



UNIVERSIDADE DE  
**COIMBRA** FACULDADE  
DE  
MEDICINA

MESTRADO INTEGRADO EM MEDICINA – TRABALHO FINAL

ANA MARGARIDA DA SILVA RIBEIRO

***Prep-PNA/SmashMedicine: A Escrita de Perguntas de Escolha  
Múltipla pelos Estudantes como Método de Estudo***

PROJETO DE INVESTIGAÇÃO

ÁREA CIENTÍFICA DE OUTRAS ÁREAS DA MEDICINA/EDUCAÇÃO MÉDICA

Trabalho realizado sob a orientação de:

PROFESSOR DOUTOR LINO MANUEL MARTINS GONÇALVES

DR. HUGO CAMILO FREITAS DA CONCEIÇÃO

MAIO/2020

# **Prep-PNA/SmashMedicine: A Escrita de Perguntas de Escolha Múltipla pelos Estudantes como Método de Estudo**

Projeto de Investigação

Trabalho final de 6º ano do Mestrado Integrado em Medicina

Ana Margarida da Silva Ribeiro

Aluna do 6º ano do Mestrado Integrado em Medicina

Faculdade de Medicina, Universidade de Coimbra, 3000-075, Coimbra, Portugal

anamargaridasilvaribeiro@gmail.com

Dr. Hugo Camilo Freitas da Conceição

Subdiretor do Gabinete de Educação Médica

Faculdade de Medicina, Universidade de Coimbra, 3000-548, Coimbra, Portugal

Prof. Doutor Carlos Robalo Cordeiro

Diretor da Faculdade de Medicina da Universidade de Coimbra, 3000-548, Coimbra, Portugal

Diretor do Serviço de Pneumologia, Centro Hospitalar e Universitário de Coimbra, 3000-075,

Coimbra, Portugal

Prof. Doutor Henrique Miguel Marques Bom Borges Alexandrino

Faculdade de Medicina, Universidade de Coimbra, 3000-548, Coimbra, Portugal

Departamento de Cirurgia, Centro Hospitalar e Universitário de Coimbra, 3000-075, Coimbra,  
Portugal

Coimbra Institute for Clinical and Biomedical Research (iCBR) área de Environment Genetics and  
Oncobiology (CIMAGO), Faculdade de Medicina, Universidade de Coimbra, 3000-548, Coimbra,  
Portugal

Prof<sup>ª</sup>. Doutora Maria Del Carmen Bento Ribeiro

Faculdade de Medicina, Universidade de Coimbra, 3000-548, Coimbra, Portugal

Centro Hospitalar e Universitário de Coimbra – Hospital Pediátrico, 3000-602, Coimbra, Portugal

Prof<sup>ª</sup>. Doutora Lèlita da Conceição dos Santos

Faculdade de Medicina da Universidade de Coimbra, 3000-548, Coimbra, Portugal

Serviço de Medicina Interna, Centro Hospitalar e Universitário de Coimbra, 3000-075, Coimbra,  
Portugal

Prof. Doutor Rufino Silva

Faculdade de Medicina, Universidade de Coimbra, 3000-548, Coimbra, Portugal

Diretor de Serviço de Oftalmologia, Centro de Responsabilidade Integrado em Oftalmologia, Centro Hospitalar e Universitário de Coimbra, 3000-075, Coimbra, Portugal

Prof<sup>a</sup>. Doutora Margarida Figueiredo Dias

Clínica Universitária de Ginecologia, Faculdade de Medicina, Universidade de Coimbra, 3000-548, Coimbra, Portugal

Serviço de Ginecologia, Centro Hospitalar e Universitário de Coimbra, 3000-075, Coimbra, Portugal

Dr<sup>a</sup>. Joana Oliveira

Clínica Universitária de Ginecologia, Faculdade de Medicina, Universidade de Coimbra, 3000-548, Coimbra, Portugal

Serviço de Ginecologia, Centro Hospitalar e Universitário de Coimbra, 3000-075, Coimbra, Portugal

Prof. Doutor António Ferreira de Macedo

Professor Catedrático de Psiquiatria, Faculdade de Medicina, Universidade de Coimbra, 3000-548, Coimbra, Portugal

Inv<sup>a</sup>. Doutora Ana Telma Pereira

Investigadora Auxiliar

Instituto de Psicologia Médica, Faculdade de Medicina da Universidade de Coimbra, 3000-548, Coimbra, Portugal

Prof. Doutor Lino Manuel Martins Gonçalves

Professor Catedrático de Cardiologia, Faculdade de Medicina, Universidade de Coimbra, 3000-548, Coimbra, Portugal

Diretor de Serviço de Cardiologia, Centro Hospitalar e Universitário de Coimbra, 3000-075, Coimbra, Portugal

Azinhaga de Santa Comba, 3000-354, Coimbra

lgoncalv@ci.uc.pt

# ÍNDICE

RESUMO .....	5
ABSTRACT.....	6
LISTA DE ABREVIATURAS .....	7
INTRODUÇÃO .....	8
Importância da avaliação.....	8
Métodos e Instrumentos de Avaliação do Conhecimento.....	10
Propriedades psicométricas dos testes .....	12
Tipologias de Perguntas de Escolha Múltipla e <i>guidelines</i> para a sua construção .....	14
Enquadramento legislativo e evolução histórica da Prova Nacional de Acesso .....	17
Adaptação à Prova Nacional de Acesso.....	19
Construção de Perguntas de Escolha Múltipla pelos estudantes.....	20
Apresentação do Projeto .....	23
OBJETIVOS .....	24
MÉTODOS.....	25
CRONOLOGIA.....	31
LIMITAÇÕES.....	32
AGRADECIMENTOS.....	33
REFERÊNCIAS BIBLIOGRÁFICAS .....	34
ANEXOS .....	39
I – Plataforma <i>SmashMedicine</i> ( <a href="http://www.smashmedicine.pt">www.smashmedicine.pt</a> ).....	39
II – Planeamento das tarefas a realizar antes do projeto iniciar, antes da competição iniciar e durante a competição .....	43
III – Guião para Construção de Perguntas e Testes de Escolha Múltipla .....	45

## RESUMO

A avaliação nas escolas médicas nacionais e na prova final de acesso à Formação Especializada assenta predominantemente em testes escritos, baseados em perguntas de escolha múltipla (PEM).

Enquanto que no passado se defendia que as PEM avaliavam apenas a recordação de factos isolados, agora sabemos que, se corretamente construídas, as PEM podem ser tão válidas para testar o raciocínio clínico e a capacidade de resolução de problemas como outros instrumentos ou tipologias de teste.

Com o intuito de dirigir o seu estudo a este tipo de testes, os estudantes de medicina procuram, frequentemente, bancos de perguntas formativos.

O presente projeto pretende ultrapassar algumas das desvantagens destes bancos, nomeadamente, o seu elevado custo, a ausência de garantia de controlo da qualidade das suas perguntas e a sobrecarga horária dos docentes que os impede de se dedicarem à construção de perguntas formativas. Aliando estas desvantagens ao conhecimento de estudos prévios do sucesso da escrita de PEM pelos estudantes, o principal objetivo deste projeto é o de apoiar a preparação dos estudantes do 6º ano da Faculdade de Medicina da Universidade de Coimbra para a Prova Nacional de Acesso, através da criação de um banco de perguntas formativas construídas pelos mesmos.

Integrado no internacional *SmashMedicine*, que congrega 5 escolas médicas europeias, este projeto consiste numa competição entre os estudantes de 6º ano, em diferentes fases, estruturadas em torno das especialidades clínicas avaliadas na prova. A organização do projeto consta de 5 etapas: 1) a escrita de PEM por parte dos estudantes, 2) a revisão das mesmas pelos pares, 3) a revisão especializada pelos docentes, 4) realização de um teste-piloto para avaliar a qualidade das perguntas e por fim, 5) a criação de um banco de perguntas formativas.

**Palavras-Chave:** Educação Médica, Avaliação, Perguntas de Escolha Múltipla, Banco de perguntas, Escrita de perguntas, Prova Nacional de Acesso, Métodos de avaliação, Abordagens ao ensino

## ABSTRACT

Assessment, in both medical schools and in the Portuguese Medical Licensing Examination, consists mostly of written examinations with multiple-choice questions (MCQ).

Although it was previously defended that MCQ assessed merely the recall of fact, now it is known that, if correctly written, this method can be as valid as others in assessing clinical competence and problem solving.

In order to study in conformity with the way they will be assessed, medical students frequently favor online question banks.

This project intends to overcome some of the disadvantages of these question banks, being them their high cost, absence of guarantee in quality control of the questions and the lack of time of teaching staff to develop high quality formative questions. Combining these disadvantages together with the success shown in previous studies of students writing multiple-choice questions, the main objective of this project is supporting sixth year medical students of Faculty of Medicine, University of Coimbra, in their preparation for the Medical Licensing Examination by creating a question bank with questions written by the students.

Embodied in the international *SmashMedicine*, that brings together 5 European medical schools, this project consists of a competition among sixth year students in the different medical subjects assessed in the licensing examination. This project is summarized in a 5 step approach: 1) student question writing, 2) student peer review, 3) senior clinician review, 4) pilot test to evaluate the quality of the questions and 5) multiple-choice question bank creation.

**Key Words:** Medical Education, Assessment, Multiple-Choice Questions, Question bank, Question writing, Portuguese Medical Licensing Examination, Assessment methods, Learning approaches

## LISTA DE ABREVIATURAS

ACSS – Administração Central do Sistema de Saúde

ANEM – Associação Nacional de Estudantes de Medicina

CEMP – Conselho de Escolas Médicas Portuguesas

CNIM – Conselho Nacional do Médico Interno

EIT – *European Institute of Innovation & Technology*

FMUC – Faculdade de Medicina da Universidade de Coimbra

GEM – Gabinete de Educação Médica

GPNA – Gabinete para a Prova Nacional de Acesso

NEM/AAC – Núcleo de Estudantes de Medicina da Associação Académica de Coimbra

NBME – *National Board of Medical Examiners*

OSCEs – Exames Clínicos Objetivamente Estruturados

PEM – Perguntas de Escolha Múltipla

PNA – Prova Nacional de Acesso

PNAS – Prova Nacional de Avaliação e Seriação

SBA – *Single-Best Answer*

USMLE – *United States Medical Licensing Examination*

# INTRODUÇÃO

## Importância da avaliação

A avaliação das aprendizagens tem assumido um papel predominante nas preocupações dos decisores e gestores curriculares das escolas médicas em todo mundo<sup>1 2 3</sup>, sendo este facto traduzido num volume assinalável de publicações científicas sobre esta temática, com centenas de artigos publicados todos os anos em revistas de educação médica tais como a *Medical Education* e a *BMC Medical Education*, nas quais, no ano de 2019, foram publicados cerca de 172 e 388 artigos sobre a avaliação, respetivamente.

A avaliação é uma componente fundamental tanto do ensino como da aprendizagem, uma vez que pretende espelhar o conhecimento que os estudantes adquirem e o nível em que este se encontra.<sup>1</sup> Permite ainda, ao próprio, detetar e resolver as falhas na sua aprendizagem.<sup>4</sup> Para avaliar corretamente, será necessário delinear objetivos a cumprir, criar oportunidades para que estes se cumpram, verificar a extensão do seu cumprimento e reformulá-los, de modo a adaptar o ensino e melhorar a aprendizagem.<sup>5</sup> Perante esta perspetiva, a avaliação deve ser vista como um método de promoção de aprendizagem e não de produção de resultados<sup>6</sup>, sendo um dos seus objetivos nucleares o *feedback* transmitido aos estudantes acerca da sua prestação passada, de modo a criar uma base sobre a qual reparar os erros anteriormente cometidos.<sup>2 6</sup>

A avaliação pode classificar-se em dois tipos: formativa e sumativa.

