



Unlversidade do Minho

Instituto de Educação e Psicologia
Departamento de Psicologia

Natércia Margarida da Conceição Ferreira Morais

Percepções do Ensino pelos Alunos:

Proposta de Instrumento de Avaliação para o Ensino Superior

Braga, 2005

Dissertação apresentada ao Instituto de Educação e Psicologia da Universidade do Minho para provas de mestrado em Psicologia, área de Psicologia Escolar, sob orientação do Professor Doutor Leandro da Silva Almeida, da Universidade do Minho.

*“A paciência é uma árvore de raízes
amargas, mas de frutos muito doces”*

Aforismo persa

Agradecimentos

Ao Professor Doutor Leandro da Silva Almeida, pelo seu apoio incondicional; pelo seu feedback incisivo que trouxe mais rigor e clareza ao meu trabalho e pela autonomia concedida dentro do processo de orientação.

À Professora Doutora Irene Montenegro, pelos momentos de ajuda e pela forma generosa como tem valorizado o meu trabalho.

À Ana Sofia e à Ana Sirage, pela cumplicidade e apoio demonstrados ao longo desta caminhada.

Ao Francisco Mendes, pela sua colaboração fundamental no processo de informatização do questionário administrado.

A todos os alunos e docentes pela sua colaboração no processo de recolha de dados.

Aos meus pais, irmãos e amigos o suporte afectivo para a realização desta tarefa.

Ao meu marido, pelo orgulho que demonstra o qual me faz desejar ir sempre mais longe.

À Universidade do Minho, que a seu lado fui crescendo profissionalmente.

Resumo

A necessidade de maior responsabilidade pela melhoria da qualidade do ensino tem sido um tópico conciliador de interesse crescente no contexto do ensino superior. Em muitos casos, a avaliação da docência tomando as percepções dos alunos tem sido uma fonte importante da qualidade do ensino. Além disso, constitui uma medida directa de satisfação dos alunos, fornecendo informação útil ao docente sobre as suas práticas e feedback à instituição para uma gestão mais adequada dos seus recursos.

Este trabalho procurou validar o questionário *Percepções do Ensino pelos Alunos* (PEA) inspirado no questionário *Student's Evaluation of Educational Quality* (SEEQ; Marsh, 1987, 1991, 2001), com base numa amostra de 41 962 questionários recolhidos no 2º semestre de 2003-04 na Universidade do Minho.

No estudo de validação, optou-se por trabalhar aspectos psicométricos dos itens do questionário de modo a aferir a sua adequabilidade no contexto Ensino Superior em Portugal. No que concerne à análise de consistência interna do PEA, os resultados mostraram níveis elevados de consistência de .97. A par destes resultados, uma análise dos coeficientes de correlação, conjugada com uma análise de conteúdo dos itens, sugere a redução do número de itens por dimensão do PEA. A estrutura factorial não replica os 9 factores/dimensões de Marsh (2001), embora, os resultados se organizem de forma consistente em 5 factores/dimensões: valor das aprendizagens; valor sócio-relacional do processo ensino/aprendizagem; valor do ensino e de outros aspectos do currículo; valor da avaliação e confronto pessoal do aluno com as características (exigências) da disciplina.

Os resultados na análise de variância (MANOVA) confirmaram a existência de efeitos principais do ano e da tipologia de aula em 2 dimensões, cuja diferença é favorável aos alunos dos anos curriculares mais avançados e quando as percepções dos alunos se reportam ao desempenho docente no contexto das aulas práticas. Os efeitos de interacção do ano da disciplina e tipologia de aula foram encontrados em 6 dimensões do PEA. Mais uma vez, as percepções mais positivas verificam-se junto dos anos intermédios e finais dos cursos e no contexto das aulas práticas. Por sua vez, a análise de variância (ANOVA) tomando 4 áreas de curso aponta para percepções mais positivas junto dos alunos frequentando cursos da área das Ciências da Educação, na generalidade das dimensões do PEA.

Os resultados obtidos com a validação do questionário *Percepções do Ensino pelos Alunos* abrem caminhos para a sua utilização na investigação e intervenção a propósito da avaliação da qualidade do Ensino Superior.

