

Testes Escritos de Resposta Objectiva na FMUC: considerações sobre os seus formatos, fidedignidade (consistência) e qualidade das perguntas

Os testes de resposta objectiva (escolha múltipla, verdadeiro-falso, correspondência, etc.) têm vindo a granjear uma aceitação crescente junto da comunidade docente nesta Faculdade, verificável pelo elevado número de Unidades Curriculares que actualmente a eles recorrem para avaliarem sumativamente os seus alunos. Pensamos que esta realidade se fica a dever, em larga medida, a três ordens de razões: por um lado, o aumento significativo do número de alunos, associado a um tempo reduzido para correcção das provas no contexto do Processo de Bolonha, tem pressionado a aplicação de testes cuja correcção se revele mais fácil e célere; por outro lado, o reconhecimento de que também estes testes, sendo construídos de acordo com as recomendações técnicas neste domínio, podem avaliar níveis de complexidade cognitiva mais elevados, tais como a aplicação de conhecimentos ou a resolução de problemas; por último, a existência de um Serviço, prestado pelo Gabinete de Educação Médica, que faculta a leitura óptica dos testes e produz relatórios que coligem informações respeitantes à seriação e classificação dos alunos, à distribuição das respostas, à análise dos itens que os integram e aos valores de fidedignidade (consistência) do teste considerado globalmente.

A quase ubiquidade desta tipologia de testes associada aos dados que já pudemos obter (das Unidades Curriculares que beneficiam do Serviço de Análise de Testes) levam-nos agora a fazer uma breve síntese de algumas das características dos testes e tendências até agora verificadas, bem como a sugerir algumas recomendações para um uso mais eficaz e pedagogicamente válido dos testes de resposta objectiva.

Formatos

A literatura neste domínio descreve diversos formatos ou tipologias de testes de resposta objectiva, em que se incluem a escolha múltipla tradicional (1 opção certa em n), passando pelo verdadeiro-falso, extended-matching, até abordagens mais recentes e promissoras como a dos testes de Key-features. Há que salientar que a abordagem, pelo menos daqueles a que tivemos acesso, é esmagadoramente pelo formato tradicional de escolha múltipla (1 opção certa em n), com utilização pontual de verdadeiro-falso, ainda que nunca como formato exclusivo. Esta realidade é, em certa medida, positiva, face às reservas que os estudos e entidades tão relevantes como o NBME têm demonstrado relativamente ao uso do formato de verdadeiro-falso. De forma taxativa, este organismo afirma: *“we recommend that you do not use true/false questions”*, por se entender que estas são substituídas com vantagem por outros formatos. O reverso desta realidade é revelado por um uso incipiente de tipologias alternativas, como a de extended-matching ou os testes de Key-features, que os estudos têm revelado como promissoras, especialmente no domínio da educação médica.

Existe ainda outra dimensão importante a analisar quando nos referimos aos formatos dos testes, que têm que ver com o tipo de estímulo que é fornecido. Assim, podemos ter perguntas ricas em contexto, em que o enunciado da pergunta apresenta um contexto rico e detalhado, normalmente baseado num cenário clínico realista e que incita à aplicação do conhecimento e/ou à resolução de problemas, em oposição a perguntas pobres em contexto, que são desprovidas de contextualização, em que habitualmente o enunciado é curto e que revelam uma maior tendência para se situarem a um nível mais básico, de mera evocação de factos isolados. A este nível, a realidade ainda demonstra muito caminho a percorrer, com a certeza, porém, de que os esforços dos docentes estão a ser feitos, reconhecendo estes as vantagens das perguntas ricas em contexto, baseadas em problemas clínicos. O panorama é agora bastante mais positivo, com diversos docentes a manifestarem a intenção de incluírem um número cada

vez maior de perguntas ricas em contexto nos seus testes, com as vantagens e os resultados positivos que daí advirão.

Muito mais poderia ser dito acerca das regras orientadoras para a construção de perguntas de resposta objectiva, ou dos erros técnicos a evitar na construção destas, mas estas questões extravasam evidentemente o escopo deste artigo (sobre este assunto Vd. Essências Educare nº 7: “Construção de Testes de Escolha Múltipla”).