Na primeira, a preocupação central reside na recolha de dados para reorientação do processo de aprendizagem, facilitando a identificação de falhas, de aprendizagens ainda não conseguidas e aspetos a melhorar.<sup>1 4 7 8</sup> Implica providenciar *feedback* em momentos-chave no percurso da aprendizagem dos estudantes, definindo os passos necessários para uma melhoria no processo ensino-aprendizagem.<sup>1 4 7 8</sup> Este processo aprofunda ainda a autoaprendizagem, permitindo que o estudante utilize o *feedback* que lhe foi dado e o aplique na melhoria da sua aprendizagem.<sup>8</sup>

A avaliação sumativa, por outro lado, está fortemente associada aos conceitos afins de classificação e seriação.<sup>1 4 8</sup> Tem como objetivo principal a verificação do nível de aquisição dos conhecimentos e competências por parte dos estudantes, em momentos específicos, com o propósito de reportar a diferentes entidades que podem incluir as instituições de ensino, os próprios estudantes e outras entidades governativas ou reguladoras com interesse direto na seleção ou recrutamento futuro daqueles.<sup>1 4 8</sup>

O enfoque excessivo na avaliação sumativa tem consequências importantes nas abordagens e comportamentos dos estudantes perante o ensino e a aprendizagem<sup>1 4 8</sup>. Este é um aspeto sobejamente reconhecido na ciência educacional, com tradução no célebre chavão: “*Assessment drives learning*”.<sup>2 4 6 9</sup> Trata-se de uma ideia-chave em Educação Médica, que enfatiza o papel determinante da planificação dos métodos e instrumentos de avaliação na aprendizagem.<sup>2 4 6 9</sup> Sintetizando, esta ideia parte do princípio de que os estudantes focam o seu esforço e estudo, baseados na perceção que têm dos objetivos e conteúdos que serão avaliados e sobre a forma como serão avaliados.<sup>2 6 10</sup> Isto advém de anos sucessivos de experiência educativa assente na primazia da obtenção de notas positivas em avaliações e não da perceção da importância dos conteúdos avaliados e da integração de conceitos.<sup>11</sup> A situação descrita mostra a dicotomia entre aprendizagens superficiais e profundas. Estas revelam estímulos, atitudes e comportamentos cognitivos muito diversos, categorizados nas abordagens ao ensino.<sup>8 12</sup>

Os comportamentos de carácter cognitivo, afetivo e psicossocial que revelam o modo como os estudantes percecionam, interagem e respondem ao ambiente de aprendizagem, constituem as abordagens ao ensino.<sup>12</sup> Através da sua identificação, as instituições de ensino podem implementar formatos de ensino compatíveis com estas abordagens.<sup>12</sup> No entanto, é importante referir o seu carácter mutável, podendo ser influenciadas pela motivação incutida pelo professor, perceção da importância da temática e ainda, pelo método de avaliação.<sup>12</sup>

As abordagens podem definir-se como profundas, superficiais e estratégicas.<sup>8 12 13</sup>

A aprendizagem profunda pretende combinar o aprender com o perceber, sendo constituída por uma base de conhecimentos e ideias prévias e pressupondo uma interação ativa com novos conteúdos.<sup>12</sup> Este processo implica o relacionamento das evidências e o registo de conclusões.<sup>12 13</sup>

No entanto, há conceitos que são melhor interiorizados se memorizados ou praticados até se tornarem automáticos. Aqui entra a aprendizagem superficial, sendo esta uma aceitação passiva de informação e de ideias que têm como enfoque os requisitos da avaliação.<sup>12 13</sup>

A aprendizagem estratégica acentua que a eficácia da aprendizagem e a obtenção de melhores resultados, pressupõem uma combinação das abordagens referidas – a profunda e a superficial. Pode, no entanto, levar a compreensão fragmentada dos conteúdos e integração pobre dos mesmos.<sup>12 13</sup>

Atualmente, com a quantidade de informação disponível e o mercado de trabalho em constante mutação, considera-se que deve ser dado ênfase à aprendizagem que pode ser transformada e aplicada a novas circunstâncias em detrimento de uma aprendizagem factual que se restringe ao contexto em que foi ensinada. Conclui-se, portanto, que a aprendizagem profunda deve ser mais encorajada e desde mais cedo.<sup>12 13</sup>

## Métodos e Instrumentos de Avaliação do Conhecimento

É possível identificar um conjunto muito alargado de metodologias e de instrumentos de avaliação no domínio da educação médica<sup>3 4 14</sup>, sendo a sua maior ou menor utilidade determinada, em grande medida, pelas competências a avaliar e pelo estágio ou nível do percurso curricular em que o estudante se encontra.<sup>9</sup> Esta ideia é bem ilustrada pela pirâmide proposta por Miller<sup>15</sup>, e resumida na Figura 1.

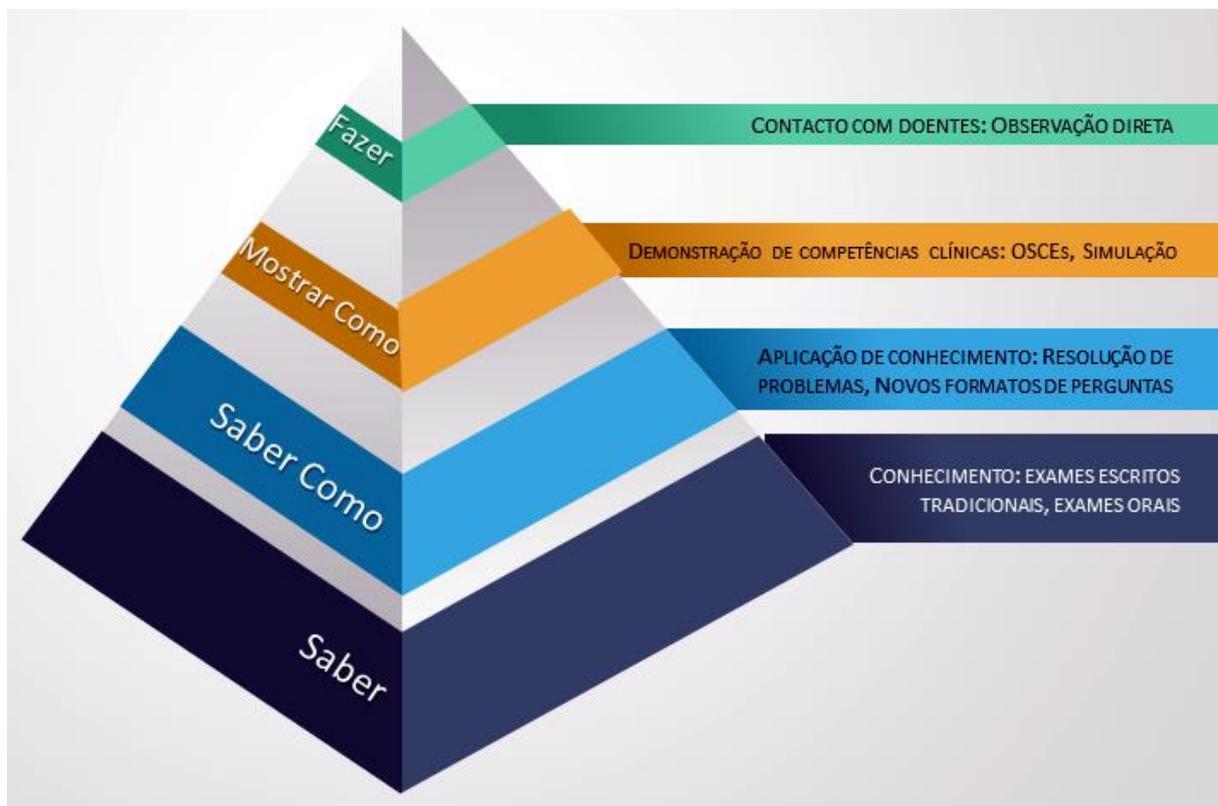


Figura 1. Pirâmide de Miller<sup>15</sup>

A Pirâmide de Miller<sup>15</sup> é usada como uma estrutura para guiar a avaliação nos profissionais de saúde.<sup>10</sup> Dividindo-se em quatro degraus (saber, saber como, mostrar como e fazer), os desenvolvimentos mais atuais na área de Educação Médica concentram-se no andar superior, “fazer”, que é o caso das tipologias de avaliação Simulação Clínica e Exames Clínicos Objetivos e Estruturados (OSCEs), sendo habitualmente agrupados na categoria de métodos de *assessment in the workplace*.<sup>10</sup> A avaliação nos níveis inferiores é dirigida ao

conhecimento factual, aplicação do mesmo e demonstração de aptidões, sendo já consagrada e escrutinada por exames escritos.<sup>10</sup> Atualmente, a definição de competência médica engloba todos os degraus da pirâmide, não dando primazia a nenhum e interpretando a pirâmide como um contínuo em que, para chegar aos degraus superiores se tem de passar pelos inferiores.<sup>16</sup> Mais ainda, abandonou-se a ideia de que para avaliar uma competência seria necessário um método específico, substituindo esta pela utilização de programas de avaliação, integrando vários instrumentos para avaliar a mesma competência, permitindo uma visão global do conhecimento e competência dos estudantes.<sup>10 16</sup> Assim, aquando do desenho de uma avaliação e escolha do método mais indicado para avaliar os conteúdos pretendidos, é importante o conceito de que “nenhum método é ideal para avaliar tudo”.<sup>1 4 5 9 10 16</sup>

Muitas destas metodologias e instrumentos de avaliação estão para além do âmbito deste trabalho, que se centra nos métodos mais adequados à testagem de conhecimentos e da sua aplicação, mais concretamente os relativos aos testes escritos que incluem Perguntas de Escolha Múltipla (PEM).

Comparativamente a outro tipo de instrumentos de avaliação, as PEM permitem que: 1) o número de itens a avaliar seja consideravelmente maior, discriminando mais os resultados e possibilitando uma maior abrangência de temas, potenciando assim uma maior fidedignidade e validade de conteúdo dos testes; 2) a correção e classificação seja feita por um computador, tornando o processo mais económico e rentável, tanto em termos de recursos materiais como humanos e de tempo; 3) comprovadamente, a prestação que os estudantes tinham em exames deste tipo, previa o desempenho que teriam como clínicos em certas situações; 4) por se tratarem de respostas fechadas, seja reduzida a subjetividade da correção, e assim, a quantidade de erro e viéses.<sup>17</sup>

A construção de qualquer tipo de PEM constitui um desafio, uma vez que, se devidamente desenhadas, conseguem avaliar a capacidade de resolução de problemas e de raciocínio clínico, mas com um desenho mais pobre podem, tendencialmente, avaliar fatos triviais ou níveis cognitivos inferiores.<sup>14 4 3</sup>

O universo de PEM pode ser dividido em duas grandes famílias: as perguntas que implicam que o examinando escolha todas as respostas corretas a uma questão (Verdadeiro/Falso) e as que requerem que os examinandos escolham a resposta mais correta para determinada pergunta (*Single-Best Answer* (SBA)).<sup>18</sup>

A família Verdadeiro/Falso acresce, ao examinando, uma maior dificuldade relativamente às perguntas SBA, uma vez que requiere um julgamento adicional necessário, que não se

relaciona diretamente com o conhecimento, dado que as opções apresentadas não são completamente verdadeiras ou falsas, detendo um grau de ambiguidade.<sup>18</sup> O examinador deve escrever as opções de forma homogênea, clara, de comprimento semelhante e o menos ambíguas possível, evitando termos como “frequentemente/geralmente” ou “muitas vezes”, que podem gerar discórdia, de modo a poder classificá-las universalmente como totalmente verdadeiras ou falsas.<sup>18</sup> Contudo, de modo a construir corretamente a pergunta, os autores caem muitas vezes na avaliação de recordação de um conteúdo isolado.<sup>18</sup>

A aplicação de conhecimento, integração e síntese são mais bem avaliadas usando SBA.<sup>18</sup> Estas podem classificar-se como de tipo R (“*Extended-Matching*” em conjuntos de 2-20 itens) ou tipo A (4 ou mais opções, itens singulares ou conjuntos)<sup>18</sup>, constituindo estas últimas o enfoque deste trabalho.

