

## EDUCAÇÃO MÉDICA PEDIÁTRICA

### Avaliação do Sistema de Avaliação de Alunos da Disciplina de Pediatria I – 1.ª Parte

LEONOR LEVY \*

\* Professora Auxiliar da Faculdade de Medicina de Lisboa

#### Resumo

Apresentam-se os objectivos da avaliação dos alunos da Disciplina de Pediatria I, os diferentes métodos de avaliação e as suas qualidades e limitações; são ainda explanados as razões do estudo e os seus objectivos. Apresenta-se ainda a cotação de cada método de avaliação, exame teórico, exame prático, avaliação contínua e trabalho de campo, assim como a determinação do Índice de Discriminação e Índice de Facilidade dos exames teóricos, a obtenção do grau de satisfação de docentes e discentes sobre o sistema de avaliação, a criação de uma base de dados e o tratamento estatístico.

**Palavras-Chave:** Pediatria I, Avaliação, Índice de Discriminação, Índice de Facilidade, Grau de satisfação de docentes e discentes, base de dados, estatística.

#### Summary

#### Evaluation of the Assessment System of Paediatrics I Students – Part I

The objectives of the assessment of students of Paediatrics I – Part I, the different assessment methods, their advantages and limitations are presented, as well as the reasons and objectives of the study. The numerical results of each assessment methods – theoretical examination, practical examination, continuous assessment and field work, as well as the Discrimination Index and the Facility Index of the theoretical examinations, satisfaction level of teachers and students concerning the assessment system, the construction of a data base and statistical analysis are also presented.

**Key-Words:** Paediatrics I, Assessment, Evaluation, Discrimination Index, Facility Index, Level of satisfaction, data base, statistics.

#### Introdução

O I Mestrado em Educação Médica realizado em Portugal pressupôs a frequência do Curso «Diploma on Medical Education», ministrado por docentes da Universidade de Cardiff. A obtenção do «Diploma on Medical Education», dependeu, para além da frequência do curso, da aprovação em diferentes provas efectuadas ao longo do curso. Para a obtenção do título de *Mestre em Educação Médica* em Portugal foi necessária a elaboração de uma Tese de Mestrado em Educação Médica, avaliada através de uma dissertação e discussão da Tese por um Júri constituído por três Professores.

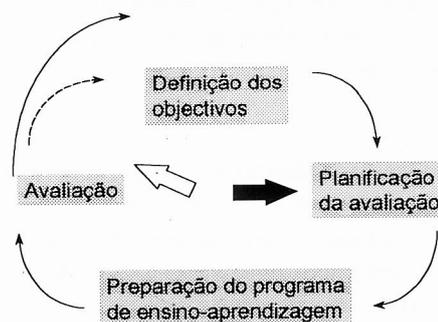
O tema escolhido para o estudo conducente a essa mesma Tese, foi a «Avaliação do sistema de avaliação de alunos da Disciplina de Pediatria I».

O estudo foi efectuado durante o ano lectivo de 1998/1999. Esta Tese de Mestrado contém seis partes.

A primeira parte do estudo consiste na introdução, razões, objectivos, e material e métodos do estudo.

O planeamento dos processos de ensino-aprendizagem e da sua avaliação estão patentes na «Espiral da Educação» (Gráfico 1). Com efeito, após a definição das tarefas e dos objectivos educacionais, deve ser planificado um sistema de avaliação, cujo maior contributo será o de fornecer as bases de um juízo de valor que permitirá tomar as melhores decisões pedagógicas. Este sistema de avaliação poderá incluir a avaliação dos discentes, a avaliação dos docentes e ainda, a avaliação de todo o programa de ensino-aprendizagem.

GRÁFICO 1  
Espiral da Educação



Correspondência: Leonor Levy  
Faculdade de Medicina de Lisboa  
Email: [leonor@mail.telepac.pt](mailto:leonor@mail.telepac.pt)

Aceite para publicação em 20/08/2001.  
Entregue para publicação em 20/08/2001.

A modificação de um programa ou das técnicas de ensino, sem mudar o sistema de avaliação, tem todas as hipóteses de não modificar seja o que for. Modificar o sistema de avaliação, mesmo sem mudar o programa de ensino-aprendizagem, terá um maior impacto sobre a natureza e a qualidade da aprendizagem, do que mudar o programa sem mudar o sistema de avaliação <sup>(1)</sup>.