Indicadores de qualidade das perguntas (análise de distractores, Índice de Facilidade e Índice de Discriminação)

Quando computamos os resultados da leitura óptica de um teste de escolha múltipla, pretendemos não apenas obter as classificações dos alunos e médias e desvio-padrão do teste mas também alguns indicadores que permitam aferir da qualidade e fidedignidade do teste e das perguntas que o integram. Assim, a análise por nós efectuada permite apurar a distribuição de frequências de resposta nas opções-chave e distractores, bem como Índices de Facilidade e de Discriminação para cada uma das perguntas do teste. Este breve estudo incidiu sobre 14 dos testes por nós tratados em 2010, que obedeceram a duas condições: apenas se incluiu um teste por cada unidade curricular e apenas foram considerados os que foram respondidos por mais de 40 alunos, por imperativos de relevância estatística. Analisemos, então, o que vimos verificando relativamente a cada um dos indicadores considerados individualmente.

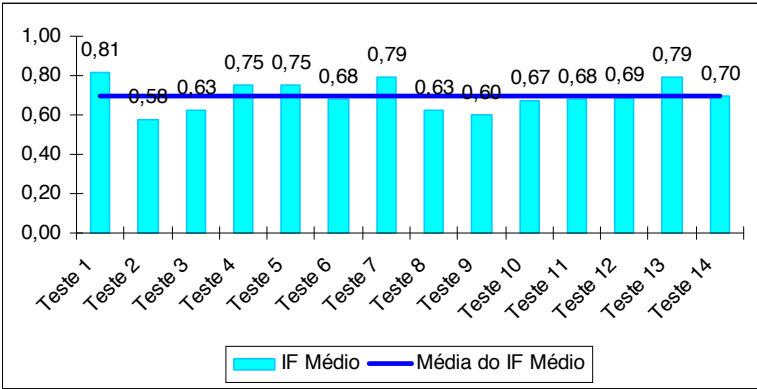
Índice de Facilidade (IF)

Para testes em que existe apenas uma opção correcta de resposta, como é o caso vertente, o Índice de Facilidade (IF) é tão somente a percentagem de alunos que respondem a um item correctamente, sobre o total. Este índice varia entre 0 e 1, sendo que, quanto mais elevado o valor, mais fácil é o item para o grupo de alunos respondentes.

No nosso caso (à semelhança do que é geralmente recomendado), estabelecemos como bons os itens situados entre 0,30 e 0,80. Fora deste intervalo, sinalizamos os itens como fáceis ou difíceis.

Podemos verificar (Gráfico 1) que, em termos médios, os IF se encontram dentro dos limites recomendados, isto é, que permitem o máximo de discriminação entre os alunos, conseguindo, no entanto, que uma percentagem significativa de alunos responda correctamente às questões. A Média dos IF Médios, de 0,70, é deveras adequada para testes de cinco opções, como são a maioria (90%) dos que estão em apreço.

Gráfico 1. Índices de Facilidade (Média)



Há, no entanto, alguns motivos para reflexão: em termos médios, metade das perguntas apresentam valores de IF superiores a 0,74, e 25% acima de 0,87. Isto significa que existem demasiadas perguntas demasiado fáceis, o que em muitos casos é compensado pela inclusão, no mesmo teste, de perguntas demasiado difíceis. Deparamo-nos com uma imensa terra de ninguém, por preencher, que deveria ser constituída por

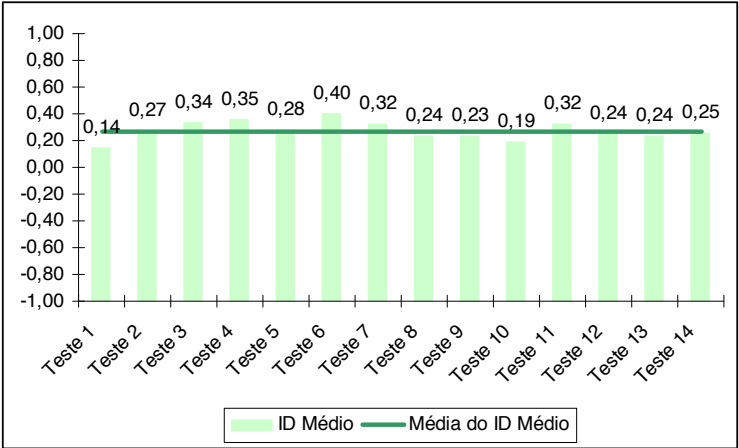
perguntas de IF intermédio, considerando que os estudos têm apurado que a fidedignidade (consistência) é maior quando os itens de IF moderado são predominantes. Acresce que os testes que incluem muitas perguntas de IF moderado tendem a ter uma maior dispersão de resultados ao longo da escala de respostas.