As perguntas do tipo A descrevem normalmente um cenário, constituindo o seu enunciado ou tronco, seguido de um *lead-in* que ajuda a clarificar a tarefa pedida aos estudantes, podendo esta revestir a forma de uma questão ou de uma afirmação, ambas estabelecendo a relação entre a pergunta e a lista de opções. Nesta lista, apenas uma questão será a correta e as restantes (três a cinco) os distratores.<sup>18</sup> À semelhança da família Verdadeiro/Falso, a sua escrita deve ser clara e sem ambiguidade.<sup>18</sup> No entanto, as opções não necessitam de ser inteiramente verdadeiras ou falsas, uma vez que o *lead-in* faz geralmente referência à “resposta mais correta, provável ou frequente”.<sup>18</sup> Todas as opções devem ser homogêneas e construídas de forma a que o examinando consiga responder à pergunta tapando as opções.<sup>18</sup>

A escrita de SBA é recomendada pelo *National Board of Medical Examiners* (NBME) por ultrapassar a ambiguidade gerada pela família Verdadeiro/Falso, uma vez que as opções não necessitam de ser inteiramente corretas ou erradas.<sup>18</sup>

## **Propriedades psicométricas dos testes**

A utilidade de um instrumento de avaliação é descrita como o produto de todas as suas propriedades psicométricas – validade, consistência ou reprodutibilidade, viabilidade, aceitabilidade e impacto educacional.<sup>9 19</sup> Deve ser estabelecido um compromisso entre estes cinco critérios aquando da escolha de um método, ao invés da opção por apenas um deles.<sup>19</sup>

Os dois atributos de maior importância – validade e consistência – são propriedades retiradas dos dados da avaliação e não inerentes ao método, daí que a escolha do teste deva basear-se no que se pretende avaliar, podendo o mesmo tipo de teste ser válido e reprodutível para avaliar certos conteúdos e outros não.<sup>9</sup>

A validade retrata a extensão na qual um teste avalia o que pretende avaliar, bem como se as inferências da competência dele retirado, são as corretas.<sup>9 14 16 20 21 22</sup> Dependendo da forma como é construído, um exame de PEM pode ser tão ou mais válido para testar níveis de conhecimento superiores e competências de raciocínio clínico do que outros métodos: se for adequadamente construído, poderá avaliar não só a memorização de factos, mas também a capacidade de interpretação, síntese e aplicação de conhecimentos.<sup>21</sup> Isto porque a validade não é determinada pelo formato da resposta, mas sim pelo conteúdo da pergunta.<sup>23</sup>

A consistência refere-se à reprodutibilidade dos resultados de um teste se realizado em diferentes circunstâncias<sup>9</sup>, indicando em que medida o resultado de um teste é generalizável.<sup>23</sup> Uma vez que os itens de um exame são uma pequena amostra de todas as possíveis perguntas relevantes que poderiam ser feitas, a consistência indica em que medida o resultado obtido em determinada questão pode ser extrapolado para outras.<sup>23</sup> Não obstante, mede também em que medida certo exame pode ser reproduzível se realizado noutro contexto espaço-temporal, ou corrigido por diferentes avaliadores.<sup>9</sup> Alguns dos fatores que podem diminuir a consistência de um teste são: o número reduzido de perguntas; o grau de dificuldade ser demasiado baixo ou alto; pouca objetividade, resultante de erros e vieses subjetivos na avaliação dos itens; condições subótimas na realização do exame e a heterogeneidade do grupo de respondentes, na medida em que se o conhecimento do grupo for muito heterogéneo, a consistência pode ser artificialmente elevada.<sup>9</sup>

A viabilidade infere sobre a exequibilidade de determinado método nas circunstâncias em que seria realizado, avaliando a existência de constrangimentos e de recursos disponíveis para a realização do mesmo.<sup>9</sup> Mais ainda, o tempo, o esforço e o custo envolvidos na construção, administração, correção e interpretação de dados de um exame devem justificar o uso do método em causa.<sup>9</sup>

A aceitabilidade define em que extensão os *stakeholders* no processo de avaliação aceitam o método usado e os resultados que dele advêm (estudantes, professores, membros da faculdade, etc.).<sup>14</sup> Os elementos que influenciam a aceitabilidade de um instrumento de avaliação têm de ser tidos em consideração aquando da escolha do mesmo, uma vez que nenhum método não aceite pelos estudantes e professores sobrevive.<sup>19</sup>

O impacto educacional relaciona-se com o aproveitamento que pode ser feito do conhecimento da influência da avaliação nos estudantes.<sup>16</sup> Uma vez que esta capitaliza a sua motivação para obterem bons resultados e direciona o seu estudo<sup>14</sup>, o método a utilizar deve ser escolhido com base na abordagem ao estudo pretendida, de modo a guiar os estudantes

para uma melhor aprendizagem.<sup>9</sup> Mais ainda, a avaliação em si deve fazer parte do processo de aprendizagem.<sup>19</sup>

## Tipologias de Perguntas de Escolha Múltipla e *guidelines* para a sua construção

As PEM constituem o método de eleição para avaliação em Educação Médica pelo que a evolução do seu processo de construção tem um peso determinante na qualidade global e na validade dos processos de avaliação.<sup>17 24</sup> O desafio será o de assegurar que as PEM são construídas de modo a que possam avaliar informação representativa dos reais conhecimentos e competências dos examinandos.<sup>17 24</sup>

A Taxonomia proposta por Bloom<sup>25</sup>, Tabela I, hierarquiza a complexidade cognitiva dos objetivos de aprendizagem em seis domínios ou categorias ordenados gradativamente – memorização, compreensão, aplicação, análise, avaliação e criação.<sup>26</sup> As capacidades propostas em cada categoria, fornecem uma estrutura útil para a construção categorizada de PEM.<sup>27</sup> Sabendo que os estudantes determinam o esforço e abordagem de estudo consoante o que percebem que seja avaliado, ao construírem perguntas enquadradas nos domínios superiores da taxonomia, os docentes estão a inculcar estas capacidades de ordem mais elevada, *higher-order thinking skills*, nos seus estudantes.<sup>6 28</sup>

Nível	Categoria	Explicação
1	Memorizar	Recordação de factos isolados ou informação
2	Compreender	Compreensão básica, explicação de ideias e conceitos
3	Aplicar	Utilização de informação em novas situações ou problemas
4	Analisar	Criação de conexões entre ideias
5	Avaliar	Combinar informação de modo a tirar inferências, justificar uma decisão ou posição
6	Criar	Produção de trabalho original

**Tabela I.** Taxonomia de Bloom<sup>25</sup>

Tradicionalmente, as PEM eram usadas como um meio para avaliar as *lower-order thinking skills* da Taxonomia de Bloom, tais como a memorização de factos/conhecimento e assim necessitavam de ser complementadas com perguntas de resposta aberta ou exames orais, concebidos para avaliarem *higher-order thinking skills* tais como a interpretação, a aplicação de conhecimentos e a resolução de problemas. De modo a obter um processo completo de avaliação, era consumido demasiado tempo, diminuindo o número de situações clínicas que poderiam ser avaliadas, comprometendo assim, a sua validade e consistência.<sup>17</sup>

A publicação de *guidelines* para a construção de PEM, como as de Case e Swanson<sup>32</sup>; Haladyna<sup>33</sup> ou mesmo aquelas elaboradas por organizações oficiais de referência, com poderes de certificação dos candidatos ao exercício da profissão, tais como o NBME<sup>18</sup> norte-americano, permitiram aperfeiçoar o processo de escrita de PEM, tornando-o mais eficaz, menos moroso e melhoraram a qualidade dos itens construídos a nível local, aproximando-os do objetivo de avaliação de *higher-order thinking skills*.<sup>22 32</sup> A Tabela II resume algumas das características ideais de uma PEM, bem como erros e falhas técnicas a evitar.

A incorporação da tecnologia representou, também, um avanço que muito contribuiu para este ponto, já que possibilita, por exemplo, complementar vinhetas clínicas com exames complementares de diagnóstico.<sup>17</sup>

	Orientações	Erros e falhas técnicas
Enunciado	<ul style="list-style-type: none"> <li>• Deve ser, preferencialmente, constituído por vinhetas clínicas</li> <li>• A ideia principal deve constar no enunciado</li> <li>• Cada item deve cobrir apenas um assunto</li> </ul>	<ul style="list-style-type: none"> <li>• Utilização de casos atípicos</li> <li>• Redigido na negativa, usando expressões como “todas são verdadeiras exceto”</li> </ul>
Lead-In	<ul style="list-style-type: none"> <li>• Deve ser claro e fechado</li> <li>• O examinando deve conseguir responder à pergunta apenas com o enunciado e <i>Lead-In</i>, tapando as respostas (Regra “Cover the Options”)</li> <li>• Deve pedir “a respostas mais correta” e não “a única resposta correta”</li> <li>• Os verbos devem ser incluídos no <i>Lead-In</i> e não nas opções</li> </ul>	
Opções	<ul style="list-style-type: none"> <li>• Devem ser homogéneas e unidimensionais (exo: todas as opções corresponderem a diagnósticos ou exames complementares de diagnóstico)</li> <li>• Devem ser escritas de forma clara</li> <li>• Os distratores devem apresentar algum grau de plausibilidade e encontrar-se num <i>continuum</i> de menos a mais correta</li> <li>• Devem ser ordenadas de forma lógica se possível</li> <li>• Os dados numéricos devem ser apresentados de forma consistente (todos em intervalos de valores ou valores independentes)</li> </ul>	<ul style="list-style-type: none"> <li>• Demasiado longas ou complicadas</li> <li>• Utilização de termos vagos como “frequentemente” ou “geralmente”</li> <li>• Ausência de homogeneidade, criando uma miscelânea de factos</li> <li>• “Nenhuma das opções anteriores” ser usado como opção</li> <li>• Erros gramaticais em uma ou várias opções que não se enquadrem no <i>Lead-In</i></li> <li>• Utilização de termos absolutos como “sempre” e “nunca”</li> <li>• A opção correta ser mais longa, mais completa e mais específica</li> <li>• A mesma palavra ser incluída no tronco e opções</li> <li>• Convergência (a opção correta inclui elementos em comum com as outras)</li> </ul>

**Tabela II.** Características ideais e falhas técnicas a evitar na construção de PEM<sup>18</sup>

## Enquadramento legislativo e evolução histórica da Prova Nacional de Acesso

Em Portugal, a necessidade de alteração da prova de acesso ao internato médico tem vindo a ser sentida há vários anos. A seriação de estudantes para o ingresso no internato médico manteve-se inalterável ao longo de quarenta anos, aproximadamente, vigorando desde os anos 70, o vulgarmente denominado “*Harrison*”. Estruturado como um teste de PEM, do tipo *SBA*, este baseava-se unicamente na obra americana *Harrison’s Principles of Internal Medicine*<sup>34</sup>, dando primazia à memorização de conteúdos nele abordados e pouco adequados à realidade epidemiológica portuguesa.<sup>35</sup>

Perante o panorama atual, pautado pela insuficiência de vagas de ingresso no internato médico, o acesso à especialidade tornou-se ainda mais competitivo, proporcionando um ambiente favorável à mudança do método de seriação, permitindo que uma nova prova emergisse, ganhando contornos definitivos.<sup>35</sup>

Exigia-se, desta forma, uma seriação mais justa e equitativa, com maior poder discriminativo, baseada no raciocínio clínico, na qual os concorrentes demonstrassem a integração do conhecimento adquirido na sua formação ao longo do curso de medicina e adequado à prática clínica expectável de um médico sem formação especializada.<sup>35</sup>

A alteração da prova foi proposta, inicialmente, em 2004, pelo Conselho Nacional do Internato Médico (CNIM), sendo que em 2007 foi criada, pelo mesmo, uma comissão que pretendia elaborar propostas para a modificação do exame de acesso à especialidade, até então designado Prova Nacional de Seriação. Entre os anos de 2009 e 2010, e após a realização de provas piloto com o apoio do NMBE, o Conselho das Escolas Médicas Portuguesas (CEMP) propôs ao Ministério da Saúde um modelo alternativo de prova, descrito como sendo “mais adequado e fiável para seriar os candidatos”.<sup>36</sup> Em dezembro de 2011, segundo o despacho nº 16696/2011, foi criado um Grupo de Trabalho para a revisão do regime do internato médico<sup>37</sup> cujo relatório final propunha a implementação de um novo modelo de prova com alteração dos conteúdos, considerando o modelo vigente desajustado.

<sup>36</sup>

Perante este relatório, e ao abrigo do despacho nº 13092/2012, de outubro de 2012, foi criado um Grupo de Trabalho para a definição do novo modelo de prova nacional de seriação relativa ao acesso ao internato médico a vigorar a partir do ano 2015.<sup>38</sup> Nesse seguimento, em maio de 2015, o Decreto-Lei nº 86/2015 procede à definição do regime jurídico da formação médica especializada com vista à obtenção do grau de especialista e estabelece os princípios gerais a que deve obedecer o respetivo processo.<sup>39</sup>

Em janeiro de 2016 foi criada, segundo o Despacho nº 642/2016, uma Comissão Nacional para o desenvolvimento do novo modelo da Prova Nacional de Avaliação e Sriação (PNAS) para o acesso ao internato médico. Esta Comissão foi presidida pelo Professor António Sarmiento e contou com a contribuição das escolas de medicina, Ordem dos Médicos, representantes das comissões responsáveis pelos exames anteriores, Administração Central do Sistema de Saúde (ACSS), Ministério da Saúde, Associação Nacional de Estudantes de Medicina (ANEM) e os Colégios das Especialidades de cada uma das especialidades a constar na nova prova.<sup>40</sup> Em março de 2017, esta Comissão remeteu à tutela o relatório com os pareceres finais e conclusivos, sendo uma das suas recomendações, a necessidade da profissionalização da estrutura responsável pelo processo de avaliação dos médicos para o acesso à formação especializada. Assim, em fevereiro de 2018, dando cumprimento ao artº. 35 do Decreto de Lei nº 13/2018, foi criado o Gabinete para a Prova Nacional de Acesso (GPNA)<sup>41</sup>, sendo no Despacho nº 3255/2018 de março de 2018 definidas a sua composição, missão e competências.<sup>42</sup>

Na sequência da publicação destes normativos, em maio de 2018, é publicado o Despacho nº4412/2018, que implementa o novo modelo da Prova Nacional de Acesso (PNA) e a matriz de conteúdos e referências bibliográficas relativas ao exame de 2019.<sup>43</sup>

A prova-piloto decorreu em novembro de 2018, sensivelmente um ano antes da primeira edição da PNA, e pretendeu testar procedimentos, aspetos logísticos e de validade, e outras propriedades psicométricas dos novos itens da prova.