Abstract

The need for greater accountability for quality improvement in teaching has become a major issue in higher education in recent years. In many cases, student evaluations of teaching have become a considerable source of data on teaching quality. Furthermore, it has the added benefit of being a direct measure of student's satisfaction, as a performance indicator of teaching effectiveness for feedback's institution to better resources management.

The purpose of this study was to validate the questionnaire named *Percepções do Ensino pelos Alunos* (PEA), inspired in the Student's Evaluation of Educational Quality Questionnaire (SEEQ, Marsh, 1987, 1991, 2001) with a sample of 41 962 surveys collected in the 2nd semestre of 2003-04 in the University of the Minho.

In the validation study of PEA, the psychometrics aspects have been worked in order to evaluate his adequacy in the context of Portugal's higher education. In what concerns the analysis of internal consistency of the PEA, the results have shown high levels of consistency of .97. However, correlation coefficients scores between items of the same factor as well as qualitative analyses suggest a reduction of the number of items by dimension of the PEA. Factor analyses do not reply the 9 factors in Marsh's studies (2001), however, the results seem very consistently organized in five factors/dimensions: value of learning; value of social aspects of teaching and learning; value of teaching and other aspects of the curriculum; value of the evaluation and personal confrontation of the student with the characteristics (requirements) of the unit.

The results of variance analyses (MANOVA) confirmed the existence of main effects of the year and type of class in 2 factors, whose difference is favourable to students of the more advanced curricular years in practical classes. The effects of year and type of class interactions have been found in 6 factors of PEA. Once again, the most positive perceptions were found in intermediate and final curricular years and in practical classes. In turn, the variance analyses (ANOVA) taking different domains, suggest high perceptions in the science education areas in the generality of PEA factors.

The obtained results encouraged the use of the *Percepções do Ensino pelos Alunos* questionnaire to support the research and intervention on teaching quality in higher education.

Índice

Introdução Geral	1
------------------------	---

Parte I

Capítulo 1	13
1.1. Introdução	13
1.2. Necessidade da avaliação dos docentes e da docência	14
1.3. Uma actividade docente para além da sala de aula	17
1.4. Objectivos prosseguidos na avaliação docente	18
1.5. Avaliação pelos alunos	23
1.5.1. Resultados alcançados	26
1.6. Formas alternativas de avaliação	32
1.6.1. Auto-avaliação do docente.....	33
1.6.2. Avaliação pelos pares.....	34
1.6.3. Avaliadores externos	35
1.7. Práticas de avaliação do ensino e dos docentes universitários em diferentes países	36
1.7.1 Balanço da situação portuguesa.....	40
1.8. Em síntese.....	42
Capítulo 2	46
2.1. Introdução	46
2.2. Papel dos alunos na avaliação dos docentes	47
2.3. A formação dos docentes para a docência.....	49
2.4. Tipologia de questionários	52
2.5. Dimensões consideradas nos questionários	54
2.6. Estudos da precisão e validade dos resultados em tais questionários	60
2.7. Em síntese.....	67

Parte II

Capítulo 3	72
3.1. Introdução	72
3.2. Objectivos, questões e hipóteses de estudo	73
3.3. Amostra.....	74
3.4. Instrumento	79
3.5. Procedimentos.....	83
Capítulo 4	85
4.1. Introdução	85
4.2. Análise dos itens	86
4.3. Consistência interna.....	88
4.4. Dimensionalidade	97
Conclusão	116
Bibliografia	122
Anexos	137