Índice de Discriminação (ID)

Este índice refere-se à capacidade das perguntas do teste em diferenciar entre os alunos, com base no domínio que estes demonstram sobre o conteúdo testado. O ID varia entre -1 e 1, e proporciona informação sobre o quão efectivo é um item a discriminar entre os piores e os melhores alunos num determinado teste. No nosso caso, utilizamos como referências o ¼ superior e ¼ inferior do conjunto de alunos e estabelecemos os seguintes intervalos aceitáveis: 0 a 0,10, fraco; 0,10 a 0,30, razoável; maior que 0,30, bom. É de salientar que o ID pode ser negativo, o que sucede quando o grupo dos piores alunos responde mais acertadamente que o grupo dos melhores, numa dada questão, o que é um sinal preocupante, como veremos adiante.

A análise dos dados recolhidos na FMUC, que é sintetizada no gráfico seguinte, indica-nos que as médias do ID se situam todas acima do limiar mínimo de aceitabilidade (0,10). A Média do ID Médio (0,27) aproxima-se bastante de valores considerados bons para este indicador, revelando, deste modo, uma razoável capacidade discriminativa das perguntas de teste, ainda que não a um nível de excelência, que deve ser almejado. Deve dizer-se a este respeito, que índices de facilidade óptimos são condição necessária, mas não suficiente para atingir níveis altos de discriminação.

Gráfico 2. Índices de Discriminação (Média)



Todavia, alguns problemas são-nos revelados quando nos detemos a observar a distribuição quartílica dos valores de ID. É evidente uma percentagem elevada de perguntas de teste com ID abaixo dos limites recomendados (cerca de 25% dos itens com ID inferior a 0,10). É ainda problemático o facto de, em 9 dos 14 testes, encontrarmos itens com ID negativos. Isto significa que os piores alunos responderam mais correctamente a essas perguntas do que os melhores, o que permite inferir que algo de errado se passa, ou com o item de avaliação (chave errada, pergunta falaciosa, incongruências gramaticais, etc.), ou com o próprio método de ensino (eventualmente, os alunos poderão estar a assimilar ideias ou conceitos erróneos).

Análise de Distractores

A análise dos distractores não é mais do que a análise da distribuição de frequências de resposta dos alunos às várias opções disponíveis. O modo como essa distribuição se efectiva permite averiguar a qualidade das opções alternativas à resposta correcta (denominadas opções distractoras).

A grande conclusão que se pode extrair da nossa amostra de testes é que existe uma quantidade considerável de opções distractoras que não são seleccionadas por nenhum aluno, ou então são no por um número insignificante. Em todos os testes analisados são várias as perguntas que

contêm uma ou mais opções que nenhum aluno assinala como resposta correcta. Isto traduz-se num problema que deve ser minimizado, porquanto estamos assim a transformar testes com 5 opções em testes com 2 ou 3 opções efectivas, reduzindo-lhes a capacidade discriminativa e aumentando a possibilidade de acerto nas respostas baseadas num conhecimento limitado e incompleto acerca dos conteúdos a testar. A este propósito, é altamente recomendável que os distractores que não são escolhidos por nenhum aluno sejam eliminados ou substituídos, em futuros testes, por opções plausíveis. Estes não estão a contribuir para a capacidade do teste discriminar entre os melhores e piores alunos.

Fidedignidade (Consistência Interna) do Teste

A consistência Interna, ao contrário dos restantes indicadores que temos vindo a analisar, não diz respeito às perguntas individuais, mas ao teste na sua globalidade. Este é um indicador que traduz a confiança que podemos depositar no nosso instrumento de avaliação de que a seriação dos alunos seria a mesma se voltássemos a aplicar o mesmo teste no futuro, ou uma forma equivalente, sem influência da primeira administração). A consistência de um teste reflecte três características de um teste:

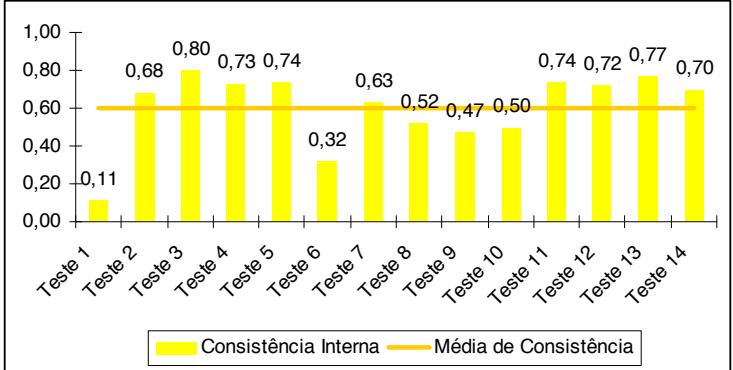
1. as intercorrelações entre os itens – quanto maior o número positivo de relações, e quanto mais fortes forem, maior a consistência;
2. A extensão do teste – um teste com mais itens tenderá a ter uma maior consistência, atendendo a que consegue avaliar uma amostra do conhecimento de maior dimensão;
3. O conteúdo do teste – geralmente, quanto mais heterogéneo for o conteúdo testado e mais divergentes forem as perguntas do ponto de vista estrutural, menor a consistência. A medida usada habitualmente para apurar a consistência é o Alpha de Cronbach (ou o coeficiente Kuder-Richardson-20, que é idêntico ao anterior quando as perguntas têm apenas uma opção correcta), que varia entre 0 e 1.