Espelhando este modelo, a Prova Nacional de Acesso ocorreu, pela primeira vez, em novembro de 2019.

O GPNA, estrutura responsável pelo processo de avaliação dos candidatos à especialidade, reconhece a necessidade de uma maior experiência e a existência de recursos limitados<sup>44</sup>, pelo que nos primeiros cinco anos da realização da PNA, o NBME irá prestar apoio técnico e disponibilizar itens para a construção do exame.<sup>44</sup>

O reconhecimento do NBME como entidade de referência neste domínio advém de se tratar de uma organização cuja missão incide na proteção da saúde da população através da avaliação dos candidatos ao exercício de profissões de saúde e de profissionais no ativo, com uma experiência acumulada de várias décadas de prática avaliativa e de investigação sobre esta temática.<sup>45</sup> No que respeita à educação médica pré-graduada, o NBME é responsável pela realização e manutenção do *United States Medical Licensing Examination (USMLE)*<sup>45</sup>, exame constituído por três passos, tendo o segundo servido de inspiração para a configuração

da nova PNA. O NMBE dedica-se ainda à formação do corpo docente das escolas médicas, contribuindo para elevar a qualidade dos itens a nível local. Estudos realizados antes e após o curso, reportam, objetivamente, a melhoria pretendida.<sup>46</sup>

### **Adaptação à Prova Nacional de Acesso**

No que concerne ao Mestrado Integrado de Medicina da Faculdade de Medicina da Universidade de Coimbra (FMUC), o formato de item de teste predominante nas avaliações finais é a PEM. Embora muitos destes testes usem as *SBA*, a grande maioria não reflete a tipologia de perguntas presente na PNA, devido à ausência de enunciados com vinhetas clínicas e situações que requeiram uma decisão clínica aplicada ao contexto, acabando por avaliar, maioritariamente, a memorização factual.

Considerando o facto de os estudantes não estarem acostumados ao esforço cognitivo associado à necessidade de um raciocínio que exige a aplicação de conhecimentos, em avaliações deste tipo, será necessário apostar na adaptação, por parte das Escolas Médicas Portuguesas, quer do currículo lecionado, quer das abordagens ao ensino-aprendizagem, de modo a alinhá-los com as características e exigências da nova prova.

Por um lado, no que toca à avaliação, foi demonstrado em estudos prévios que a qualidade da escrita de PEM melhorava com a formação dos docentes envolvidos na mesma<sup>46 47 48 28</sup>. Esta evidência reforça o interesse e a necessidade de formação do corpo docente. Esta deveria ocorrer com o propósito de uma reformulação da escrita dos exames, indo ao encontro da tipologia da PNA, bem como das principais recomendações de boa prática e *guidelines* existentes no domínio da construção de PEM.

Por outro lado, “um exame tem sempre duas faces, a da avaliação e a de formação”<sup>35</sup>, pelo que o GPNA considera que a nova prova contribui com benefícios para a educação médica em Portugal, na medida em que valoriza “um maior contacto prático com doentes e cuidado extra por parte dos estudantes em observar doentes durante as valências clínicas na faculdade”.<sup>35</sup> Desta forma, também as escolas médicas devem dar primazia a este contacto e, quando este não é possível, potenciar as alternativas à realidade clínica, tais como a simulação médica.<sup>35</sup>No que toca ao primeiro ponto, o Gabinete de Educação Médica (GEM) da FMUC tem vindo a realizar cursos de formação de escrita de PEM para docentes. Desde o ano de 2008 foram realizadas onze formações, tendo sido formados, no total, 139 docentes.

## Construção de Perguntas de Escolha Múltipla pelos estudantes

Esta adaptação, considerada essencial, tanto dos currícula, como das práticas avaliativas dos docentes, foi já iniciada em outras escolas médicas, a nível internacional, tendo sido publicados alguns estudos sobre os efeitos, ao nível da cognição e dos resultados em exames, que a formação dos estudantes no processo da construção de PEM implica.<sup>49 50 51 52 53</sup>

O facto de as PEM/SBA constituírem o formato de avaliação mais utilizado pelas escolas médicas<sup>49 52 53 54 55 56</sup> e a sua atual transposição para a PNA, tem resultado na crescente procura por bancos de perguntas, com finalidades formativas, por parte dos estudantes de medicina, uma vez que replicam, diretamente, o formato de avaliação ao qual estarão expostos.<sup>49 52</sup>

A grande dificuldade na construção dos bancos de perguntas com função formativa prende-se com a exiguidade de tempo e a grande exigência de recursos humanos necessários à sua preparação.<sup>49 53</sup> Tendo os docentes outras tarefas relacionadas com os seus deveres assistenciais, de ensino e de investigação, que concorrem com aquelas associadas à avaliação, o tempo por eles dedicado à construção de PEM é dirigido aos testes de carácter sumativo e não à escrita de perguntas que pudessem constituir um banco dirigido aos objetivos formativos locais e personalizados para cada escola.<sup>49 53</sup> Assim, os estudantes recorrem frequentemente a bancos de perguntas incluídos em produtos comerciais, já bastante disseminados no mercado. Estes poderão encerrar algumas desvantagens, nomeadamente o seu elevado custo, o facto de poderem não estar adaptados ao currículo da escola ou até à realidade clínica do país e, em última instância, não haver garantia de controlo da qualidade das suas perguntas.<sup>53</sup>

Uma abordagem alternativa será a construção colaborativa destas perguntas com a participação dos estudantes, não só contribuindo para a construção destes bancos de perguntas, mas também para benefício da sua aprendizagem,<sup>50 51</sup> uma vez que a escrita de uma PEM requer conhecimento sólido, tanto sobre o assunto a avaliar, como dos mecanismos de construção da pergunta em si; sendo esta última uma vantagem para o desenvolvimento da capacidade de análise, compreensão e aplicação do estudante, contribuindo para a aquisição de *higher-order thinking skills*.<sup>52 54 27</sup>

A construção colaborativa de perguntas pelos estudantes tem sido experimentada noutras instituições de ensino estrangeiras<sup>49 51 52 53 55 57</sup>, seguindo um formato de 5 passos:<sup>49</sup>

1. Escrita, por parte dos estudantes, de PEM

2. Revisão das mesmas pelos pares
3. Revisão das mesmas por docentes, especialistas de cada área
4. Criação de um banco de perguntas formativas
5. Avaliação estatística dos itens e dos testes

Concluíram, em termos gerais, que as perguntas escritas pelos estudantes são de elevada qualidade<sup>32 49 50 52 55 57 58 59 26</sup>, capazes de avaliar *higher-order thinking skills*<sup>32 52 55 27 57 59 26</sup>, podendo a qualidade ser até comparada com as escritas por docentes<sup>50 58</sup>, com a mínima contribuição dos mesmos<sup>49</sup> e adaptadas ao ensino local.<sup>58</sup>

Os bancos de perguntas geraram muita recetividade por parte dos estudantes<sup>53</sup>, tendo sido demonstrada uma clara utilidade dos mesmos na preparação para os exames.<sup>49 55 58</sup> Detetou-se, em diversos estudos, uma correlação positiva entre a escrita de PEM por parte de estudantes e os resultados destes em exames sumativos<sup>51 53 27 58</sup>, demonstrando que o facto dos estudantes entenderem de forma sólida os mecanismos de escrita de PEM, contribuiu para um maior conhecimento dos mesmos.<sup>32 52 57</sup> Mais concretamente ainda, conclui-se que a frequência da escrita de PEM aumenta proporcionalmente o desempenho nestes exames.<sup>51 26 60</sup>

Como benefício adicional, a escrita de PEM pelos estudantes, contribuiu também para a avaliação de necessidades, ou seja o processo de diagnóstico de falhas na aprendizagem e *performance* e ulterior tomada de ações futuras para colmatar estas falhas<sup>61</sup>, bem como da qualidade da transmissão de informação no processo de ensino.<sup>50 52 55 59</sup>

Não pode deixar de ser mencionada a importância da revisão do processo por pares<sup>49 52 56 58 26</sup>, ainda que uma das limitações à utilização do banco, tenha sido a preocupação dos estudantes no que toca à qualidade das perguntas escritas pelos pares.<sup>51 52</sup>

O método, ainda que inicialmente pouco popular (muito pelo fator tempo), foi crescendo em termos de aceitação, à medida que os estudantes o experimentaram<sup>59</sup>, concluindo-se que a escrita de PEM constitui um fator motivacional para o estudo.<sup>57</sup>

Embora a grande maioria dos estudos tenham considerado a atividade como benéfica para os estudantes, importa dizer que alguns mencionam que o grau de dificuldade atribuído à execução da tarefa é elevado<sup>50 52</sup>destacando-se o tempo que consome.<sup>52</sup>

Importa ainda referir que os estudantes incluídos nos estudos, uma vez que na grande maioria são voluntários, podem não constituir a amostra mais representativa, sendo que podem, só por si, tratar-se de estudantes mais proativos e interessados na sua educação.<sup>53</sup>

De modo a ultrapassar as dificuldades de estudos já realizados, foi sugerido que nos próximos, as faculdades onde os projetos são implementados deveriam tornar a atividade o mais benéfica e encorajadora possível para os estudantes.<sup>32</sup> Neste sentido, a proposta mais concreta passou pela incorporação de PEM redigidas por estudantes em testes sumativos<sup>51</sup><sup>52</sup><sup>57</sup><sup>59</sup><sup>62</sup>, o que poderia constituir um incentivo adicional para a participação dos estudantes, bem como aumentar o grau de aceitação da tarefa.

Outra componente que se sugere que seja introduzida, de modo a aumentar a adesão, devido à sua capacidade comprovada de aumentar a motivação, participação e consequentemente aprendizagem<sup>63</sup><sup>64</sup>, é a gamificação.

A Gamificação define-se como o uso de elementos desenhados para jogos, em contexto que não o de jogo,<sup>65</sup> ou o processo de usar o pensamento e mecanismos de um jogo para resolução de problemas e motivar utilizadores.<sup>64</sup><sup>65</sup>

O conceito baseou-se no pressuposto de que, se os jogos têm a capacidade de entreter e prender a atenção do jogador por tão longos períodos de tempo, a transferência de elementos comuns a estes jogos para outros contextos, também poderia aumentar a atenção e participação.<sup>64</sup> A Gamificação faz, ainda, uso das vias de recompensa com a utilização de pontos, crachás e diferentes níveis<sup>64</sup><sup>65</sup>, bem como da Teoria da Autodeterminação que, aplicada ao conceito de Gamificação, defende que, num espetro que vai desde “amotivação”, a motivação extrínseca e por fim motivação intrínseca, a Gamificação pode atuar como pressão extrínseca que, se desenhada de forma competente, pode acabar por incutir motivação intrínseca nos participantes.<sup>64</sup>

Introduzida em diversas áreas como Finanças, Saúde, Comunicação Social e Sustentabilidade, mais recentemente tem sido aplicada à Educação no ensino superior.<sup>65</sup> Isto deve-se, além das demais razões relacionadas com os desenvolvimentos da tecnologia, à chegada ao ensino superior da geração jovem intitulada de *millennials*, geração esta que teve acesso universal à tecnologia na sua educação.<sup>63</sup> Dada a sua capacidade de motivação e aumento da adesão, é usada em competições extracurriculares, potenciando aprendizagens que não ocorreriam de outra forma.<sup>63</sup>

## Apresentação do Projeto

O projeto *SmashMedicine* teve origem nas Universidades britânicas de Oxford e Cardiff. Surgiu após o desenvolvimento de estudos na área da escrita de PEM por parte dos estudantes <sup>49 51</sup>, como uma plataforma *online* de escrita e revisão de perguntas para posterior uso formativo, integrada num processo mais amplo de inteligência artificial e gamificação, que visava premiar, mensalmente e anualmente, os estudantes cujas perguntas revelassem maior qualidade a nível da revisão, bem como atribuindo crachás e níveis a todos os participantes, de modo a promover a adesão à plataforma.

