

Capítulo 1

Introdução

“The mind that opens to a new idea never returns to its original size”

Albert Einstein (1879 - 1955)

1.1 Antecedentes e motivação

A educação é um processo dinâmico e contínuo que acompanha o indivíduo em todas as facetas da vida. Por ser um processo de extrema importância, a sociedade, desde a camada mais baixa, e os governos, ao mais alto nível, apostam na educação como uma via através da qual o cidadão possa garantir a sua integração com os demais, evoluir e fazer aplicações dos seus conhecimentos em vários domínios do saber. Na educação escolar, o processo educativo é regido por normas, objetivos concretos e padrões com a finalidade de se alcançar o conhecimento científico. Alguns dos instrumentos habitualmente utilizados, para verificar se o conhecimento científico é alcançado e até que nível ele é alcançado, são as provas avaliativas ou exames (doravante, simplesmente, designados por testes). Concretamente, no processo de ensino e no processo de aprendizagem, os testes são utilizados para verificar se os objetivos programáticos para um determinado conteúdo curricular, unidade temática, capítulo, programa ou ciclo acadêmico foram, efetivamente, alcançados. Espera-se que os resultados obtidos em tais testes sejam o reflexo do desempenho dos examinandos, os quais poderão ditar o seu futuro em termos de distinção e progressão. Se no processo do ensino se tem, fundamentalmente, os educadores (professores ou entidades) que elaboram e aplicam os testes em função dos objetivos que pretendem alcançar então, no processo de aprendizagem, têm-se os examinandos (alunos) que devem possuir determinadas capacidades (habilidades)

para responderem positivamente aos testes.

Ao nível da Universidade de Aveiro, e especificamente no Departamento de Matemática (DMat), local onde foi desenvolvido este trabalho, as avaliações acompanham o processo de ensino e o processo de aprendizagem. A unidade curricular de Bioestatística, da responsabilidade do DMat para os alunos do 1^o ano do curso de Biologia e do 2^o do curso de Biologia e Geologia, foi escolhida para estudo nesta dissertação pois, pela primeira vez, no ano lectivo 2015/2016 mereceu destaque face ao interesse dos professores em reformular os modelos de avaliação. As avaliações foram parceladas em dois modelos: testes escritos, com 75% da nota final, e um trabalho em grupo, com os restantes 25%. Os testes consistiam de questões de escolha múltiplas com apenas uma opção correta cada, permitindo que as respostas fossem dicotomizadas e, conseqüentemente, corrigidas no critério de certo ou errado. Anteriormente, embora algumas avaliações incluíssem questões de escolha múltiplas seleccionadas a partir do moodle, o peso destas avaliações sob a nota final era de apenas 5%. O objetivo, a médio prazo, será de criar um banco de dados com perguntas parametrizadas para diferentes conteúdos e, eventualmente no futuro, com a implementação do mesmo modelo noutras unidades curriculares, comparar o desempenho dos alunos de diversos cursos ao longo de anos. Com base na mudança de paradigma nos critérios de avaliação e no propósito dos professores desta unidade curricular, vimos uma oportunidade de desenvolver um estudo que contribuísse com a efetivação de tal propósito.

1.2 Objetivo

Outrora as avaliações tinham como finalidade a obtenção de resultados classificatórios. Hoje, com os investimentos feitos no processo educativo, interesses há não só em fazer estudos comparativos dos examinandos em diferentes épocas ou dos examinandos de diferentes grupos mas também em saber o que está por detrás desses resultados (por exemplo, as habilidades cognitivas dos examinandos e a qualidade dos instrumentos de avaliação). Para lograr esses intentos há necessidade de avaliar as avaliações, isto é, verificar o nível de qualidade da própria avaliação com base em critérios adequados. Este processo é fundamental pois

permite verificar se os resultados gerados pela avaliação são úteis e contribui para obtenção de informações atinentes às limitações e potencialidades da avaliação. A avaliação da avaliação é conhecida na literatura como meta-avaliação (Scriven, 1969). Etimologicamente meta significa “depois de”, contudo, a meta-avaliação pode ser formativa (realizada continuamente visando a melhoria do processo avaliativo) quanto sumativa (realizada no fim do processo avaliativo examinando cuidadosamente os resultados com a finalidade de analisar o seu valor e detetar pontos fracos e fortes das avaliações anteriores). Para Severo & Tavares (2010), o controle da qualidade dos teste escritos são de extrema importância e a meta-avaliação permite determinar se os objetivos definidos para uma dada disciplina estão sendo alcançados e avaliados através do teste proposto. Independentemente de ser ou não formativa, a meta-avaliação deve se orientar por alguns critérios consensuais entre vários autores: a validade (interna e externa), confiabilidade, objetividade, credibilidade, utilidade ou importância, custo-benefício, relevância, abrangência, oportunidade e dissiminação (Scriven, 1991; Davidson, 2005; Stufflebeam, 2011).