A educação é um processo cujo objectivo essencial é o de facilitar as modificações do comportamento. Os tipos de modificações do comportamento que uma instituição se esforça por obter, constituem os objectivos educacionais.

Assim, a avaliação consiste em determinar em que medida cada um dos objectivos foi atingido, as qualidades das técnicas de ensino-aprendizagem e dos docentes <sup>(2, 3, 4, 1)</sup>.

A avaliação dos alunos deveria ser um processo sistemático de recolha, análise e interpretação de informações, a fim de determinar o sucesso dos alunos ou segundo Sutton <sup>(1)</sup>, «a avaliação é o processo pelo qual tentamos responder à pergunta: *Quanto é que os alunos aprenderam?*».

Para Sutton <sup>(1)</sup> a palavra *assessment* refere-se à avaliação dos desempenhos dos alunos, enquanto que a palavra *evaluation* se refere à efectividade do processo de ensino-aprendizagem; *evaluation* seria, assim, o processo pelo qual nós tentamos responder à pergunta: «*Qual é a qualidade do processo de ensino-aprendizagem?*».

Os objectivos da avaliação dos alunos (*assessment*) são inúmeros:

- Diagnosticar as necessidades de aprendizagem
- Providenciar *feed-back* aos formandos
- Ordenar os estudantes
- Tomar decisões acerca da competência dos formandos
- Assegurar o padrão mínimo de desempenhos
- Motivação dos alunos
- Selecção dos alunos
- Promoção do desenvolvimento pessoal
- Dar informação ao mercado de emprego
- Proteger a sociedade
- Ganhar prémios
- Aumentar o prestígio da instituição
- Avaliar a efectividade de um programa de ensino-aprendizagem
- Etc.....

Os métodos de avaliação devem adequar-se aos objectivos em causa, sejam eles dos domínios dos conhecimentos, das atitudes e dos gestos. Recorde-se que os autores anglo-saxónicos referem-se a estes domínios como o *Triple H-Head, Heart, Hand*.

Para cada um destes domínios existem diferentes graus de complexidade.

Para o domínio dos conhecimentos a *recordação* corresponderá a um nível menor, a *interpretação* a um nível médio e a *resolução* de problemas a um nível de maior complexidade.

Para o domínio das atitudes, a *receptividade* corresponderá a um nível baixo, a *resposta* a um nível médio e a *interiorização* a um nível mais elevado.

No domínio dos gestos, a *imitação* corresponderá a um nível mais baixo, o *controlo* a um nível intermédio e o *automatismo* a um nível mais alto.

Segundo os quatro níveis de aprendizagem da pirâmide de Miller <sup>(1)</sup>, a avaliação dos alunos poderá efectuar-se em quatro níveis; no primeiro nível, avalia-se se o aluno **sabe**, no segundo nível, avalia-se se o aluno **sabe como**, no terceiro nível, avalia-se se o aluno **mostra como** e no quarto e mais elevado nível, se o aluno **faz**.

Também as formas de avaliação diferem de acordo com os seus objectivos.

A avaliação formativa, também chamada de controlo ou de diagnóstico, tem como objectivo elucidar o aluno sobre o caminho que lhe falta percorrer para atingir os objectivos educacionais.

A avaliação formativa consiste na avaliação dos progressos do aluno, desde o início até ao fim do programa de ensino-aprendizagem, permite ajustar as actividades pedagógicas em função do progresso do aluno e não pode ser utilizada pelo docente para sancionar o aluno <sup>(4, 5)</sup>.

A avaliação de certificação, também chamada de sumativa, de sanção ou cumulativa é efectuada habitualmente no fim de um período de ensino-aprendizagem e tem como objectivo a protecção da sociedade impedindo os incompetentes do exercício da profissão e serve tradicionalmente para classificar e ordenar os estudantes, decidir da progressão de um curso, para a obtenção de um diploma e para a progressão numa carreira.

Existem numerosos métodos de avaliação, como provas escritas para avaliação dos conhecimentos, sob a forma de perguntas de escolha múltipla, perguntas de resposta curta, perguntas de resposta curta modificada ou «*modified essay question*», perguntas de desenvolvimento, interpretação e resolução de problemas, provas orais para avaliação dos conhecimentos e provas práticas para avaliação de conhecimentos atitudes e gestos, em situações reais ou em simulações.

Todos os métodos de avaliação têm qualidades e defeitos.