Após o sucesso a nível local, estas Universidades desenvolveram protocolos com outras instituições a nível europeu, nomeadamente a de Barcelona, Grenoble, Sorbonne e, atualmente, a FMUC.

Considerando a convergência de objetivos, cujo fim último é a melhoria da preparação dos estudantes para a PNA, desenvolveu-se o projeto conjunto Prep-PNA/SmashMedicine. A Figura 2 retrata a análise SWOT do projeto, analisando as suas forças, fraquezas, ameaças e oportunidades.

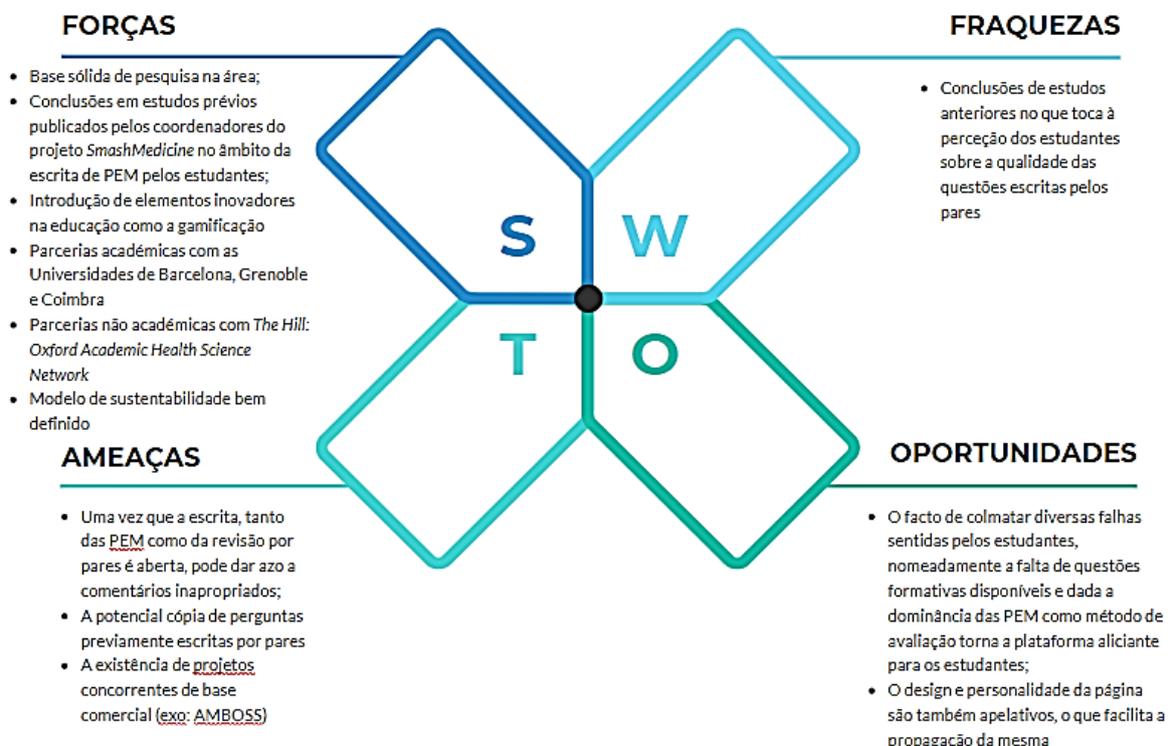


Figura 2. Análise SWOT do projeto Prep-PNA/*SmashMedicine*

## OBJETIVOS

Os objetivos do presente projeto incluem:

- Potenciar a aprendizagem dos estudantes e a qualidade do seu estudo através do efeito da construção colaborativa de PEM pelos próprios;
- Aumentar o conhecimento dos estudantes na construção de PEM, de modo a auxiliar na interpretação e análise em futuros exames com esta tipologia de questão;
- Apoiar os estudantes da FMUC na preparação para a PNA;
- Criar um banco de PEM formativas para uso dos estudantes de medicina nas diferentes áreas avaliadas pela PNA;
- Potenciar a avaliação por pares e colaboração entre os estudantes da FMUC;
- Aumentar a formação de docentes na escrita PEM através de Sessões de Formação;
- Promover a colaboração entre docentes e estudantes na criação de um banco de PEM de qualidade;
- Avaliar a qualidade das PEM construídas, em primeira instância através da colaboração com os docentes, e posteriormente através de um teste-piloto que pretende avaliar as propriedades psicométricas das PEM.

## **MÉTODOS**

### **Seleção de docentes**

O critério para seleção dos docentes relacionou-se com as especialidades presentes no projeto, sendo estas as mesmas avaliadas na PNA (Medicina Interna, Cirurgia, Pediatria, Obstetrícia, Psiquiatria) e algumas adicionais, nomeadamente Cardiologia, Pneumologia e Oftalmologia. Foram endereçados convites a pelo menos um docente por área, sendo que aderiram um total de dez docentes, todos médicos no Centro Hospitalar e Universitário de Coimbra e docentes da faculdade, alguns dos quais, regentes das cadeiras da respetiva área.

Todos os convites foram enviados via *e-mail*.

### **Seleção de participantes**

A participação no projeto será aberta a todos os estudantes que frequentam, no ano letivo de 2019/2020, o sexto ano do Mestrado Integrado em Medicina da Faculdade de Medicina da Universidade de Coimbra, não existindo limite máximo de número de participantes.

A divulgação do projeto será feita através das páginas das redes sociais *Facebook* e *Instagram*, tanto da Faculdade como do Núcleo de Estudantes local, o Núcleo de Estudantes de Medicina da Associação Académica de Coimbra (NEM/AAC), onde serão partilhados cartazes referentes à atividade, bem como um vídeo promocional realizado pela *SmashMedicine*.

Será divulgado um formulário na plataforma *Google Forms*, juntamente com os cartazes acima referidos, através do qual a inscrição no projeto será feita.

### **Desenho do projeto**

O projeto Prep-PNA/SmashMedicine foi planificado em duas fases: uma anterior ao início da competição, incluindo formações, em formato de *workshop*, dirigidas a docentes e estudantes e a competição propriamente dita. A Figura 3 representa, esquematicamente, as várias etapas do projeto.

**Antes da  
Competição**

Sessão de formação em  
construção de PEM



Sessão de apresentação da  
plataforma *SmashMedicine*

**Competição**

Escrita de PEM pelos estudantes



Revisão por pares



Revisão especializada por  
docentes



Realização de um teste piloto  
para avaliar a qualidade das  
perguntas



Criação de um banco de  
perguntas formativas

**Figura 3.** Etapas do projeto Prep-PNA/*SmashMedicine*

## **Antes da competição**

### **Sessões de Formação na escrita de PEM/SBA**

Foram desenhadas duas sessões independentes de formação para a escrita de PEM/SBA, com cerca três horas cada, uma direcionada aos docentes e outra aos estudantes.

A sessão de formação inclui uma contextualização inicial sobre a nova PNA e o seu novo estilo de perguntas bem como uma explicação dos conceitos *Assessment drives Learning* e da Pirâmide de Miller. Aborda também as características e qualidades dos instrumentos de avaliação. São dados exemplos, tanto de perguntas complexas, como de perguntas que tenham por base apenas a memorização. É explicada a construção de todos os componentes de uma pergunta e de como evitar erros e falhas técnicas. Por último, abordam-se indicadores estatísticos relevantes. A sessão permite que todos estes conceitos sejam exercitados ativamente pelos participantes, através de exercícios práticos realizados, cujos resultados são analisados e debatidos por todos os presentes. No final da sessão de formação, os formandos devem estar capacitados para:

- Distinguir as características das principais famílias e tipos de testes e formatos de perguntas;
- Selecionar as perguntas e os testes baseando-se em critérios de base científica, tais como a validade ou fidedignidade;
- Construir PEM, aplicando as recomendações técnicas e *guidelines* vigentes;
- Evitar erros na construção de perguntas devidos a erros e/ou falhas técnicas;
- Avaliar os resultados da aplicação de um teste, baseando-se na análise dos indicadores estatísticos relevantes

O processo formativo dos docentes integra a formação na escrita de PEM no modelo da nova PNA, aspeto fundamental à adaptação dos exames sumativos de modo a ir de encontro a este modelo e contribuindo para uma melhor preparação dos estudantes para a prova.

Ambas as sessões serão realizadas nas instalações da faculdade e ministradas pelo Dr. Hugo Camilo, SubDiretor do GEM da FMUC e colaborador neste projeto.

## **Sessões de Apresentação da Plataforma**

Serão também realizadas duas sessões, de cerca de 45 minutos, com o intuito de apresentar a plataforma a docentes e a estudantes. Para o efeito, será solicitado a todos os participantes, que se façam acompanhar dos seus computadores.

Cada sessão incluirá uma fase inicial de informação, sobre as funcionalidades da plataforma, seguida da aplicação prática, na qual todos os participantes, estudantes e docentes, irão ser apoiados no processo de criação de conta e início de sessão na plataforma.

Mais tarde, será pedido aos estudantes para construírem, na plataforma, uma PEM, bem como para reverem as construídas pelos colegas. Aos docentes, será solicitado que avaliem os itens contruídos pelos estudantes.

Será prestado apoio permanente durante a sessão, pelos coordenadores do projeto, quer nas explicações sobre o funcionamento da plataforma, quer no processo de construção das perguntas.

## **Competição**

O início da competição está programado para fim do mês de abril, terá uma duração prevista de 6 meses, com o final previsto para outubro de 2020.

A duração prevista, bem como a sua divisão temporal, prende-se com o número de especialidades envolvidas. Isto porque a cada mês corresponde uma especialidade (salvo algumas exceções, nas quais serão fundidas duas especialidades no mesmo mês), sendo que nesse período de tempo apenas serão construídas e avaliadas perguntas no âmbito dessa área médica. As especialidades envolvidas no projeto são as mesmas que constam da PNA (Medicina Interna, Cirurgia, Pediatria, Obstetrícia, Psiquiatria) e algumas adicionais, nomeadamente Cardiologia, Pneumologia e Oftalmologia.

Durante a competição, os docentes envolvidos no projeto deverão selecionar a pergunta de maior qualidade, a nível da construção e relevância. Os autores das perguntas serão premiados com uma oferta ainda a definir pela faculdade, através da entidade financiadora do projeto *SmashMedicine, European Institute of Innovation & Technology (EIT) Health*.

Anualmente, e desta vez a nível europeu, serão escolhidos os estudantes que mais tempo dedicaram à plataforma, sendo-lhes também atribuído um prémio (em maior escala), novamente com o fundo da *EIT Health*.

O funcionamento da competição *per si* pode ser descrito como um processo de cinco etapas, sendo estas: 1) a escrita de PEM por parte dos estudantes, 2) a revisão das mesmas pelos pares, 3) a revisão especializada pelos docentes, 4) realização de um teste-piloto para avaliar a qualidade das perguntas e 5) a criação de um banco de perguntas formativas. Esta organização é renovada todos os meses, com a alteração da especialidade de enfoque.

Toda a competição é feita através da plataforma construída pelo projeto *SmashMedicine* e adaptada para a língua portuguesa (Anexo I). Esta facilita uma navegação intuitiva no processo de construção e de revisão de PEM, permite responder a perguntas já avaliadas e arrecadar prémios como crachás e mudanças de nível, consoante a quantidade e qualidade de perguntas escritas e respondidas. Permite também, a criação de um banco de perguntas formativas de qualidade que pode ser usado para o estudo dos estudantes, como é objetivo deste projeto.

## **1. Escrita de PEM**

Após a sessão de formação inicial, será facilitado, aos estudantes do 6º ano, o Guião para construção de perguntas e testes de escolha múltipla da Faculdade de Medicina da Universidade de Coimbra, de modo a orientá-los na sua escrita.

Embora sem carácter obrigatório, será sugerida aos estudantes a escrita de pelo menos cinco perguntas por especialidade, sendo o tema (dentro da especialidade) à sua escolha, apelando à relevância epidemiológica do mesmo, isto é, que as condições em avaliação sejam tão frequentes quanto representativas da realidade do nosso país.

Cada PEM terá de ser constituída por um enunciado/tronco da pergunta, o *lead-in*, e por uma lista de opções, tendo de selecionar e justificar a resposta correta obrigatoriamente, sendo também recomendada a justificação das opções erradas ou distratores.

## **2. Revisão pelos pares**

A partir do momento em que uma pergunta é construída está automaticamente disponível para ser revista pelos demais estudantes participantes.