Lista de Quadros

- Quadro 3.1:** Questionários recolhidos por licenciaturas agrupadas nos Conselhos de Cursos
- Quadro 3.2:** Questionários por Conselhos de Cursos e anos das disciplinas
- Quadro 3.3:** Descrição das dimensões do PEA e respectivo número de itens
- Quadro 4.1:** Análise da distribuição das respostas aos itens do PEA
- Quadro 4.2:** Análise da consistência dos itens do PEA
- Quadro 4.3:** Correlações dos itens na dimensão Relevância da Aprendizagem
- Quadro 4.4:** Correlações dos itens na dimensão Empenho docente
- Quadro 4.5:** Correlações dos itens na dimensão Organização/Clareza
- Quadro 4.6:** Correlações dos itens na dimensão Interação da Turma
- Quadro 4.7:** Correlações dos itens na dimensão Relação Docente/Aluno
- Quadro 4.8:** Correlações dos itens na dimensão Profundidade na Abordagem dos Assuntos
- Quadro 4.9:** Correlações dos itens na dimensão Avaliação/Classificações
- Quadro 4.10:** Correlações dos itens na dimensão Trabalhos/Leituras
- Quadro 4.11:** Correlações dos itens na dimensão Carga de Trabalho/Dificuldade
- Quadro 4.12:** Análise factorial
- Quadro 4.13:** Análise factorial do PEA no Conselho de Cursos de Ciências Económicas Empresariais e Políticas
- Quadro 4.14:** Médias e desvio-padrão dos resultados nas dimensões do PEA em função dos Conselhos de Cursos e do ano da disciplina
- Quadro 4.15:** Médias e desvio-padrão dos resultados nas dimensões do PEA em função dos Conselhos de Cursos e da tipologia de aula
- Quadro 4.16:** Análise de variância (MANOVA) dos resultados nas dimensões do PEA em função da tipologia de aula e ano da disciplina
- Quadro 4.17:** Análise de variância One Way (ANOVA) dos resultados nas dimensões do PEA em função da área do curso

Lista de Gráficos

Gráfico 3.1: Descrição da amostra por Conselhos de Cursos

Gráfico 4.1: Efeitos de Interação na dimensão Relevância da Aprendizagem

Gráfico 4.2: Efeitos de Interação na dimensão Empenhamento Docente

Gráfico 4.3: Efeitos de Interação na dimensão Interação da Turma

Gráfico 4.4: Efeitos de Interação na dimensão Docente/Aluno

Gráfico 4.5: Efeitos de Interação na dimensão Avaliação/Classificações

Gráfico 4.6: Efeitos de Interação na dimensão Carga de Trabalho/Dificuldade

Introdução Geral

A Universidade é ainda hoje considerada como uma das instituições mais prestigiadas e que terá tido como mérito a garantia da sua continuidade desde a Média até à actualidade (Soares, 2004). Para isso terá contribuído a convicção generalizada de que as universidades constituíam uma referência de qualidade para outras instituições, possuindo mecanismos próprios que asseguravam a qualidade do seu trabalho (Ferreira, 2003).

Porém, o processo de democratização do acesso ao Ensino Superior que se verificou na década de 70 em Portugal, resultante não só da afirmação do princípio de igualdade de oportunidades mas também da pressão social resultante de uma visão do ensino como instrumento privilegiado de ascensão social, induziu um crescimento acelerado deste subsector do sistema educativo (decorrente quer do aumento populacional, da melhoria das condições de vida, do alargamento da escolaridade obrigatória e da associação positiva entre níveis académicos e a inserção no mercado de trabalho). Paralelamente, a este crescimento acelerado do sistema educativo associa-se a expansão e a multiplicação do número e diversidade dos estabelecimentos de Ensino Superior (públicos e privados, universitários e não-universitários), bem como a sua pulverização geográfica (Almeida *et al.*, 2002; Estanque & Nunes, 2001; OECD, 2001). Com efeito, a este *boom* não foi alheio o desenvolvimento de duas lógicas aparentemente contraditórias orientadas quer pelo reforço da rede pública de estabelecimentos de ensino quer pela lógica de privatização deste nível de ensino. Só no Ensino Superior público aponta-se para cerca de 400 mil alunos a frequentar os 1558 cursos (995 públicos e 563 não públicos) leccionados em 168 instituições públicas e 137 privadas (DAPP, 2000). Proporcionalmente considerando o número de habitantes, Portugal possui um maior número de Instituições de Ensino Superior (IES) por comparação com os Estados Unidos da América, a Espanha, a França, a Itália ou a Holanda (Crespo, 2003).

Assim, as mudanças estruturais que conheceu o Ensino Superior não alteraram apenas a configuração da oferta do Ensino Superior, mas traduziram-se numa diversificação das origens sociais e culturais e, mesmo, geográficas dos alunos, das suas trajectórias escolares, dos seus