Os testes podem conter tipos diferentes de perguntas, como perguntas de resposta múltipla, de resposta curta, de resposta curta estruturada ou modificada, interpretação de casos clínicos e ainda perguntas de desenvolvimento.

As perguntas de resposta múltipla são difíceis de elaborar, fáceis de corrigir, têm boa fiabilidade mas uma menor validade. Encorajam a adivinha, são consumidoras de tempo na elaboração, abrangem muitas áreas e recompensam o conhecimento, habitualmente o nível mais baixo do conhecimento <sup>(6, 7)</sup>.

Os testes de perguntas objectivas podem ter diferentes formas. Alguns tipos de testes objectivos apresentam os *items* sob a forma de «verdadeiro ou falso». Estes tipos de teste são difíceis de construir de maneira a que a resposta seja totalmente verdadeira ou totalmente falsa; por outro lado, um aluno totalmente ignorante tem 50% de hipóteses de acertar na resposta certa, embora existam sugestões de que modificando estes testes, estes poderiam ser ideais para avaliar a aquisição de conceitos <sup>(8, 1)</sup>.

O género mais familiar das perguntas de resposta múltipla, pode ser encarado como uma extensão do género «verdadeiro ou falso» e consistem numa afirmação chamada o «caule», seguida de três, quatro ou cinco respostas opcionais. A resposta correcta chama-se «chave» e as opções incorrectas chamam-se «distractores».

Uma crítica que frequentemente se faz às perguntas de resposta múltipla, é a que os alunos são encorajados a «adivinharem» a resposta certa. No entanto, existe a evidência que os alunos apenas são encorajados a adivinharem a resposta certa em duas situações: quando são totalmente ignorantes na matéria e quando se sentem com falta de tempo para acabarem o exame <sup>(1)</sup>.

Outra crítica que se faz é a que as perguntas de resposta múltipla apenas contemplam o domínio do conhecimento e ao mais baixo nível. Não é necessário que assim seja, embora se reconheça também a dificuldade de elaborar perguntas para testar o conhecimento em níveis mais elevados, como, por exemplo, a interpretação e a resolução de problemas <sup>(9)</sup>.

As perguntas de resposta curta são mais fáceis de elaborar, mas de menor fiabilidade na correcção. Abrangem menos áreas, recompensam o raciocínio, encorajam o bluff, mas por outro lado não existem sinais que ajudem a adivinhar resposta correcta <sup>(7, 10)</sup>.

A interpretação de casos clínicos pode ser efectuada através de perguntas de resposta curta modificada ou «modified essay question» (MEQ). Este tipo de avaliação permite a avaliação no domínio do conhecimento nos seus níveis mais elevados, ou seja a interpretação e resolução de problemas e parece que melhores desempenhos no MEQ estão associados a melhores desempenhos na vida clínica, embora o MEQ possa ser acusado de dificuldades de avaliação <sup>(11, 12, 13, 14, 15)</sup>.

As perguntas de desenvolvimento são de difícil correcção e apesar de poderem ter uma boa vitalidade, a sua fiabilidade é baixa, a não ser que se estructure a sua correcção <sup>(13, 16)</sup>.

A observação directa do desempenho do aluno em situação real ou em simulação parece ser a maneira mais válida de avaliação do desempenho dos alunos. Infelizmente a fiabilidade deste método parece ser muito baixa, especialmente em áreas complexas como a da comunicação interpessoal em que não existem outras formas de avaliação alternativas.

No entanto, temos que continuar a avaliar o desempenho dos alunos em áreas como a colheita de uma história clínica, a relação médico-doente ou ainda nas competências de aconselhamento o que sugere a importância de melhorar o sistema de avaliação quando recorremos à observação directa <sup>(13, 16)</sup>.

Os exames orais ou *viva* foram durante anos o método predominante ou, por vezes, o único método de fazer a avaliação dos alunos; os exames orais tradicionais têm sido postos em causa, por terem uma baixa fiabilidade, por poderem ser substituídos vantajosamente por perguntas de resposta múltipla, podendo, no entanto, serem apropriados quando se trata de fazer a discriminação entre os alunos mais bem classificados, através de perguntas tendentes a avaliar o conhecimento a níveis mais elevados <sup>(13, 16)</sup>.

Ao fazermos a escolha dos métodos de avaliação precisamos de ter em conta os objectivos da avaliação em termos de conteúdo nos domínios dos conhecimentos, atitudes e gestos e ainda o nível da avaliação propriamente dita.