A revisão por pares tem início quando os estudantes respondem às perguntas dos colegas, podendo no fim deixar um comentário no qual avaliam a qualidade da questão em causa nas suas diversas vertentes, ou seja, o comentário tanto pode dizer respeito à componente formal da questão como ao conteúdo científico. A revisão implica uma maior a colaboração entre estudantes, uma vez que permite a troca de impressões e comentários.

### **3. Revisão especializada pelos docentes**

Aos docentes envolvidos e a quem competirá a revisão especializada do processo, foi também enviado o Guião para construção de perguntas e testes de escolha múltipla da Faculdade de Medicina da Universidade de Coimbra.

A partir deste, da formação recebida e do conhecimento científico que possuem na sua área, os docentes avaliam os itens, determinando a sua passagem ou não para o banco de perguntas formativas, com base no cumprimento das regras e *guidelines* formais para a construção de perguntas, na correção gramatical e precisão da terminologia médica, no rigor do conteúdo científico e na relevância para o nível de aprendizagem a que se destinam.

Apenas após esta revisão, poderão ser disponibilizadas as perguntas para constituírem o banco de perguntas formativas.

### **4. Realização de um teste piloto para avaliar a qualidade das perguntas**

Finda a competição, as perguntas serão integradas num teste piloto que será administrado aos estudantes do 6º ano da FMUC e, posteriormente, avaliado com indicadores estatísticos relevantes, nomeadamente a consistência interna do teste, índice de dificuldade dos itens, índice de discriminação dos itens e padrão de respostas.

### **5. Construção do banco de perguntas formativas**

No decorrer do projeto, todos os itens que completarem o processo de revisão, serão compilados por forma a integrar o futuro banco de perguntas formativas. Enquanto a competição estiver a decorrer, apenas os estudantes inscritos terão acesso livre ao banco de perguntas. Após a sua conclusão, todos os estudantes da Faculdade de Medicina da Universidade de Coimbra poderão aceder-lhe de forma gratuita.

## CRONOLOGIA

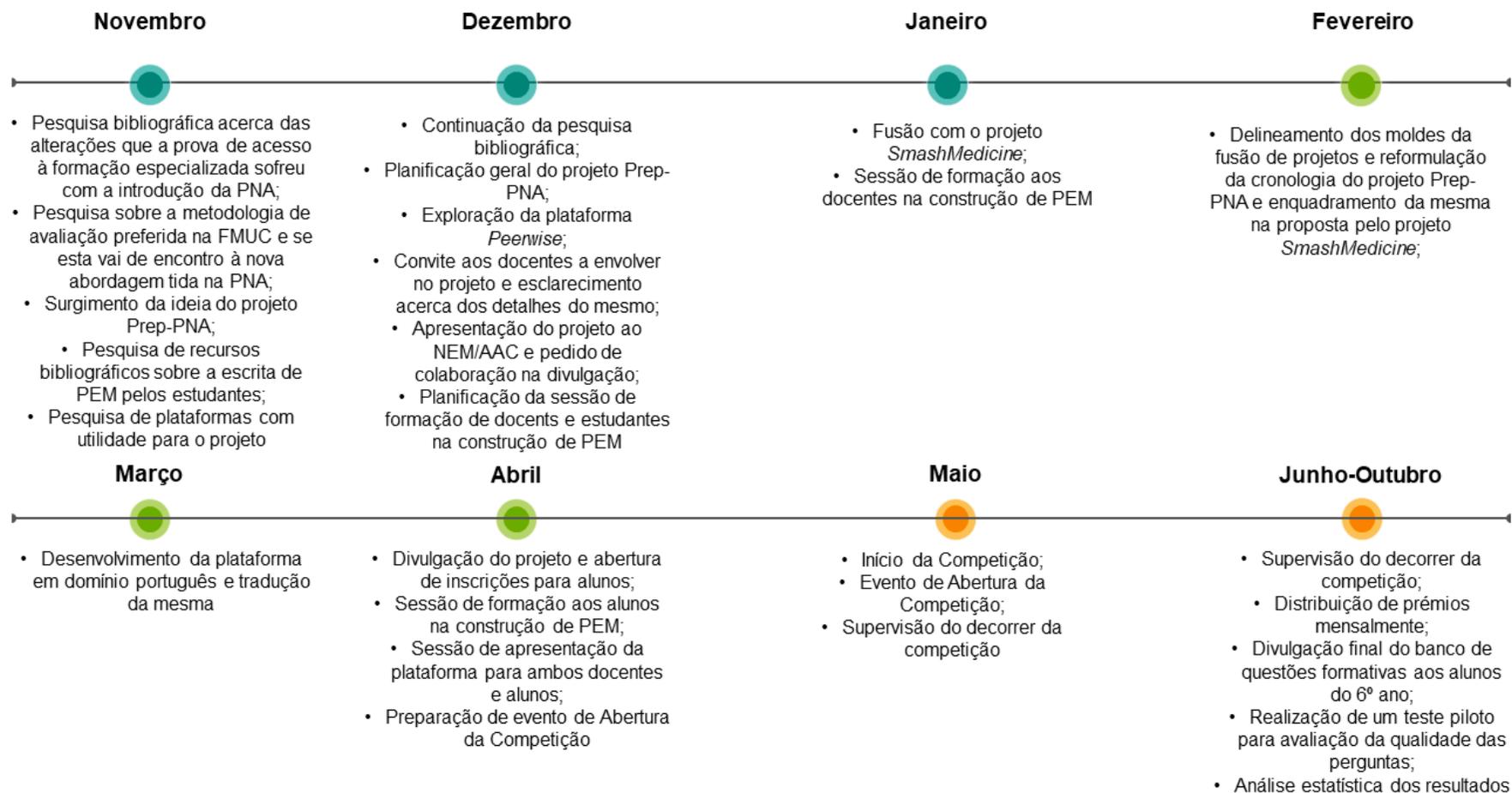


Figura 4. Cronologia do projeto Prep-PNA/*SmashMedicine*

## LIMITAÇÕES

As limitações na concretização deste projeto são recentes e devidas ao estado de pandemia que atravessamos pela infeção por SARS-CoV2, COVID-19.

Estas prendem-se com o facto de que, quer os docentes envolvidos, quer os elementos responsáveis pela coordenação do projeto *SmashMedicine* serem profissionais de saúde na primeira linha de contacto com a infeção, o que atrasou o início do projeto. Adicionalmente, o facto das aulas na FMUC estarem suspensas, impossibilita a formação presencial dos estudantes na construção de PEM e de ambos docentes e discentes no conhecimento da plataforma.

Estas limitações encontram-se atualmente a ser ultrapassadas com o uso de videoconferência.

Considera-se, no entanto, que no contexto atual, o projeto poderá constituir uma vantagem ainda maior para a preparação dos estudantes para a PNA, por se tratar de um meio digital de fácil acesso e permitir o estabelecimento de contactos regulares e melhor direcionados entre estudantes e docentes.

## **AGRADECIMENTOS**

Esta secção dedico a todos aqueles que contribuíram e que me apoiaram na realização deste projeto.

Ao Professor Doutor Lino Gonçalves e Dr. Hugo Camilo, por toda a disponibilidade e atenção, bem como contribuição técnica e científica, imprescindível à realização do Trabalho Final.

Ao Professor Doutor Carlos Robalo Cordeiro, Professor Doutor Henrique Borges Alexandrino, Professora Doutora Maria Del Carmen Bento Ribeiro, Professora Doutora Lèlita da Conceição dos Santos, Professor Doutor Rufino Silva, Professora Doutora Margarida Figueiredo Dias, Professor Doutor António Ferreira de Macedo, Dra. Joana Oliveira e Investigadora Doutora Ana Telma Pereira pela indispensável contribuição para o projeto e interesse pela pedagogia na nossa Faculdade, bem como por todo o apoio dado aos estudantes ao longo do seu percurso académico.

À minha família, pelo apoio incondicional ao longo de todos estes anos e paciência inesgotável. O maior agradecimento de todos, por fazerem de mim quem sou, e por todas as oportunidades que me deram. Nenhuma palavra seria suficiente para vos agradecer.

Aos meus amigos de Vila Real, cuja amizade sobrevive o passar do tempo e as mudanças de espaço, por celebrarem todas as minhas vitórias e atenuarem as derrotas, desde que me lembro.

Aos meus amigos e companheiros de Coimbra, por se tornarem a minha segunda família, me acompanharem e partilharem tudo o que vivi nestes anos, que, afinal, foram mesmo os melhores da minha vida. Um agradecimento adicional porque sem vocês não acabava o curso. Uma vez Coimbra, para sempre Saudade.

Ao NEM/AAC e à sua XXI Direção, por me ensinarem que consigo sempre mais do que o que penso, quando carregada em ombros de gigantes, como vocês. Por me despertarem o interesse pela área e partilharem do espírito inconformista. Por serem tanto de uma equipa, como de uma casa.

À Clara, à Leonor, à Lara e ao André, por serem as minhas pessoas.

## REFERÊNCIAS BIBLIOGRÁFICAS

1. Ferris H, O' Flynn D. Assessment in Medical Education ; What Are We Trying to Achieve ? *Int J High Educ.* 2015;4(2):139-144.
2. Lombardi MM. Making the Grade : The Role of Assessment in Authentic Learning. *Educ Learn Initiat.* 2008:1-16.
3. Ribeiro JSN. Organização da Avaliação da Aprendizagem no Curso de Medicina: Estudo de Caso. 2011.
4. Epstein RM. Assessment in Medical Education. *N Engl J Med.* 2007;356(4):387-396.
5. Suskie L. *Assessing Student Learning: A Common Sense Guide.* 2nd ed. Jossey-Bass; 2004.
6. Bezuidenhout MJ, Alt H. ' Assessment drives learning ' : Do assessments promote high-level cognitive processing ? *South African J High Educ.* 2011;25(6):1062-1076.
7. Gardner J. *Assessment and Learning.* 2nd ed. Sage; 2012.
8. Harlen W, James M. Assessment and Learning : differences and relationships between formative and summative assessment. *Assess Educ Princ Policy Pract.* 1997;4(3):365-379.
9. Sood R, Singh T. Medical Education Assessment in medical education : Evolving perspectives and contemporary trends. *Natl Med J India.* 2012;25(6):357-364.
10. Van der Vleuten C, Currie E, Schuwirth LWT, Scheele F, Driessen E, Hodges B. The assessment of professional competence : building blocks for theory development. *Best Pract Res Clin Obstet Gynaecol.* 2010;24(6):703-719.
11. Harlen W. The role of Assessment in Developing Motivation for Learning. In: Gardner J, ed. *Assessment in Learning.* Sage; 2012:171-182.
12. Mirghani HM, Ezimokhai M, Shaban S, van Berkel HJ. Superficial and Deep Learning Approaches among Medical Students in an Interdisciplinary Integrated Curriculum. *Educ Heal.* 2014;27(1):10-14.
13. Samarakoon L, Fernando T, Rodrigo C, Rajapakse S. Learning styles and approaches to learning among medical undergraduates and postgraduates. *BMC Med Educ.* 2013;13(1):42.
14. Norcini JJ, Mckinley DW. Assessment methods in medical education. *Teach Teach Educ.* 2007;23(3):239-250.
15. Miller G. The Assessment of clinical skills/competence/performance. *Acad Med.* 1990;65(9):63-67.

16. Schuwirth LWT, Van der Vleuten C. Changing education, changing assessment, changing research? *Med Educ.* 2004;38(8):805-812.
17. Pugh D, De Champlain A, Touchie C. Plus ça Change , plus C ' Est Pareil : Making a Continued Case for the Use of MCQs in Medical Education. *Med Teach.* 2019;41(5):569-577.
18. Paniagua M, Swygert K, Billings M, et al. *Constructing Written Test Questions For the Basic and Clinical Sciences.* Philadelphia,PA: National Board of Medical Examiners; 2016.
19. Van der Vleuten C. The Assessment of Professional Competence : Developments , Research and Practical Implications. *Adv Heal Sci Educ.* 1996;1(1):41-67.
20. Schuwirth L, Van der Vleuten C. Written assessment (ABC of learning and teaching in medicine). *Br Med J.* 2003;326(7390):643-646.
21. McCoubrie P, McKnight L. Single best answer MCQs : a new format for the FRCR part 2a exam. *Clin Radiol.* 2008;63(5):506-510.
22. Haladyna TM. *Developing and Validating Multiple-Choice Test Items.* 3rd ed. Routledge; 2004.
23. Schuwirth LWT, Van Der Vleuten CPM. Different written assessment methods: What can be said about their strengths and weakness? *Med Educ.* 2004;38(9):974-979.
24. Collins J. Writing Multiple Choice Questions for Continuing Medical Education Activities and Self-Assessment Modules. *Radiographics.* 2006;26(2):543-551.
25. Bloom B. *Taxonomy of Educational Objectives. Vol 1: The Cognitive Domain.* New York: McKay; 1956.
26. Galloway KW, Burns S. Doing it for themselves: Students creating a high quality peer-learning environment. *Chem Educ Res Pract.* 2015;16(1):82-92.
27. Bates S, Galloway KW, Riise J, Homer D. Assessing the quality of a student-generated question repository. *Phys Rev Spec Top - Phys Educ Res.* 2014;10(2).
28. Abdulghani HM, Irshad M, Haque S, Ahmad T, Sattar K, Khalil MS. Effectiveness of longitudinal faculty development programs on MCQs items writing skills: A follow-up study. *PLoS One.* 2017;12(10).
29. Coughlin PA, Featherston C. How to Write a High Quality Multiple Choice Question ( MCQ ): A Guide for Clinicians. *Eur J Vasc Endovasc Surg.* 2017;54(5):654-658.
30. Hingorjo MR, Jaleel F. Analysis of One-Best MCQs : the Difficulty Index , Discrimination Index and Distractor Efficiency. *JPMA - J Pakistan Med Assoc.*

62(2):142.