Uma consistência elevada significa que as questões nele incluídas tenderam a demonstrar coesão, isto é, os alunos que responderam correctamente a uma determinada questão revelarão maior probabilidade de ter respondido correctamente a outras questões.

Usámos como guidelines orientadoras as geralmente referidas na literatura:

- >0,90 – excelente, ao nível dos melhores testes standardizados;
- 0,80-0,90 - muito bom para um exame de uma unidade curricular;
- 0,70-0,80 – bom, dentro dos valores da maioria dos testes. Haverá provavelmente poucos itens que poderiam ser melhorados;
- 0,60-0,70 - relativamente baixo; o teste deve ser conjugado com outros instrumentos de avaliação; alguns itens provavelmente podem ser melhorados;
- 0,50-0,60 - sugere-se a revisão do teste, que tem que ser inevitavelmente suplementado por outros instrumentos de avaliação;
- <0,50 - consistência questionável. O teste não deve contribuir grandemente para a classificação dos alunos e necessita de revisão.

Gráfico 3. Consistência Interna (KR20)



A análise dos dados recolhidos (Gráfico 3) permite rapidamente extrair uma conclusão óbvia que merece a nossa reflexão: a consistência média é de apenas 0,60, o que é tangente ao limiar mínimo de aceitabilidade, de acordo com os nossos critérios de referência. Todavia, uma análise mais aprofundada revela que dois dos testes apresentam valores particularmente baixos (0,11 e 0,32), que penalizam significativamente o resultado médio global. A constatação desta realidade levou a que o Gabinete de Educação Médica lançasse um alerta aos docentes responsáveis, no sentido da revisão futura dos testes e da necessidade imperiosa de complementar estes testes com outros instrumentos de avaliação sumativa. Em geral, verifica-se que o número reduzido de perguntas incluídas nestes testes (40 e 7 perguntas, respectivamente) teve um impacto negativo significativo no valor de Consistência, mas não linear. Isto é, embora um dos testes apresentasse um nº de perguntas bastante superior a outro, isso não significou uma melhor consistência, pelo contrário. O Coeficiente de Correlação entre a Consistência e o número de perguntas do teste revelou-se positivo, mas moderado (0,31), como esperado.

Conclusões

Face ao exposto anteriormente, concluímos que os testes escritos de escolha múltipla realizados na FMUC revelam um comportamento globalmente positivo, segundo a perspectiva de análise que aqui adoptámos. Todavia, detectam-se igualmente áreas a aperfeiçoar, designadamente no que concerne à necessidade de melhoria dos índices estatísticos relevantes, da demanda por um maior rigor estrutural e de conteúdo que devemos imprimir às perguntas de teste, bem como na exploração de novos e promissores formatos de teste que vêm demonstrando resultados positivos, divulgados na literatura especializada. Outra grande conclusão a extrair é a de que este tipo de análise e as recomendações que daí emanam não seria possível sem o contributo de um sistema parcialmente automatizado de leitura dos testes e subsequente tratamento e reflexão sobre os dados apurados.

Causa Nostra

Curso TIPS em Reumatologia



Nos dias 8 e 9 do passado mês de Fevereiro, decorreu, nos H.U.C., mais um Curso TIPS de formação para docentes, como sempre dedicado aos temas do desenvolvimento curricular e da inovação pedagógica. Nesta ocasião, porém, este Curso assumiu um pendor diverso do que vinha acontecendo, já que se tratou do primeiro que resultou da organização conjunta entre o Gabinete de Educação Médica (GEM) e a Regência de uma Unidade Curricular/Valência: a Valência de Reumatologia do curso de Medicina (Unidade Curricular de Patologia Médica). As novidades introduzidas foram várias, sendo de ressaltar o facto da organização do curso ter partido da iniciativa da Regência, e dos próprios conteúdos e métodos formativos terem sido concebidos e apresentados em plena articulação entre aquela e o GEM e orientados para um público-alvo específico: todos aqueles com responsabilidades no ensino da Reumatologia. Pedimos ao Prof. Doutor José António Pereira da Silva, Regente da Valência de Reumatologia e co-organizador do Curso TIPS, que nos deixasse as suas impressões sobre esta iniciativa, que em seguida reproduzimos:

P:O que o motivou a propor esta colaboração com o GEM?