Um bom sistema de avaliação deverá ainda ter validade, fiabilidade, ser reprodutível e exequível.

A avaliação de um programa de ensino-aprendizagem só tem sentido se tiver um objectivo. Os processos ligados a uma avaliação deverão ser encarados como parte do desenvolvimento profissional dos professores, a fim de adquirirem as competências que possam conduzir a progressos em educação <sup>(1)</sup>.

Ao planear uma avaliação deverá ser importante considerar:

- A razão pela qual a avaliação está a acontecer
- Por quem é que a avaliação é exigida
- O que vai ser avaliado
- Quando é que a avaliação terá lugar
- Como é que a avaliação vai ser efectuada
- Que informações vão ser recolhidas e analisadas
- Quais as perguntas que vão ser feitas
- Como é que as perguntas vão ser formuladas
- Quem levará a cabo a avaliação

Esta avaliação poderá ser interna, ser for levada a cabo por um membro da equipa responsável pela produção do programa de ensino-aprendizagem.

Uma avaliação será externa, se for efectuada por alguém não envolvido no processo de ensino-aprendizagem.

Uma avaliação será formativa, se o processo de avaliação decorrer a tempo de fazer eventuais mudanças; pelo contrário, uma avaliação será sumativa, quando o processo de avaliação decorre no fim de um processo de ensino-aprendizagem. Este tipo de avaliação permite o julgamento de como decorreu o sistema de ensino-aprendizagem e de como o processo de ensino-aprendizagem foi aceite por docentes e discentes.

O *timing* de uma avaliação pode também sofrer do *efeito hawthorne*, ou seja o efeito da novidade de uma maneira diferente de fazer as coisas.

A avaliação de um processo de ensino-aprendizagem pode ser efectuada de uma maneira formal ou informal e os resultados serão tanto mais úteis se as informações forem recolhidas sistematicamente.

As fontes de informação para esta avaliação incluem:

- A opinião dos docentes
- Os resultados dos testes e das outras formas de avaliação
- Observação pessoal
- A opinião dos alunos

As opiniões de docentes e discentes podem ser recolhidas quer verbalmente através de entrevistas ou quer através do preenchimento de questionários.

Para Sutton <sup>(1)</sup>, as entrevistas têm alguns inconvenientes, como o grande consumo de tempo, a ineficácia em termos de resultados fiáveis e a tendência dos entrevistados responderem de uma maneira «politicamente correcta».

Existem diferentes tipos de questionário, mas essencialmente os questionários devem ter algumas características comuns:

- Simples e rápidos de preencher
- Fáceis de analisar e interpretar
- Válidos e fiáveis

As palavras que integram um questionário são de importância vital; os questionários não devem induzir a resposta ou serem emocionalmente enviesados ou ainda embaraçar ou levar a sentimentos de inferioridade ou desadequação nas pessoas que a eles respondem.

### Razões do estudo

A Disciplina de Pediatria I tem como Professor regente o Professor Doutor João Gomes-Pedro. O número dos seus docentes tem variado ao longo dos anos; no ano lectivo de 1998/1999, durante o primeiro semestre,

havia onze docentes a que correspondiam onze turmas para as aulas práticas; no segundo semestre do ano lectivo, os assistentes livres deixaram de colaborar no ensino prático da Disciplina de Pediatria I e o número de docentes ficou reduzido a seis.

Ao iniciar o Programa de Ensino-Aprendizagem da Disciplina de Pediatria I, os alunos tomam conhecimento dos seus conteúdos, objectivos gerais, métodos de ensino-aprendizagem e métodos de avaliação.

Os objectivos gerais da Disciplina de Pediatria I são:

- Adquirir conhecimentos teóricos e práticos conducentes à promoção da Saúde, Bem-estar e prevenção das doenças nas crianças e adolescentes.
- Adquirir atitudes e capacidades de comunicação eficazes na abordagem de cada criança e adolescente, quando integrados em diferentes ambientes.
- Aprender os conceitos de diferenças individuais, de resiliência e da dinâmica do desenvolvimento bio-psico-social inerente às crianças e adolescentes.
- Aprender a Semiologia ao longo do ciclo de vida até ao fim da adolescência, nos seus diferentes parâmetros.
- Integrar os conceitos anteriores numa perspectiva conjunta de Educação e Saúde da criança e sua família, no contexto da sua circunstância e dos seus Direitos.