planos vocacionais de ingresso e das suas motivações/representações (Almeida *et al.*, 2002; Arroteia, 1996; Balsa *et al.*, 2001; Braga da Cruz *et al.*, 1995; Fernandes, 2001; Gago *et al.*, 1994). Com efeito, a massificação verificada pode ser explicada pela intersecção de vários factores. A inexistência de uma política de *numerus clausus*, e a maior importância atribuída à educação como um factor de desenvolvimento económico do País, conduziram a uma aposta importante neste subsector educativo, tal como é visível nos valores de investimento atribuídos dentro do orçamento de Estado comparativamente aos outros países da OECD (Ferreira, 2003; Soares, 2004). Por sua vez, o número e o grau de qualificações docentes evoluiu de uma política restritiva que comprometia a evolução da carreira dos docentes universitários e não dava reconhecimento científico dos doutoramentos realizados no estrangeiro, até uma situação em que está estabelecida legalmente a carreira docente, e onde, apesar de não ideais, existem condições adequadas para a investigação² (Ferreira, 2003). Assim, a expansão e a diversificação do ensino superior em Portugal, aquando da entrada na União Europeia, conduziram à evolução algo “descontrolada” da oferta formativa, de professores e de alunos. Só agora aparece reconhecido, por exemplo, que a qualidade da investigação desenvolvida nas instituições é variável assim como o número, o nível e a qualificação dos docentes nas diversas instituições de ensino superior nacionais. A este propósito, Simão, Santos e Costa (2003) interrogam-se acerca da legitimidade, e muitas vezes da transparência, desta evolução, que julgam ter sido “apressada” e assente numa lógica de funcionamento de mercado e não do “resultado de tradições culturais e científicas”.

Esta expansão e diversificação aparecem, muitas vezes, associadas a um forte desequilíbrio do binómio qualidade - quantidade (Simão, 2003ab) e que fez com que, particularmente nos finais da década de 80 e na década 90, se tivesse verificado uma massificação acentuada no ensino superior (Almeida *et al.*, 2001; Conceição *et al.*, 1998; Braga da Cruz *et al.*, 1995; Eurydice, 2000; Simão, Santos & Costa, 2003). Com a heterogeneidade instalada, a que Simão (2003b) denominou “multiplicidade anárquica” ou “estrangulamento e ruptura” (Braga da Cruz *et al.*, 1995), aliadas ao crescimento desenfreado, não houve tempo e espaço para a monitorização da qualidade. Face a estas circunstâncias, emergiu a necessidade de se criar referenciais de qualidade passíveis de garantir a coerência do sistema no quadro dos princípios que orientam a autonomia dos estabelecimentos de ensino superior. Tal preocupação encontra-se expressa no documento de orientação para a política do ensino superior sobre

¹ Possibilidade de dispensa de aptidão para alunos com classificação superior a uma data nota.

² Na sua maioria, esta investigação é apoiada externamente por financiamentos conseguidos pelas próprias equipas de investigação.

autonomia e qualidade (ME/SEES, 1997). De facto, na revisão de 1997, a autonomia universitária passou a ser condicionada não só por uma lei própria, mas também pela qualidade de ensino. Assim, em Portugal a lei da Avaliação passou a estar intimamente relacionada com a lei da Autonomia, referindo-se que, a “*autonomia só verdadeiramente se compreende, justifica e ganha sentido entrelaçada com a ideia de responsabilidade*”, sendo que essa responsabilidade configura uma prestação de contas do seu funcionamento em órgãos políticos e à sociedade em geral, através de um processo de avaliação que combina avaliação interna e externa. A uma instauração da lógica da crescente autonomia das Universidades, sobretudo nos países mais avançados da Europa, corresponde uma fórmula que combina maior autonomia, maior auto-regulação e responsabilização (*accountability*) susceptível de gerar maior produtividade e eficiência. De acordo com Amaral (1998), deixa de haver interferência do Governo para passarem a existir leis de avaliação para aferir a qualidade das instituições (e.g. lei n.º 1/2003, de 6 de Janeiro) e o aumento consequente de dispositivos de controlo interno (ministérios, ordens, equipas de avaliação externa). Com efeito, durante a década de 80, o movimento de autonomia parece reconhecer a individualidade e a especificidade de cada uma das instituições, concedendo-lhes uma auto-determinação para a definição dos processos, embora na prática, mantendo um controlo sobre os resultados obtidos (Seixas, 2003).