31. Hrynchak P, Glover Takahashi S, Nayer M. Key-feature questions for assessment of clinical reasoning : a literature review. *Med Educ.* 2014;48(9):870-883.
32. Case S, Swanson DB. *Constructing Written Test Questions For the Basic and Clinical Sciences.* 3rd ed. Philadelphia,PA: National Board of Medical Examiners; 2002.
33. Haladyna TM, Downing SM, Rodriguez MC. A Review of Multiple-Choice Item-Writing Guidelines for Classroom Assessment. *Appl Meas Educ.* 2004;15(3):309-333.
34. Kasper DL, Fauci AS, Hauser SL, Longo DL, Jameson JL LJ. *Harrison's Principles of Internal Medicine.* 19th ed. McGraw-Hill Education; 2015.
35. Ribeiro JC, Villanueva T. A Nova Prova Nacional de Seriação The New Medical Licensing Examination in Portugal. *Acta Med Port.* 2018;31(6):293-294.
36. Revisão do Regime do Internato Médico. 2012.
37. Despacho nº 16696/2011. *Diário da República II Série Nº236*.:48210.
38. Despacho nº 13092/2012. *Diário da República II Série Nº193*.:33489.
39. Decreto-Lei nº 86/2015. *Diário da República I Série Nº 98*.:2669-2674.
40. Despacho nº 642/2016. *Diário da República II Série Nº 9*.:1450-1451.
41. Decreto-Lei nº 13/2018. *Diário da República I Série Nº 40*.:1077-1085.
42. Despacho nº 3255/2018. *Diário da República II Série Nº 63*.:9303.
43. Despacho nº 4412/2018. *Diário da República II Série Nº 86*.:12582-12587.
44. Ferreira C. Gabinete para a Prova Nacional de Acesso. *NorteMédico.* 2018;(76):10-15.
45. Haist SA, Katsufakis PJ, Dillon GF. The Evolution of the United States Medical Licensing Examination (USMLE) Enhancing Assessment of Practice-Related Competencies. *JAMA.* 2013;310(21):2245-2246.
46. Karthikeyan S, O'Connor E, Hu W. Barriers and facilitators to writing quality items for medical school assessments – a scoping review. *BMC Med Educ.* 2019;19(1):123.
47. Naeem N, van der Vleuten C, Alfaris EA. Faculty development on item writing substantially improves item quality. *Adv Heal Sci Educ.* 2012;17(3):369-376.
48. Abdulghani HM, Ahmad F, Irshad M, et al. Faculty development programs improve the quality of Multiple Choice Questions items' writing. *Sci Rep.* 2015;5:9556.
49. Harris BHL, Walsh JL, Tayyaba S, Harris DA, Wilson DJ, Smith PE. A Novel Student-Led Approach to Multiple-Choice Question Generation and Online Database Creation,

- With Targeted Clinician Input. *Teach Learn Med.* 2015;27(2):348-352.
50. Wagener S, Timbil S, Schultz JH, et al. Development of a competency-based formative progress test with student-generated MCQs: Results from a multi-centre pilot study. *GMS Z Med Ausbild.* 2015;32(4).
  51. Walsh JL, Harris BHL, Denny P, Smith P. Formative student-authored question bank: Perceptions, question quality and association with summative performance. *Postgrad Med J.* 2018;94(1108):97-103.
  52. Grainger R, Dai W, Osborne E, Kenwright D. Medical students create multiple-choice questions for learning in pathology education: A pilot study. *BMC Med Educ.* 2018;18(1):201.
  53. Walsh J, Harris B, Tayyaba S, Harris D, Smith P. Student- written single- best answer questions predict performance in finals. *Clin Teach.* 2016;13(5):352-356.
  54. Palmer EJ, Devitt P. Constructing multiple choice questions as a method for learning. *Ann Acad Med Singapore.* 2006;35(9):604.
  55. Gooi ACC, Sommerfeld CS. Medical school 2 . 0 : How we developed a student-generated question bank using small group learning. *Med Teach.* 2015;37(10):892-896.
  56. Malau-Aduli BS, Zimitat C. Peer review improves the quality of MCQ examinations. *Assess Eval High Educ.* 2012;37(8):919-931.
  57. Green DH. Student Generated Exams: Testing and Learning. *J Mark Educ.* 1997;19(2):43-53.
  58. Denny P, Luxton-Reilly A, Simon B. Quality of student contributed questions using peerwise. *Conf Res Pract Inf Technol Ser.* 2009;95:55-63.
  59. Jobs A, Twesten C, Göbel A, Bonnemeier H, Lehnert H, Weitz G. Question-writing as a learning tool for students - Outcomes from curricular exams. *BMC Med Educ.* 2013;13(1):89.
  60. Song D. Student-generated questioning and quality questions: A literature review. *Res J Educ Stud Rev.* 2016;2(5):58-70.
  61. Gupta K, Slezzer C, Russ-Eft D. *A Practical Guide to Needs Assessment.* 2nd ed. (Sons JW&, ed.). Pfeiffer; 2011.
  62. Baerheim A, Meland E. Medical students proposing questions for their own written final examination : evaluation of an educational project. *Med Educ.* 2003;37(8):734-738.

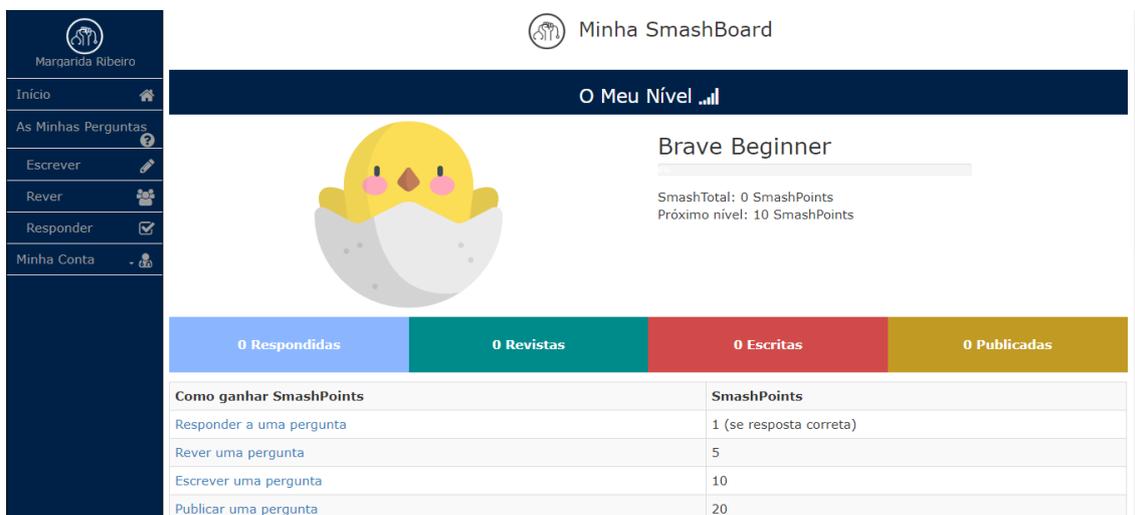
63. Nevin CR, Westfall AO, Rodriguez JM, et al. Gamification as a tool for enhancing graduate medical education. *Postgrad Med J*. 2014;90(1070):685-693.
64. Chrystal R, Walsh C, Swinger N, et al. Gamification in Action: Theoretical and Practical Considerations for Medical Educators. *Acad Med*. 2018;93(7):1014-1020.
65. Deterding S, Dixon D, Khaled R, Nacke L. From Game Design Elements to Gamefulness : Defining Gamification. 2011:8-15.

# ANEXOS

## I – Plataforma *SmashMedicine* ([www.smashmedicine.pt](http://www.smashmedicine.pt))



Fig.1 e 2 – Página Inicial



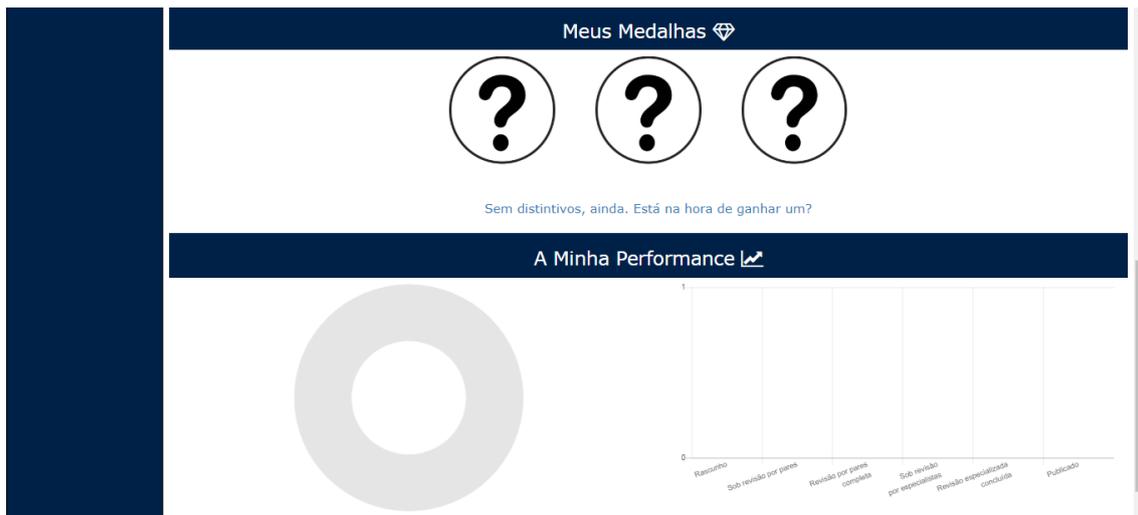


Fig. 3 e 4 – “SmashBoard”. Secção onde o estudante pode consultar o seu nível, medalhas e performance, bem como o estado das perguntas que escreveu.

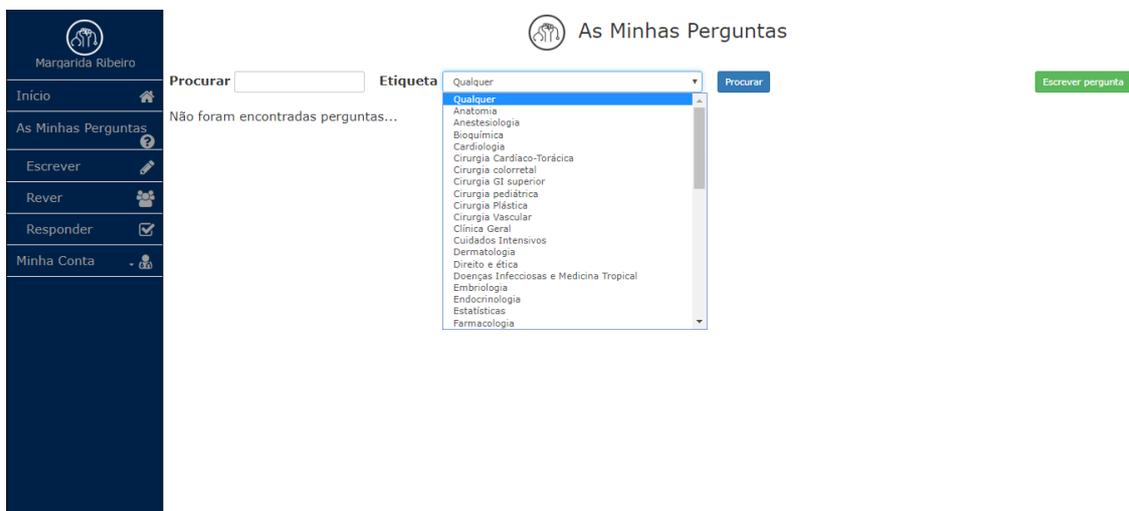


Fig. 5 - “As minhas Perguntas”. Secção destinada à consulta, por parte do estudante, das perguntas escritas pelo mesmo.