Os objectivos específicos da Disciplina de Pediatria I estão contidos numa lista de conhecimentos, atitudes e desempenhos, codificados de acordo com o grau de exigência pretendida, segundo a metodologia do «Blueprint» <sup>(17)</sup>.

A avaliação dos alunos da Disciplina de Pediatria I, tem sido apenas de certificação ou sumativa. Esta avaliação compreende um exame final com um teste escrito e uma prova prática, contando ainda com uma avaliação contínua do aluno nas aulas práticas e a avaliação de um relatório escrito sobre o trabalho de campo proposto aos alunos.

O teste escrito é composto por perguntas de resposta múltipla e perguntas de resposta curta. Os resultados dos testes têm revelado uma curva de Gauss desviada para a direita, sugerindo a facilidade das perguntas.

Não tem existido uma estruturação da avaliação das provas práticas tendente a melhorar a fiabilidade deste tipo de provas, assim como não tem existido uma estruturação da avaliação contínua dos alunos nas aulas práticas ou da avaliação do relatório sobre o trabalho de campo.

A presente situação tem implicações várias:

Os estudantes não se sentem satisfeitos com o sistema de avaliação, não incorporam os seus novos conhecimentos na prática e tendem a esquecer aquilo que aprenderam.

O sistema de avaliação não discrimina entre os bons e os maus estudantes e tem pouca fiabilidade.

Os docentes não estão satisfeitos com a presente situação, investem pouco tempo na educação e na sua actualização científica.

O prestígio da Faculdade de Medicina de Lisboa sofre com esta problemática.

Pensamos que o sistema de avaliação de alunos da Disciplina de Pediatria I não satisfaz nem os estudantes nem os docentes.

Um melhor sistema de avaliação implicaria uma maior satisfação de docentes e discentes, uma maior ligação entre conhecimentos teóricos e a sua aplicação na prática, traduzindo-se em melhores desempenhos nas situações reais; um melhor sistema de avaliação aliado às necessárias modificações do programa de ensino-aprendizagem, poderia aumentar a motivação dos alunos, condição essencial para um melhor desenvolvimento das capacidades de auto-aprendizagem e auto-aperfeiçoamento ao longo da sua vida activa.

Pensamos ainda que um melhor sistema de avaliação dos alunos se repercutiria positivamente na satisfação dos docentes, levando a um maior investimento na educação e na actualização científica a ela ligada, aumentando o prestígio da Faculdade de Medicina de Lisboa.

Dadas as dificuldades existentes, nomeadamente a escassez de docentes e as suas responsabilidades assistenciais, o *ratio* docente/discente existente nas aulas práticas, sem esquecer as limitações logísticas, não nos foi ainda possível alterar o actual sistema de avaliação dos alunos na Disciplina de Pediatria I.

Pensamos, no entanto, que existe uma condição subjacente a qualquer tipo de mudança, que é a avaliação do que actualmente se pratica em termos de sistema de avaliação.

É assim necessário fazer uma avaliação do sistema de avaliação na Disciplina de Pediatria I, a fim de fazer o diagnóstico da situação e, se for o caso, proceder às modificações necessárias.

## Objectivos

O objectivo deste estudo foi a *Avaliação do Sistema de Avaliação dos Alunos da Disciplina de Pediatria I*. Este objectivo pressupõe a:

- Avaliação dos diferentes sistemas de avaliação dos alunos.
- Avaliação do grau de satisfação dos docentes da Disciplina de Pediatria I, quanto ao sistema de avaliação.

- Avaliação do grau de satisfação dos alunos da Disciplina de Pediatria I, quanto ao sistema de avaliação.
- Comparação dos resultados obtidos nas diferentes provas de avaliação nos dois semestres da Disciplina de Pediatria I.
- Comparação do grau de satisfação dos docentes nos dois semestres da Disciplina de Pediatria I.
- Comparação do grau de satisfação dos alunos nos dois semestres da Disciplina de Pediatria I.

## Material e Métodos

Este estudo efectuou-se no ano lectivo de 1998/1999, nos dois semestres da Disciplina de Pediatria I; no primeiro semestre, houve apenas uma chamada e no segundo semestre, duas chamadas.

Assim, optou-se por avaliar separadamente cada método de avaliação dos alunos, ou seja, três exames teóricos, três exames práticos, três avaliações contínuas e três trabalhos de campo; optou-se ainda por avaliar os graus de satisfação de docentes e alunos apenas distribuídos por dois semestres.