É neste quadro de ponderação da quantidade-qualidade que tem granjeado força o movimento de questionamento da qualidade do Ensino Superior e a instauração progressiva de mecanismos de avaliação. Logicamente que não é fácil definir qualidade. Aliás, vários autores sublinham a dificuldade em encontrar uma definição de qualidade no Ensino Superior, chegando mesmo a colocar a questão “*can we improve quality if we cannot define it?*” (Bouge & Saunders, 1992, p.1 in Ferreira, 2003). Emergem, de imediato, as perguntas: O que é a qualidade? Qual é a sua essência? São perguntas de difícil resposta e, face à esterilidade da sua procura, sugere-se que não haverá uma definição de qualidade (Castro, 1993; Green, 1995; Vroeijensteijn, 1995) ou, então, que existem tantas definições quantos os actores interessados no ensino superior (alunos, docentes, governo, empregadores, etc.), multiplicados pelos objectivos ou dimensões que estes actores distinguem (Santiago, 1997).

A qualidade no Ensino Superior é multidimensional e de natureza pouco precisa, sendo que qualquer tentativa de legislar uma única definição estará votada ao fracasso (Brennan & Shah, 2000). Aliás, Green (1995) alerta para o facto de existirem conceitos divergentes habitualmente usados quando se fala em qualidade no Ensino Superior: conceito tradicional, conformidade face a

uma especificação ou a um padrão, ajustada ao objectivo, eficácia em atingir objectivos institucionais e satisfação dos clientes. No entanto, reconhecer esta diversidade dos possíveis significados não nos permite alhear da responsabilidade de trabalhar para manter e promover a qualidade (Vroeijerstijn, 1995). Aceitar a legitimidade dos argumentos apresentados, implica reconhecer a improbabilidade de encontrar um método único de avaliação que possa emergir para satisfazer os propósitos de todas as partes interessadas e envolvidas no processo (UNESCO, 1998). Assim, teremos que deixar de falar de “qualidade” e passar a considerar “qualidades” no Ensino Superior (Vroeijerstijn, 1995).

Adicionalmente, o modelo de excelência e auto-avaliação da EFQM (*European Foundation for Quality Management*) também destaca a multidimensionalidade inerente à qualidade no Ensino Superior. Este modelo incide sobre os “meios” para a qualidade (liderança, pessoas, política e estratégia, parcerias, recursos e processos) e sobre “resultados” da qualidade (resultados-pessoas, resultados-clientes; resultados-sociedade; e resultados-chave do desempenho) através da promoção transversal da inovação e da aprendizagem (Gama, 2004). Porém, a transposição deste conceito de origem industrial para o mundo universitário está longe de ser linear e muito menos amplamente aceite pela comunidade académica. Vários autores das Ciências da Educação têm vindo a alertar para os riscos associados à importação linear de modelos cuja génese se encontra associada ao mundo empresarial (Castro, 1993), por exemplo, a consideração dos alunos como clientes (Coaldrake & Stedman, 1998; Lourtie, 1998).

Apesar de todos estes pontos menos claros, a procura da qualidade é hoje apanágio do Ensino Superior em Portugal, integrando os seus objectivos. Estando tentos à complexidade destas Instituições, mesmo assim importa reflectir sobre o momento presente e o questionamento que é feito em torno dos seus objectivos e missão. Este questionamento é indissociável da qualidade do ensino/aprendizagem e, daquilo que mais nos interessa nesta dissertação, a avaliação da qualidade do ensino.

A qualidade no ensino superior tem ocupado, nos últimos anos, um lugar de destaque no discurso dos agentes educativos, nas preocupações dos decisores políticos e na literatura especializada (Marsh, 2001; Simões, 2000). Por sua vez, esta significativa expressão da temática representa, em larga medida, a expansão para o domínio educativo de uma filosofia de gestão que, desde há muito, se vem tentando implementar nas empresas industriais e de serviços (Rego, 2003; Caetano & Vala, 2002), ganhando por isso, particular relevância a questão do investimento da qualidade na educação terciária. Com efeito, a qualidade e a excelência são amplamente referidas

e objecto dos processos de avaliação das instituições, cursos e docentes, não porque tais características não existam nas Instituições de Ensino Superior, mas porque nelas assenta a “prestação pública de contas” por parte das instituições, assim como o seu financiamento ou a aplicação de um financiamento diferenciado entre as instituições.