R: Pedi ao GEM para organizar um curso TIPS com os docentes da cadeira de Reumatologia por considerar que é essencial que todos os que ensinam conheçam os princípios básicos da educação médica, para que possam ter uma prática pedagógica reflexiva e estruturada, condições indispensáveis à melhoria do ensino e factor valioso para a sua satisfação profissional.

Acredito ainda que o simples facto de dedicarmos dois dias a pensar sobre “como ensinamos” é já factor de progresso e um importantíssimo contributo para o espírito de grupo e co-responsabilidade do corpo docente.

Novos Cursos TIPS FLEX: flexíveis, adaptáveis e à medida das necessidades

Apresente a sua proposta, diga-nos o que a sua Unidade Curricular necessita!

O Gabinete de Educação Médica passa a disponibilizar, de forma permanente, uma nova tipologia formativa, desenvolvida a partir da experiência com a Valência de Reumatologia, que consiste na oferta de cursos de duração e conteúdo variável, adaptados às necessidades e características específicas das diferentes unidades curriculares. Esta nova oferta formativa implica um trabalho de coordenação e organização conjuntos entre o GEM e os Regentes e outros docentes responsáveis que manifestem interesse em construir uma formação pedagógica adequada aos interesses e necessidades específicos da sua Unidade Curricular.

Os **Cursos TIPS Flex** incluem sempre um tronco comum que será desenvolvido em todos os cursos, que inclui debate e reflexão sobre tópicos relativos ao desenvolvimento curricular e aos métodos e técnicas de ensino-aprendizagem e de avaliação. A premissa básica que presidirá a todo e qualquer curso será sempre a do primado da Aprendizagem Activa, pelo que os métodos formativos a utilizar serão sempre fortemente influenciados por esta linha de orientação fundamental.

P: Que valor acrescentado traz este tipo de formação aos docentes e aos alunos da U.C.?

R: Para muitos docentes é uma absoluta novidade tomarem consciência de que há investigação muito profunda sobre as metologias de ensino de adultos e aperceberem-se de como uma abordagem estruturada ao que fazem resulta num aumento considerável de qualidade e eficácia do ensino.

Um dos aspectos mais surpreendentes para a maior parte dos docentes reside na mudança de paradigma do professor como alguém que ensina para alguém que facilita a aprendizagem, o aumento da responsabilidade do aluno como agente de aprendizagem, a mudança do foco pedagógico para o que se passa na cabeça do aluno e não na boca do professor.

Trata-se de uma pequena mudança que tem um enorme alcance no processo de ensino-aprendizagem mas também na atitude do aluno. No entanto esta mudança de paradigma não é fácil de apreender nas suas múltiplas nuances, um processo que exige reflexão, debate e manuseamento do conceito. O Curso TIPS é óptimo para isso.

Os alunos podem, numa primeira fase, reagir negativamente a esta nova responsabilidade mas rapidamente se apercebem que este novo modelo corresponde a um grau muito mais elevado de respeito pelas capacidades intelectuais do aluno e uma eficiência incomparável na construção do profissional de saúde, algo muito mais complexo do que um repositório enciclopédico de informação.





P: Pensa que outras U.C. poderão beneficiar de experiências deste tipo?

R: Não tenho qualquer dúvida: todas as unidades curriculares, todos os docentes e todos os alunos têm a beneficiar se estes cursos se multiplicarem. A escola beneficia com todos e cada um.

A ideia de reunir o corpo docente de uma só disciplina tem alguns aspectos negativos, como sejam o empobrecimento na exposição a outras experiências pedagógicas e menor interacção entre as diversas disciplinas. Acredito, contudo, que os benefícios são maiores: maior abertura ao diálogo, maior pertinência dos exemplos para todos e para a unidade curricular, maior contributo para o espírito de grupo na disciplina. A presença do regente reforçará a importância que este atribui a esta formação e aos ganhos de qualidade que dela espera.

P: Que balanço faz do curso realizado nos dias 8 e 9 de Fevereiro?