Tratou-se de um estudo prospectivo, exploratório, não experimental e sem ocultação.

Para efeitos de avaliação do sistema de avaliação da Disciplina de Pediatria I, foram objecto de estudo todos os alunos que se apresentaram ao exame teórico e prático da Disciplina de Pediatria I, que efectuaram o trabalho de campo proposto e respectivo relatório.

O número de alunos avaliados foi de 70 no primeiro semestre, e no segundo semestre foi respectivamente de 35, para a primeira chamada e 41, para a segunda chamada.

Para efeitos de avaliação do grau de satisfação dos docentes da Disciplina de Pediatria I, quanto ao sistema de avaliação, foram objecto de estudo todos os docentes que preencheram e entregaram o respectivo questionário anónimo.

O número de docentes da Disciplina de Pediatria I variou de acordo com o semestre em estudo; assim, no primeiro semestre, os docentes eram em número de onze, enquanto que no segundo semestre, o número de docentes ficou reduzido a apenas seis.

Para efeitos de avaliação do grau de satisfação dos alunos da Disciplina de Pediatria I, quanto ao sistema de avaliação, foram objecto de estudo todos os alunos que preencheram e entregaram o respectivo questionário anónimo.

Ao longo do 1.º semestre do ano lectivo de 1998/1999, os docentes da Disciplina de Pediatria I reuniram-se mensalmente a fim de preparar o sistema de avaliação

dos alunos do 1.º semestre. Houve apenas uma época de exames.

No 2.º semestre, os alunos exigiram duas épocas de exames; houve assim uma primeira e uma segunda chamadas.

Ao longo do 2.º semestre do ano lectivo de 1998/1999, os docentes da Disciplina de Pediatria I reuniram-se mensalmente a fim de preparar o sistema de avaliação dos alunos do 1.º semestre.

A metodologia seguida em ambos os semestres foi semelhante, diferindo apenas entre si, os conteúdos dos testes.

Todos os docentes contribuíam com sugestões para a elaboração dos testes teóricos finais e para a estruturação das outras formas de avaliação, ou seja, os exames práticos, a avaliação contínua e a avaliação do relatório do trabalho de campo.

A avaliação dos alunos foi efectuada no final de cada semestre através de um exame teórico, de uma prova prática, avaliação de um relatório escrito sobre o trabalho de campo efectuado ao longo do semestre e avaliação contínua do aluno pelo seu assistente.

A pontuação final (máximo de 20 valores), constou da soma da nota do teste (máximo de 10 valores), da prova prática (máximo de 5 valores), do relatório sobre o trabalho de campo (máximo de 2,5 valores) e da avaliação contínua (máximo de 2,5 valores).

Definiu-se o Critério Mínimo de aprovação, quer para o teste, quer para o somatório de todos os tipos de avaliação (teste, prova prática, avaliação contínua e avaliação do trabalho de campo).

Como Critério Mínimo de aprovação do teste foi definida a nota de 5 valores, correspondendo a 50% da nota máxima possível; como Critério Mínimo de aprovação na Disciplina de Pediatria I, foi definida a nota de 10 valores, correspondendo a 50% da nota máxima possível.

#### Exame teórico:

O exame teórico da Disciplina de Pediatria I do 1.º semestre, consistiu num teste escrito com 44 perguntas de escolha múltipla e 8 de resposta curta; a maior parte das perguntas de resposta múltipla integravam testes da mesma disciplina efectuados em anos anteriores, tendo sido apenas uma minoria de perguntas (20%) elaboradas para o efeito.

O exame teórico da Disciplina de Pediatria I do 2.º semestre – 1.ª chamada, consistiu num teste escrito com 25 perguntas de escolha múltipla, duas perguntas de interpretação de casos clínicos e 20 de resposta curta; a maior parte das perguntas de resposta múltipla integravam testes da mesma disciplina efectuados em anos anteriores, tendo

sido apenas uma minoria de perguntas (20%) elaboradas para o efeito.

O exame teórico da Disciplina de Pediatria I do 2.º semestre – 2.ª chamada, consistiu num teste escrito com 26 perguntas de escolha múltipla, duas perguntas de interpretação de casos clínicos e 19 de resposta curta; a maior parte das perguntas de resposta múltipla integravam testes da mesma disciplina efectuados em anos anteriores, tendo sido apenas uma minoria de perguntas (20%) elaboradas para o efeito.