A avaliação da qualidade no ensino superior não foi “inventada” na década 90. Contudo, e até aos meados da década de 80, o debate e as práticas situaram-se, em larga medida, dentro do próprio sistema (Green, 1995; Vroeijensteijn, 1995) ou, se realizado no exterior, assumia contornos pouco claros e abstractos (Perry, 1991). Esta postura não exclui certamente a ideia de que as universidades constituíam uma referência de qualidade para outras instituições. Com efeito, a Universidade possuía mecanismos próprios que asseguravam a qualidade do seu trabalho (qualificações dos alunos na admissão e obtenção do diploma; prestações dos docentes para a sua evolução na carreira) e que durante muito tempo contribuiu para a “*defensive wall behind which the academy has been able to hide*”.

É neste contexto que, em Portugal, tem vindo a ser cada vez mais notória a preocupação com as questões relativas à avaliação do ensino, tomando a figura do docente. Com efeito, o interesse pelo estudo da qualidade do desempenho docente não é um fenómeno recente, tendo-se desenvolvido nos Estados Unidos, pelo menos desde a década 20. Contudo, entre nós, só na última década o tema ganhou a devida expressão (e.g. tendências como o movimento a favor da qualidade total, educação para o desenvolvimento sustentável⁴, avaliação institucional, sistemas de financiamento, autonomia e avaliação, Declaração de Bolonha). Aliás, a emergência destes movimentos tem trazido, ao de cima, a importância da função docente, a par das funções de investigação e de prestação de serviços à comunidade.

A evidência teórica e empírica acumulada nas últimas décadas permite-nos concluir que a avaliação do ensino, conceptualizada como uma actividade de maior visibilidade social das Universidades nacionais e estrangeiras, é uma tarefa desafiadora (que dificilmente pode ser operacionalizada com base num critério único de avaliação), exigindo, por isso, a adopção de instrumentos de natureza multidimensional em complementaridade com outras fontes informativas que atestem essa qualidade (auto-avaliação docente, avaliação por pares, observadores externos, *portfólios*). Desde logo, esta diversidade de meios decorre da complexidade deste processo

³ “...*parede defensiva atrás da qual a academia tem conseguido esconder-se...*” (Harvey & Knight, 1996 in Ferreira, 2003).

⁴ Estratégia Nacional de Educação para o Desenvolvimento Sustentável (ENEDS) Ponto 5.6: Universidade e Institutos Politécnicos (Freitas, 2004).

avaliativo e, sobretudo, da diversidade de dimensões e de variáveis a considerar. A par desta complexidade, importa a conceptualização da docência no ensino superior baseada numa visão inovadora da sua função e que se deverá traduzir no envolvimento dos docentes nos processos de formação no sentido de atingir uma duas missões mais importantes: o sucesso académico dos alunos (Tavares, 2001).

Porém, a revisão da literatura sugere que dois propósitos/funções têm orientado a avaliação da qualidade do desempenho docente (i) a que tem procurado servir objectivos sumativos, de natureza retroactiva, utilizada para a prestação de contas, e (ii) a que destaca a emergência de concepções mais descritivas, marcada por objectivos formativos, de natureza mais proactiva, destinada ao desenvolvimento profissional como suporte à melhoria progressiva do sistema. Claro está que, sempre que possível, a opção deve ser por um modelo multidimensional susceptível de abarcar múltiplos aspectos do ensino/aprendizagem, procurando por sua vez, a harmonização destas duas perspectivas num compromisso claro e sério entre as abordagens do ensino ditas tradicionais (centração no docente e valorização dos resultados) e as abordagens do ensino/aprendizagem (centração no aluno, papel mediador do docente e o enfoque nos processos).

A partir da caracterização das experiências no campo da avaliação dos docentes no Ensino Superior e do estudo de casos de experiências mais significativas em termos nacionais e internacionais, é possível concluir que as Instituições de Ensino Superior têm utilizado várias fontes de informação pela consciência de que a avaliação não deve ser realizada a partir de indicadores únicos e isolados, dando-lhes um carácter participativo, e, por outro lado, procurando preservar a confidencialidade dos dados relativos aos docentes.