Fig.6 e 7 – “Escreva uma nova pergunta”. Secção destinada à escrita de novas perguntas.

Fig.8 – “Rever”. Secção destinada à revisão de perguntas escritas pelos colegas.

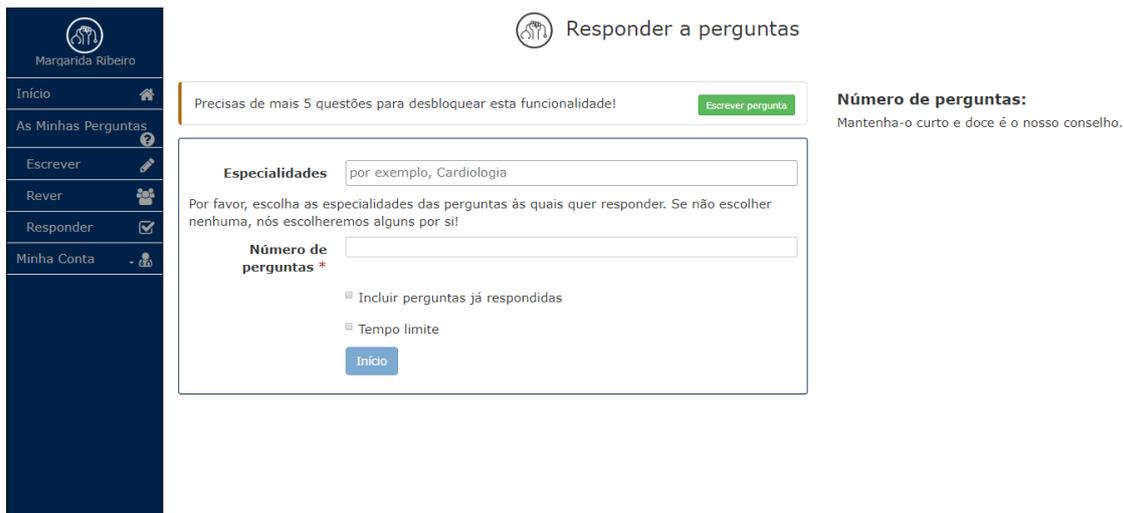


Fig.9 – “Responder a perguntas”. Secção destinada à resposta de perguntas já revistas.

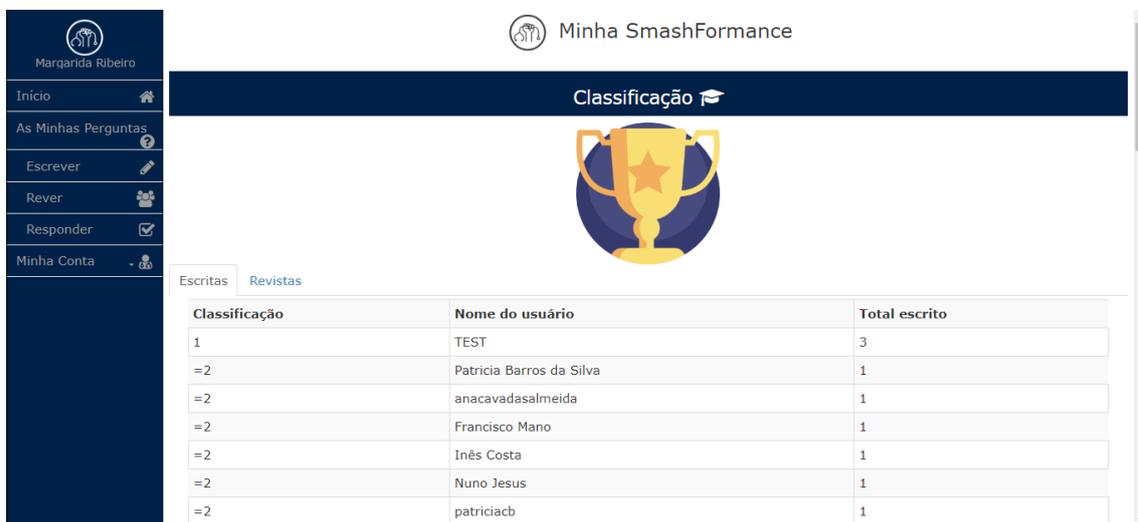


Fig.10 – “Minha SmashFormance”. Secção destinada à classificação geral dos estudantes.

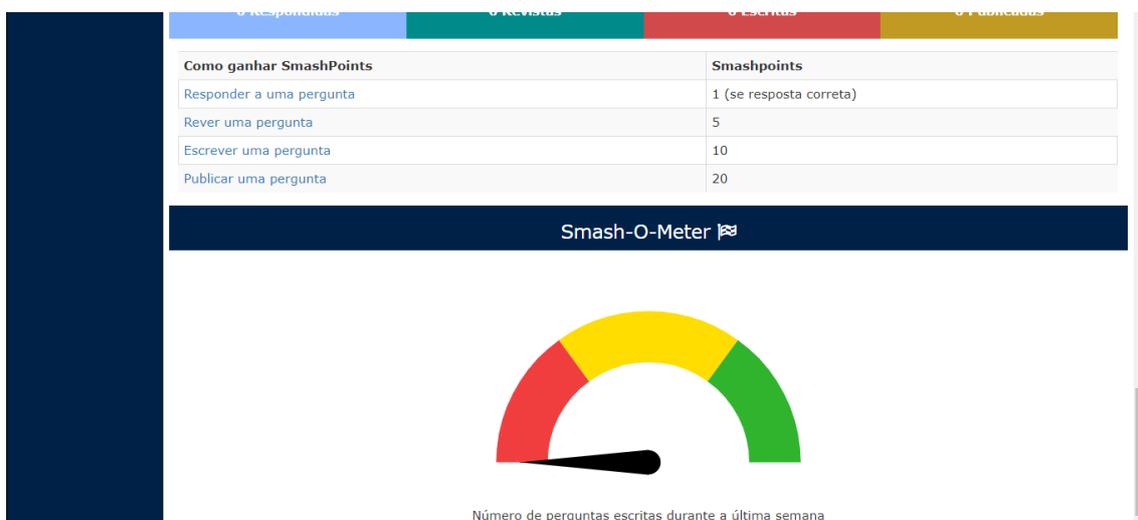


Fig. 11 – “Smash-O-Meter”. Secção destinada ao contador de perguntas semanal.

## II – Planeamento das tarefas a realizar antes do projeto iniciar, antes da competição iniciar e durante a competição

TAREFAS	
Antes do projeto iniciar	<ul style="list-style-type: none"> <li>• Pesquisa de recursos bibliográficos acerca das alterações que a prova de acesso à formação especializada sofreu com a introdução da PNA, em detrimento do vulgo “Harrison”</li> <li>• Pesquisa sobre a metodologia de avaliação mais utilizada na FMUC e a sua consistência com a nova abordagem tida na PNA</li> <li>• Pesquisa de recursos bibliográficos na área da Educação Médica sobre métodos de preparação para exames da tipologia presente na PNA</li> <li>• Criação do projeto Prep-PNA</li> <li>• Pesquisa de recursos bibliográficos sobre a escrita de PEM por estudantes</li> <li>• Pesquisa de plataformas online com utilidade para o projeto, isto é, com possível funcionalidade de criação de um banco de perguntas online, escrita e revisão das mesmas</li> <li>• Exploração da plataforma Peerwise</li> <li>• Delineamento das diferentes fases do projeto e cronologia do mesmo</li> <li>• Construção da lista de especialidades médicas a incluir no projeto, tendo em conta a sua relevância para a preparação dos estudantes para a PNA</li> <li>• Construção de uma lista de docentes passíveis de integrar o projeto, em função da seleção das especialidades médicas</li> <li>• Elaboração de guião a enviar como convite aos docentes</li> <li>• Convite aos docentes a envolver no projeto</li> <li>• Apresentação do projeto ao NEM/AAC, pedindo colaboração na divulgação do mesmo</li> <li>• Planificação das sessões de formação aos docentes e discentes na construção de PEM</li> <li>• Após fusão com o projeto SmashMedicine, reformulação da cronologia do projeto Prep-PNA para enquadramento nos moldes do projeto SmashMedicine</li> <li>• Organização da ordem das especialidades na competição, de acordo com a disponibilidade dos docentes;</li> <li>• Delineamento de critérios para definir os vencedores dos prémios mensais</li> <li>• Definição do tipo de prémios a atribuir mensalmente</li> <li>• Tradução da plataforma SmashMedicine</li> <li>• Criação da identidade visual do projeto e realização do material de imagem para a divulgação do mesmo</li> <li>• Criação do guião para divulgação junto dos estudantes</li> <li>• Criação do formulário para inscrição dos estudantes no projeto</li> <li>• Divulgação e abertura de inscrições para estudantes</li> </ul>
Antes da competição iniciar	<ul style="list-style-type: none"> <li>• Definição da data para a sessão de formação dos docentes em construção de PEM</li> <li>• Realização da sessão de formação dos docentes em construção de PEM</li> <li>• Definição da data para a sessão de formação dos estudantes em construção de PEM</li> <li>• Realização da sessão de formação dos estudantes em construção de PEM</li> <li>• Fornecimento do Guião para construção de perguntas e testes de escolha múltipla do GEM a ambos</li> </ul>

	<ul style="list-style-type: none"> <li>• Definição de data para a sessão de apresentação da plataforma SmashMedicine tanto para docentes como discentes</li> <li>• Realização da sessão de apresentação da plataforma SmashMedicine</li> <li>• Organização do evento de abertura da competição</li> </ul>
<p><b>Durante a competição</b></p>	<ul style="list-style-type: none"> <li>• Supervisão do decorrer da competição, recorrendo a feedback de todos os envolvidos, docentes e discentes</li> <li>• Distribuição de prémios mensais</li> <li>• Divulgação final do banco de perguntas formativas a todos os estudantes do 6º ano do Mestrado Integrado em Medicina da FMUC</li> </ul>

# 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
<b>"One-Best-Answer" (Resposta Única)</b>	Tipo A (convencional, 4 ou mais opções, itens simples ou sets) Tipo B (4 ou 5 opções de correspondência em sets de 2-5 itens) Tipo R (Extended-Matching em sets de 2-20 itens)	Ao contrário da família de verdadeiro/falso, este tipo de questões torna claro o número de opções a selecionar. 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	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.
<b>Verdadeiro/Falso</b>	Tipo C (A/B/Ambos/nenhum) Tipo K (verdadeiro-falso complexos) X (itens de verdadeiro-falso simples) Simulações (tais como os Patient Management Problems (PMP))	Esta família de testes requer que um aluno seleccione todas as opções que são "verdadeiras".  Exemplo:  <i>Afirmações verdadeiras sobre a fibrose quística incluem:</i> <ol style="list-style-type: none"> <li>1. A incidência de FQ é de 1:2000.</li> <li>2. As crianças com FQ geralmente morrem na adolescência.</li> <li>3. Homens com FQ são estéreis.</li> <li>4. A FQ é uma doença autossómica recessiva.</li> </ol>	O <i>National Board of Medical Examiners</i> dos EUA deixou de utilizar esta família de testes nos seus exames de acesso à profissão. 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"

## 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 subseqüentes, 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, seleccione (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, seleccione a (p.ex., 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 as 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 linha 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 As opções não devem ser complicadas ou duplas
Evite as negativas (p.e.x., "exceto") e duplas negativas. Se necessário, enfatizar com maiúsculas e negrito e usar opções apenas com uma palavra	Evite termos vagos ou imprecisos, tais como "frequentemente", "geralmente", "raramente", "algumas vezes", etc.
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 usar " <i>nenhum dos anteriores</i> " como opção
Evite a verbosidade (excesso de palavras), minudências irrelevantes ou conteúdo destinado a induzir em erro o aluno	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. Envolver, 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.e.x., 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:

- 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).
- 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).
- 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												Discriminação							
	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	%						
P1	1	3,0%	1	3,0%	1	3,0%	5	15,2%	25	75,8%	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	0,75	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 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:**



### 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 (3<sup>rd</sup> 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 falsepositives. *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 <b>V</b> se cumprir a recomendação
<b>Enunciado</b>	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	
<b>Opções</b>	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		
<b>Formatação e CONTEÚDO</b>	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.		