Todos os docentes da Disciplina de Pediatria I contribuíam para a preparação do teste.

Os testes tiveram lugar em dois anfiteatros, o anfiteatro de Fisiologia e o de Farmacologia da Faculdade de Medicina de Lisboa e tiveram a duração de duas horas. Cotação dos testes:

Os docentes da Disciplina de Pediatria elaboraram a cotação de cada uma das perguntas do teste e depois de somados os resultados de todas as perguntas e através de uma regra de três simples, comparou-se o resultado final obtido com a máxima nota possível de dez valores.

Foi determinado o Índice de Facilidade para cada pergunta de resposta múltipla dos testes.

O Índice de Facilidade (F) define a proporção de candidatos testados que obtiveram a resposta certa e obtêm-se dividindo o número de candidatos que obtiveram a resposta certa pelo número de candidatos que fizeram o teste.

$$F = \frac{R}{N} \text{ (candidatos que têm a resposta certa)} \\ \text{N (candidatos que fazem o teste)}$$

Foi determinado o Índice de Discriminação para cada pergunta de resposta múltipla dos testes.

O Índice de Discriminação (D) dá a medida de como cada item de um teste é capaz de discriminar entre os candidatos mais capazes e os menos capazes.

O Índice de Discriminação obtêm-se através da fórmula:

$$D = \frac{A - B}{n} = \frac{A - B}{0.25N}$$

Para obter o Índice de Discriminação foi necessário:

- Ordenar os candidatos, de acordo com os resultados totais obtidos no teste, por ordem decrescente.
- Identificar o «top quarter», ou seja, os alunos cujos resultados estão acima do percentil 75 e o «bottom quarter», ou seja, os alunos cujos resultados estão abaixo do percentil 25.
- Para cada item, determinar quantos candidatos acima do percentil 75 (top quarter) tiveram a res-

posta certa e quantos candidatos abaixo do percentil 25 (bottom quarter) tiveram a resposta certa.

- O Índice de Discriminação é dado pela equação e varia entre +1 e -1. Um poder discriminativo menor que 0,2, faz com que as perguntas sejam retiradas e entre 0,2 e 0,3, sugere que as perguntas sejam revistas.

A – número no «top quarter» que tiveram a resposta certa

B – número do «bottom quarter» que tiveram a resposta certa

n – número que representa um quarto do número total de candidatos

N – número total de candidatos

Exame prático:

O exame prático dos alunos efectuou-se nos dias imediatos ao do teste, numa Unidade de Internamento, na Urgência ou no piso 5 do Serviço de Obstetrícia e teve a duração aproximada de 30 minutos.

A prova prática constou da colheita da história clínica e a observação de uma criança, pelo aluno. A avaliação do desempenho dos alunos foi feita através do preenchimento e cotação de uma grelha de observação em termos de avaliação nos domínios dos conhecimentos, atitudes e gestos.

O júri da prova prática constou de dois docentes da Disciplina de Pediatria I, um dos quais era o docente que acompanhou o aluno nas aulas práticas.

Regras do preenchimento da grelha de observação do exame prático e da sua cotação:

- Os alunos foram classificados por 2 docentes da Disciplina de Pediatria I.
- Os alunos deveriam colher 2 itens da história, ter 2 cuidados antecipatórios e fazer 2 manobras no exame objectivo da criança de qualquer grupo etário.
  - Todas as 6 atitudes foram cotadas da seguinte maneira: Não faz = 0, Faz mal = 1, Faz = 2, Faz bem = 3...
- Os itens da colheita de história, os 2 cuidados antecipatórios e as 2 manobras no exame objectivo da criança também foram cotados da mesma maneira.
- Para obter a cotação do exame prático, fez-se a média entre os resultados dos 2 assistentes para cada subgrupo (atitudes, história, cuidados antecipatórios e exame objectivo da criança) e somaram-se; depois, aplicou-se a regra de 3, comparando o resultado obtido com o máximo possível de 5 valores.

A avaliação contínua dos alunos foi efectuada durante as aulas práticas ao longo do semestre de aulas e constou da apreciação do aluno pelo respectivo docente quanto à assiduidade, pontualidade, interesse e intervenção nas aulas práticas.