Uma das linhas de estudo que se tem destacado é a que se dedica a investigar a avaliação da qualidade do desempenho pedagógico docente tomando as percepções dos alunos com base na administração de questionários (forma de avaliação que goza de maior suporte empírico por comparação com outros métodos). De acordo com alguns autores, a legitimidade da avaliação do ensino pelos alunos ganha força ao defender-se que os alunos são os primeiros agentes da aprendizagem e, como clientes, têm o direito de participar, em certa medida, na avaliação do ensino. Contudo, alguns docentes resistem a ser avaliados pelos alunos, argumentando que a influência subjectiva das suas opiniões tende a influenciar negativamente os resultados nas avaliações. Todavia, a revisão da literatura defende que a avaliação da qualidade docente inferida com base nas percepções dos alunos, quando positivamente assumida, pode facultar uma

avaliação de tipo diagnóstico (sinalização dos pontos fortes e fracos da actividade docente⁵) no sentido do desenvolvimento e aperfeiçoamento profissional dos docentes. Por outro lado, as avaliações feitas a partir da recolha de informações junto dos alunos são, no seu conjunto, fiáveis e estáveis; dependem mais do docente do que da disciplina; são relativamente válidas e pouco afectadas pelos enviesamentos potenciais e assumem-se como ferramentas úteis de melhoria da qualidade do ensino, quando combinadas com estratégias concretas de intervenção (Marsh, 2001).

Não é pois de admirar que existam muitos estudos centrados na componente estrutural dos questionários, onde é possível constatar a presença de dimensões/facetas conceptualmente distintas para avaliar a qualidade do ensino. Contudo, a revisão da literatura confirma a ideia de que as teorias do ensino/aprendizagem não têm sido explicitamente referenciadas para o desenvolvimento destes instrumentos, colocando por sua vez, o desafio de maior integração das concepções cognitivistas, construtivistas e interaccionistas do ensino/aprendizagem, bem como de análises empíricas com recurso à análise factorial, análises multimétodo-multitraço e discussão do constructo com base no *feedback* dos alunos e docentes. Um olhar atento sobre a literatura permite verificar a existência de um elenco significativo de factores associados às práticas e aos comportamentos dos docentes com efeitos positivos ou negativos sobre os alunos e, globalmente, sobre o próprio processo de ensino/aprendizagem. De resto, com base na literatura revista sobre dimensões que definem a qualidade do ensino no ensino superior e, em particular, a qualidade do docente, admite-se que existe um relativo consenso entre alguns investigadores sobre as características que os definem. Desenvolveremos este tópico mais à frente nesta dissertação.

Objectivos gerais deste estudo e sequência dos capítulos

Assim, o presente estudo enquadra-se numa temática que tem vindo a merecer alguma atenção por parte das Instituições de Ensino Superior. Nestas começa a ser visível um interesse crescente pelas questões da avaliação da qualidade do ensino, em particular a avaliação do desempenho docente. Neste contexto, o objectivo fundamental desta dissertação procura contribuir

⁵ Na área da Gestão de Recursos Humanos, fala-se em *Swot Analysis* (identificação e análise de pontos fortes, pontos fracos, oportunidades e ameaças).

para uma apresentação e discussão da investigação internacional neste domínio e, ao mesmo tempo, contribuir, na sua parte empírica, para a adaptação e validação de um instrumento de avaliação. Para o efeito, tomaremos as percepções que uma amostra representativa de alunos da Universidade do Minho possuem acerca do processo ensino/aprendizagem com base nesse questionário. Estamos face a um instrumento que procura avaliar a actividade docente com base nas percepções dos alunos por referência às dimensões representativas da sua actuação na sala de aula, previamente identificadas na literatura especializada.

A realização deste trabalho foi, sobretudo, um resultado concreto de uma tentativa de organização das leituras e ideias da autora a respeito do fenómeno em análise. Desde logo emerge a crença de que os aspectos da relação pedagógica e a análise da sua influência no desempenho dos alunos do ensino superior, não têm sido objecto de estudos sistemáticos. Por outro lado, temos a percepção que a avaliação contínua da qualidade do ensino/aprendizagem, da qual os docentes constituem variáveis cruciais, ganha cada vez maior importância numa época onde se perspectiva, até 2010, a constituição de um Espaço Europeu de Ensino Superior, marcada pelas aproximações sucessivas à certificação da qualidade do ensino. Logicamente que todas estas leituras e reflexões fazem sentido no quadro da sua pertença ao Gabinete de Avaliação e Qualidade do Ensino (GAQE) e às actividades que aí se desenvolvem desde 2004.