A Teoria Clássica de Testes (TCT) e a Teoria de Resposta ao Item (TRI), abordadas nesta dissertação, possibilitam a partir dos resultados obtidos nas avaliações, respetivamente, a obtenção de estatísticas que permitem avaliar a qualidade de um teste e a definição de modelos para medir as habilidades dos examinandos e estimar os níveis de dificuldades, de discriminação e de acerto casual presentes nas perguntas. Os modelos da TRI estabelecem relação entre a probabilidade de um indivíduo responder corretamente a uma pergunta e as suas habilidades. Em particular, em contexto educativo, os modelos da TRI são utilizados para modelar a probabilidade de um aluno responder corretamente a uma questão de um teste em termos da capacidade do indivíduo e das características da questão.

Nesta dissertação pretendemos fazer uma meta-avaliação das questões parametrizadas de escolha múltiplas contidas nos testes de Bioestatística do último ano letivo 2015/2016. Mais ainda, com vista a experimentar as dificuldades inerentes à construção de potenciais questões com enunciados não estáticos (variáveis cada vez que se recorre a ela), na presente dissertação contribuimos ainda com um conjunto de questões parametrizadas que podem ser incorporadas num banco de questões de Estatística.

Fruto deste trabalho de meta-avaliação e parametrização de exercícios, outras questões, como as que se seguem, foram despertando o nosso interesse: como os professores podem elaborar ou encontrar questões adequadas ao tipo de alunos que têm? Até que ponto questões parametrizadas podem ajudar os professores a minimizar a tarefa de elaborar periodicamente questões para avaliar os mesmos conteúdos? Como os professores podem elaborar ou encontrar questões que garantem confiabilidade e validade dos resultados? Será que a forma como os professores encaram um teste é a mesma com que os alunos o encaram?

1.3 Pertinência

A TCT foi reconhecida depois da divulgação do trabalho de Gulliksen (1950) e a TRI, depois da divulgação do trabalho de Lord & Novick's (1968). Estas teorias têm suas origens na Psicometria. Embora não sejam teorias novas, a sua difusão e aplicação não são de domínio de muitos investigadores. A realização deste estudo, com recurso a TCT e TRI, constitui um contributo para os profissionais ligados ao processo educativo. Áreas como pesquisa de mercado, pesquisa de marketing, pesquisas psicológicas, satisfação do cliente, produção de indicadores sócio-económicos, só para destacar algumas, podem aplicar estas teorias. Efetivamente, quanto mais estatísticos, ou outros profissionais que usam a Estatística nas suas pesquisas, souberem da aplicabilidade destas teorias, maiores serão as hipóteses de entenderem melhor os problemas associados à avaliação e, conseqüentemente, obterem resultados satisfatórios, consistentes e válidos.

As avaliações devem satisfazer às reais intenções do avaliador e corresponder às expectativas dos avaliados. Nesse contexto, a avaliação tornar-se-á num fator motivador para a aprendizagem, um meio de distinguir os mais empenhados e direcionar melhor os indivíduos às áreas de seu interesse e domínio próprio, num instrumento que facilite a atividade docente. A TCT e TRI são teorias indicadas para suprir essas necessidades e, como mencionado acima, estas teorias e suas aplicações serão abordadas neste trabalho.

1.4 Organização

Esta dissertação, cujo estudo focaliza a aplicação da TCT e dos modelos unidimensionais da TRI para dados dicotómicos, está estruturada em mais três capítulos para além deste primeiro e foi usada a plataforma *SageMathCloud* com recurso à linguagem \LaTeX . No segundo capítulo são apresentados, a título modelo, quatro exercícios parametrizados. Cada enunciado é acompanhado da respetiva descrição, proposta de resolução e breves comentários de alguns erros cometidos pelos alunos. O terceiro capítulo ficou reservado à abordagem da TCT e da TRI. Em relação à TRI, para além da descrição dos respetivos modelos, são apresentadas as abordagens para estimação das habilidades dos examinandos e dos parâmetros dos itens. Finalmente, tem-se o quarto capítulo onde se faz a aplicação da TCT e da TRI aos dados reais, relativos às avaliações na unidade curricular de Bioestatística realizadas em 2015/16 na Universidade de Aveiro, bem como a discussão dos resultados. Para epílogo da dissertação, algumas considerações finais, em jeito de conclusão, limitações e sugestões para trabalhos futuros, são apresentadas.