A avaliação de um relatório escrito sobre o trabalho de campo efectuado ao longo do semestre pelo aluno constou da apreciação do relatório escrito em termos da legibilidade, apresentação, conteúdo, capacidades criativas e críticas do aluno.

#### **Avaliação do grau de satisfação dos docentes da Disciplina de Pediatria I, quanto ao sistema de avaliação:**

A avaliação do grau de satisfação dos docentes da Disciplina de Pediatria I, quanto ao sistema de avaliação, foi efectuada através de um questionário segundo uma escala de Likert de 1 a 5, que foi distribuído a todos os docentes. O questionário foi preenchido anonimamente pelos docentes e entregue depois de finda a avaliação de todos os alunos.

#### **Avaliação do grau de satisfação dos alunos da Disciplina de Pediatria I, quanto ao sistema de avaliação:**

A avaliação do grau de satisfação dos alunos da Disciplina de Pediatria I, quanto ao sistema de avaliação foi efectuada através de um questionário segundo uma escala de Likert de 1 a 5, que foi distribuído a todos os alunos que se apresentaram no exame teórico da Disciplina de Pediatria I. O questionário foi preenchido anonimamente pelos alunos e entregue depois do exame prático.

Os dados obtidos foram introduzidos numa base de dados em SPSS 8 para Windows.

O estudo estatístico foi efectuada através do SPSS 8 para Windows.

O estudo estatístico iniciou-se com a análise univariada e a apresentação das variáveis nominais, ordinais e intervaladas em termos de estatística descritiva e representações gráficas.

As variáveis nominais foram estudadas em termos de frequência simples e moda.

As variáveis ordinais foram estudadas em termos de frequência simples e moda.

As variáveis de tipo intervalado foram estudadas em termos de estatística descritiva e foram ainda apresentadas em histograma com curva normal sobreposta, em diagrama de caule e folhas e caixa de bigodes.

As associações de variáveis foram estudadas através de correlações bivariadas, ou através do teste t quando se tratava de curvas normais e através dos testes não paramétricos de Mann-Whitney e de Kruskal-Wallis quando não se verificava os pressupostos para a aplicação do teste t e da ANOVA. Em todas as determinações foi escolhido como limiar de significância, o valor de p de 0,05.

#### Bibliografia

1. Sutton RA. An Introduction to Assessment & Evaluation Processes and Procedures. University College Cardiff.
2. Curzon LB. Teaching in further education. An outline of principles and practice, 4th edn. London: Cassell, 1994.
3. Eaton DGM, Levene MI. Student feedback: influencing the quality of teaching in a paediatric module. *Medical Education* 1997; 31: 190-3.
4. Guilbert JJ. Guide Pédagogique pour les personnels de Santé. Genève, OMS, 1981.
5. Rolfé I, McPherson J. Formative assessment: how am I doing? *Lancet* 1995; 345: 837-9.
6. Ferland JJ, Dorval J, Levasseur L. Measuring higher cognitive levels by multiple choice questions: a myth? *Medical Education* 1987; 21: 109-13.
7. Harris D. Assessment of students. *Br J Hospital Medicine* 1992; 9: 586-9.
8. Richardson R. The multiple choice true-false question: what does it measure and what could it measure? *Medical Teacher* 1992; 2/3: 201-4.
9. Van der Vleuten CPM, Newble DI. How can we test clinical reasoning? *Lancet* 1995; 345: 1032-4.
10. Wakeford RE, Roberts S. Short answer questions in an undergraduate qualifying examination: a study of examiner variability. *Medical Education* 1984; 18: 168-73.
11. Feletti GI, Smith EKM. Modified Essay Questions: are they worth the effort? *Medical Education* 1986; 20: 126-32.
12. Knox JDE. What is... a Modified Essay Question? *Medical Teacher* 1989; 1: 51-7.
13. Newble D, Cannon R. A Handbook for medical teachers. United Kingdom, Kluwer Academic Publishers 1994.
14. Rabinowitz HK. The modified essay question: effect of author location on student performance. *Medical Education* 1986; 20: 318-20.
15. Rabinowitz HK. The modified essay question: an evaluation of its use in a family medicine clerkship. *Medical Education* 1987; 21: 114-8.
16. Newble DI. Assessing clinical competence at the undergraduate level. *Medical Education* 1992; 26: 504-11.
17. Metz JCM, Stoelinga GBA. Blueprint 1994: training doctors in The Netherlands. Objectives of undergraduate medical education. University of Nijmegen (Ed) 1994.