R: Um balanço extremamente positivo: saímos todos a saber mais, a reflectir mais nas melhores maneiras de ensinar, a esperar mais de nós próprios e dos nossos alunos. Contruímos, em conjunto, objectivos específicos de aprendizagem para a nossa unidade curricular, com que nos comprometemos, debatemos os métodos mais apropriados, identificámos comportamentos desejáveis e indesejáveis em alunos e docentes, reforçámos o empenho comum em procurar excelência no ensino e encontrámos muitos motivos novos para retirarmos satisfação do nosso trabalho docente e da pertença a este grupo.

Um último ponto. Tomaram parte no curso todos os médicos do Serviço de Reumatologia e não apenas os docentes do pré-graduado. Ensinar e comunicar é parte integrante do trabalho do médico - todos têm a ganhar e todos sentiram que ganharam.

Literatura em Educação Médica

A integração das novas tecnologias no ensino, e em contexto de aula, em particular, continua a ser um desafio para docentes e decisores curriculares. Os autores do artigo que revemos nesta edição propõem-nos um modelo para promover a integração das tecnologias no ensino, fundado nas teorias da aprendizagem, nos objectivos de aprendizagem a alcançar e no modo como a tecnologia pode suportar estes processos.

Irmgard U. Willcokson, Cynthia L. Phelps (2010) 'Keeping learning central: a model for implementing emerging technologies', Medical Education Online 2010, 15: 4275 -DOI: 103402/meo.v15i0.4275

O modelo, denominado eTIME, começa por criar uma solução preliminar através da consideração da tecnologia, da teoria e do problema a resolver. Após este passo, são aplicados métodos de design instrucional para chegar à implementação final. Os autores propõem três etapas na fase preliminar: 1. definição do problema de ensino-aprendizagem a resolver, 2. Consideração da teoria de aprendizagem e 3. Correspondência das potencialidades tecnológicas com o problema identificado. Para chegar à etapa de implementação final, os autores recomendam que se considere também a definição clara e explícita dos objectivos de aprendizagem e a determinação das características e necessidades do grupo específico de alunos, através de algum tipo de inquérito de avaliação.

Espera-se que este modelo sirva como um referencial e uma metodologia estratégica para a implementação de novas tecnologias no contexto de sala de aula.

ExcelênciasEDUcare

Inovação na Metodologia de Avaliação das Aprendizagens: o caso da Propedêutica Médica

Nesta edição damos a conhecer uma inovação educacional impulsionada pelos docentes da Unidade Curricular de Propedêutica Médica, do Mestrado Integrado em Medicina, que tomaram a iniciativa de reformular todo o processo de avaliação prática dos seus alunos. As mudanças impuseram-se pelo reconhecimento de fragilidades no sistema de avaliação vigente e pela vontade de introduzir mudanças na avaliação que melhor pudessem servir os interesses de alunos e docentes e aumentar a qualidade intrínseca do processo avaliativo. O modelo que serviu de referencial para a inovação foi a metodologia

OSCE (Objective Structured Clinical Examination), ainda que com adaptações decorrentes das limitações logísticas e de recursos ainda existentes para operacionalizar uma abordagem tão exigente como esta. O OSCE procura avaliar, nas suas múltiplas dimensões, a competência clínica de forma planificada, estruturada e objectiva pela observação directa da performance e interacção de médico/aluno-paciente, ao longo de um conjunto de estações. Uma das particularidades desta metodologia define-se na rotatividade, isto é, os participantes passam por todas as estações estruturadas de forma rotativa num intervalo de tempo pré-definido.

A Regente da U.C. de Propedêutica Médica, Profª Doutora Lelita Santos, acedeu a responder-nos a algumas questões que procuram clarificar os motivos das modificações operadas e o modo como todo o processo se desenrolou.

P: Descreva-nos genericamente a estrutura do exame e os moldes em que decorre.

R: O exame foi estruturado por 3 estações consecutivas: Est. 1 - "História Clínica"; Est. 2 - "Exame objectivo geral e do Aparelho Respiratório" e Est. 3 - "Exame clínico do Aparelho Cardio- Vascular". As Estações foram localizadas na Enfermaria. Em cada estação estavam um doente e um examinador que colocava questões, de uma listagem, consideradas essenciais para a avaliação das competências referidas. Esta listagem contém 10 grupos, cada um com 20 questões de exame clínico diferentes. O aluno responde e executa 3 questões/manobras por cada estação, excepto na 1ª em que apresenta a anamnese de acordo com uma checklist. Cada grupo de questões tem 3 níveis de cotação e um tempo máximo de execução que, no nosso modelo, corresponde a 20 minutos para o total das 3 estações.



