <i>lead-in</i> , deve instruir os alunos para selecionarem a melhor opção (que pode ser melhor do que as restantes, mas não tem que necessariamente ser a única correta)	
	Evite a verbosidade (excesso de palavras), minudências irrelevantes ou conteúdo destinado a induzir em erro o aluno	
Escreva o enunciado como uma pergunta (um <i>lead-in</i> claro) ao invés de uma afirmação incompleta		
Opções	Certifique-se que as opções distratoras são todas plausíveis e elabore o maior número de opções que puder. A investigação sugere que 3 opções plausíveis podem ser suficientes.	
	Certifique-se que apenas uma das opções é a resposta <i>mais correta</i> .	
	A linguagem nas opções deve ser coerente e estas devem estar colocadas por uma ordem lógica	
	As opções devem ser independentes, sem sobreposições	
	As opções devem ser homogéneas em conteúdo e unidimensionais (p. ex., todas diagnósticos, todas exames laboratoriais ou tratamentos)	
	As opções devem ser tão breves (curtas) quanto possível.	
	Evite ou use cuidadosamente “ <i>Nenhum dos anteriores</i> ” ou “ <i>Todos os anteriores</i> ”	
	Escreva as opções na afirmativa; evite negativas tais como NAO	
	Aproveite os erros típicos dos alunos para escrever as suas opções distratoras	
	Evite o uso de termo absolutos, tais como “sempre”, ou “nunca”, que podem dar pistas sobre a resposta correta	
	Evite opções opostas ou mutuamente exclusivas, porque podem sinalizar que uma delas está correta.	
	Não use linguagem utilizada nos manuais de referência para a resposta correta, e não para os distratores	
	Evite elementos repetidos nas opções, porque assim a pergunta pode prestar-se a uma estratégia de convergência (a resposta correta geralmente inclui a maioria dos elementos em comum com as outras opções)	
	A posição da resposta correta deve variar (p.ex., não colocar todas as respostas corretas na alínea c))	
	A resposta correta deve ser melhor que as restantes opções e claramente correta (<i>não significa que as opções distratoras tenham que ser incorretas, apenas menos corretas que a resposta-chave</i>)	
As opções devem ter extensão semelhante. As palavras repetidas nas opções devem ser movidas para o enunciado		
As opções não devem ser complicadas ou duplas		
Evite termos vagos ou imprecisos, tais como “frequentemente”, “geralmente”, “raramente”, “algumas vezes”, etc.		
Quando são usadas opções numéricas, devem ser listadas por ordem numérica (ou cronológica), num único formato		
Formatação e Conteúdo	Cada item deve incidir sobre conteúdos e objetivos de aprendizagem importantes. Evite conteúdo trivial, demasiado específico ou demasiado genérico	
	Os itens devem ser construídos para um nível de dificuldade adequado. Evite introduzir dificuldade irrelevante	
	Os itens devem testar a aplicação do conhecimento ao invés da mera recordação de informação e factos isolados	
	Use material novo para testar níveis de complexidade cognitiva mais elevada (e não apenas a recordação)	
	Formate os itens verticalmente ao invés de horizontalmente	
	Minimize a quantidade de leitura em cada item.	
Edite e teste os itens.		