Aluna aguardando pela entrada na Estação nº3



Aluno e Examinador na Estação nº 1



Exame Objectivo Geral

P: Que motivos a levaram a optar por esta abordagem à avaliação prática?

R: Os principais objectivos da avaliação prática, na disciplina de Propedêutica Médica são, testar o desempenho de competências clínicas e a aplicação dos conhecimentos adquiridos nas aulas práticas através de uma Anamnese estruturada, a comunicação oral dos dados clínicos do doente e a realização do exame clínico completo.

O exame prático clássico, em que o aluno responde às questões colocadas por um Assistente e faz algumas manobras de exame

físico, revela-se muitas vezes insatisfatório, por várias razões, que se prendem sobretudo, com critérios subjectivos do próprio examinador e com a pouca abrangência das competências avaliadas. O facto de existirem muitos alunos para avaliar, num curto período de tempo, exige um formato de exame, que sendo eficaz e justo, encurte os tempos de avaliação. Assim, a opção pela metodologia OSCE pode ser mais eficiente, pelo que este formato foi adaptado, no ano lectivo de 2009-2010, às avaliações práticas da disciplina.

P: Como referiu, este é um modelo que revela semelhanças com o OSCE. Esta aproximação é intencional?

R: Este modelo não representa o OSCE na íntegra, por dificuldades logísticas práticas que foram detectadas nos "treinos" previamente realizados entre os docentes. No entanto, contém objectivos e estrutura semelhantes: testa o desempenho de competências clínicas e a aplicação prática de conhecimentos; todos os examinados são confrontados com tarefas semelhantes; a classificação para cada estação é padronizada e isenta de subjectividade; cada estação consiste num cenário clínico; o aluno é avaliado através de uma checklist, exclusiva para cada estação, onde são testadas competências específicas. Na prática, para este exame, foi necessária uma organização prévia sólida e uma disciplina rigorosa no cumprimento do tempo previsto em cada estação. Sendo o exame feito na Enfermaria, foi necessário ter em conta variáveis relativas aos doentes e aos docentes: os exames de diagnóstico programados, a higiene diária, a "vontade" de colaborar, as intercorrências clínicas imprevisíveis, a visita médica e as múltiplas tarefas clínicas dos docentes (também médicos do hospital).

Igualmente, o facto de serem muitos alunos e, simultaneamente, poucos dias para a avaliação, dificultou a realização dos exames, agravado pela ausência de apoio administrativo.

No entanto, no balanço final, este formato de avaliação revelou-se mais adequado do que o anterior pois as falhas antes detectadas, quer pelos docentes, quer pelos alunos, foram ultrapassadas, nomeadamente: o desequilíbrio das classificações inerentes à subjectividade, as diferenças na exigência dependentes dos avaliadores, a variedade das competências avaliadas e o tempo total dispendido nos exames. Ou seja, houve maior uniformização dos critérios e equilíbrio nas classificações. Embora o tempo de preparação prévia tenha sido mais longo, os ganhos de eficiência nos dias do exame foram muito evidentes.

P: Que expectativas tem quanto ao futuro da avaliação das aprendizagens no Ciclo Clínico, e na sua Unidade Curricular, em particular?

R: Na aquisição de competências clínicas em Medicina o momento de avaliação é fundamental. O grande número de alunos no curso de Medicina condiciona não só a qualidade do ensino prático, que tem de continuar a ser feito à cabeceira do doente, baseado nos casos clínicos, como a avaliação das competências. Assim, é necessário procurar outras formas de avaliação que possam estimular os alunos a praticarem mais as competências clínicas e se sintam encorajados a desenvolver aptidões que ultrapassem o domínio do conhecimento apenas teórico. É evidente que, para conseguir avaliações adaptadas à prática clínica, é necessária uma estrutura de ensino prático adequada, com mais docentes nas áreas clínicas e menos alunos por cada turma prática, satisfazendo as exigências de um ensino de competências práticas essenciais ao futuro médico.

A Educação Médica nos Media

Inauguramos este novo espaço na presente edição, que pretende ser de divulgação de todo o tipo de referências noticiosas sobre educação médica nos principais meios de comunicação social, nacionais e internacionais, quer nos suportes tradicionais quer nas plataformas tecnológicas, como a Internet.

"Do You Have the "Right Stuff" to Be a Doctor?"

The New York Times

Link: <http://www.nytimes.com/2010/01/15/health/14chen.html>

Nesta primeira abordagem ao tema, destacamos um artigo publicado num dos mais renomados periódicos, a nível mundial: o The New York Times. Este jornal, na sua edição de 15 de Janeiro passado, incluía um artigo intitulado: "Do You Have the "Right Stuff" to Be a Doctor?", que aborda o tema da adequação dos critérios de admissão de novos alunos às escolas médicas, deixando no ar a interrogação sobre se as metodologias agora utilizadas são as mais apropriadas para seleccionar os melhores e mais competentes médicos do futuro.

A autora do artigo explora a possibilidade, agora revelada num estudo recente, publicado no The Journal of Applied Psychology, de que existe um tipo de teste que é melhor predictor do sucesso futuro dos alunos em Medicina do que os tradicionais testes de conhecimento, ou do que as médias escolares: os testes de personalidade.

Durante cerca de uma década, uma equipa de três psicólogos seguiu mais de 600 alunos de medicina na Bélgica. No início do estudo, os investigadores administraram um teste de personalidade estandardizado e avaliaram cada estudante em cinco dimensões diferentes da personalidade: extraversão, neuroticismo, abertura à experiência, amabilidade e conscienciosidade. Em seguida, acompanharam estes estudantes durante o seu percurso escolar, anotando as suas classificações, desempenho e taxas de desgaste. Descobriram então que os resultados nos testes de personalidade apresentavam uma fortíssima correlação com o desempenho dos alunos. As diferentes dimensões da personalidade revelaram ter correlações diferenciadas com o desempenho e sucesso dos estudantes, mas, em regra, todas se revelaram relevantes.

Um dos autores do estudo conclui que "o domínio da personalidade, do não-cognitivo, é uma área por abordar nas admissões de alunos às escolas médicas". O teste de personalidade que foi usado neste estudo é particularmente imune a diferenças culturais e de background, sendo esta maior fidedignidade uma potencial mais-valia para que este se torne um instrumento co-adjuvante dos testes tradicionais de conhecimento.

Mas estes testes de personalidade só serão efectivamente úteis quando as escolas médicas decidirem que tipo de profissional querem formar. Como nos avisa um dos autores: "Se uma escola médica só se importa com a formação de grandes investigadores, então eu dir-lhes-ia que não pesassem muito os resultados dos testes de personalidade. Mas se queremos médicos para exercerem prática clínica, então teremos que prestar maior atenção a estes resultados".

Tel 21 446 57 19
E-mail: webmaster@univadis.pt
www.univadis.pt

um serviço

medicina

univadis é uma marca registada da Merck & Co., Inc. Whitehouse Station, New Jersey, U.S.A.

Editorial

Alteração estatutária, novas escolas e velhos problemas

Confesso desconhecer os fundamentos e as intenções do legislador que subscreeu os novos estatutos da nossa Escola. Parece-me útil introduzir modificações estruturais quando baseadas na identificação de problemas essenciais, possibilitando a sua resolução com maior eficácia. Na verdade, e a título de exemplo, lembro que no domínio do ensino médico uma proposta de revisão curricular ficou esquecida nos dossiers dos anteriores Órgãos de Gestão; o Laboratório de Aptidões Clínicas, com plano de acção aprovado e iniciando os seus primeiros passos foi ignorado nos novos estatutos. E, entretanto, anunciavam-se novos grupos de ensino, sem que fosse sentida necessidade da sua reformulação, criando-se assim um clima de volatilização estrutural paralisante.

Surgem duas novas Escolas ditas Médicas, com modelo de ensino baseado em problemas, desinseridas dum contexto hospitalar e de tutores reconhecidamente qualificados, com o objectivo de formar médicos generalistas. Não nego algumas virtualidades do modelo e de até o aceitar como projecto de investigação, se não verificasse a escassez de recursos humanos e financeiros na nossa e nas restantes Escolas clássicas. É inaceitável assistir ao investimento em dois projectos, cada um para cerca de três dezenas de alunos, desnecessário e de resultados duvidosos,

Nesta Edição:

- **Editorial:** Alteração Estatutária, Novas Escolas e Velhos Problemas
- **Análise:** Testes Escritos de Resposta Objectiva na FMUC: considerações sobre os seus formatos, fidedignidade e qualidade das perguntas
- **Causa Nostra:** Curso TIPS em Reumatologia; novos Cursos TIPS Flex
- **Excelências Educare:** Inovação na Metodologia de

Avaliação das Aprendizagens: o caso da Propedêutica Médica

■ **Educação Médica nos Media:** artigo do jornal The New York Times

■ **Literatura em Educação Médica:** "Keeping learning central: a model for implementing emerging technologies"

■ **Essências Educare:** O Raciocínio Clínico e a Resolução de Casos-Problema Clínicos