Na prossecução deste objectivo, a presente dissertação encontra-se organizada em 4 capítulos, sendo os 2 primeiros dedicados à revisão da literatura no domínio e os restantes ao estudo empírico. Assumindo desde logo uma perspectiva multidimensional da avaliação da qualidade do ensino, no Capítulo 1 procuramos traçar o enquadramento da necessidade da avaliação do ensino à luz dos desafios recentes colocados às IES e, em particular, da avaliação do docente e da docência nas suas múltiplas vertentes, bem como os objectivos prosseguidos nesta avaliação (modelos existentes e as suas funções). Adicionalmente, procuramos destacar a necessidade da avaliação das percepções do ensino pelos alunos, apresentando para isso os resultados que justificam a sua inclusão num plano mais vasto de avaliação da qualidade do ensino. Por outro lado, procuramos ainda descrever algumas formas alternativas dessa avaliação (auto-avaliação docente, avaliação por pares, avaliadores externos) e os resultados alcançados com a sua utilização. Por último, fazemos referência às grandes tendências internacionais da avaliação do ensino e dos docentes universitários nos últimos anos, e, em jeito de reflexão, sugerimos algumas orientações gerais para um processo de avaliação emergente no contexto do ensino superior.

O Capítulo 2 encontra-se reservado à reflexão em torno do papel dos alunos na avaliação do ensino, e em particular às suas percepções sobre a docência, destacando a possibilidade de tais percepções encontradas nos questionários reflectirem a qualidade do processo ensino/aprendizagem. Neste sentido, apresentamos algumas das principais teorias e conceitos associados ao ensino/aprendizagem que cremos contribuirão para o esclarecimento do fenómeno em análise. Procura-se, ainda, descrever a tipologia e as dimensões consideradas nos questionários de avaliação do ensino pelos alunos. Por último, fazemos referência com base na consulta da literatura, aos estudos de precisão e validade realizados tomando os resultados obtidos com a sua aplicação.

No Capítulo 3 descrevemos a metodologia do estudo empírico que orientou a nossa investigação. Começamos por apresentar as questões e os objectivos que pautaram o estudo, a caracterização da amostra que serviu de base ao estudo de adaptação e validação do instrumento de medida utilizado (constituída por um painel transversal que integra 41 962 questionários de alunos de 49 licenciaturas e de 6 cursos de complemento/qualificação, que se encontravam a frequentar as disciplinas anuais e semestrais do 2º semestre no ano lectivo 2003-04), bem como os demais procedimentos metodológicos adoptados. O questionário utilizado para avaliar as percepções dos alunos acerca da qualidade do ensino na Universidade do Minho designa-se *Percepções do Ensino pelos Alunos* (PEA), e foi inspirado no questionário *Student's Evaluation of Educational Quality* (SEEQ; Marsh, 1987, 1991, 2001). Tratando-se de um instrumento mundialmente utilizado para efeitos de avaliação da qualidade do ensino, julgou-se importante rentabilizá-lo para efeitos deste estudo, procurando ao mesmo tempo conjugar a experiência de 14 anos de avaliação do ensino ministrado da Universidade do Minho⁶.

No Capítulo 4, apresentamos os resultados das análises estatísticas conduzidas quer considerando os resultados item a item, quer assumindo as dimensões ou facetas do questionário. Esta análise considera a amostra global de alunos repartidos pelo ano curricular a que pertencem as disciplinas cuja docência foi avaliada, pela natureza prática ou teórica das aulas, assim como pelos Conselhos de Cursos a que pertencem.

Por último, esta dissertação apresenta uma conclusão do trabalho realizado onde, sobretudo, queremos destacar as suas implicações, seja ao nível da investigação seja ao nível da avaliação e da intervenção, tendo em vista o incremento da qualidade do ensino. Claro que não

⁶ Em 1991 foi instituída, por Despacho Reitoral, a Avaliação do Ensino Ministrado através da aplicação do Questionário da Avaliação do Ensino Ministrado (QAEM) à população discente (Montenegro, 2002).

temos como pretensão mostrar um quadro completo e profundo do tema. Pretendemos, sobretudo, apresentar uma proposta de avaliação da qualidade do ensino e justificar a sua utilização, com base, respectivamente, nas análises dos resultados obtidos com o questionário e na síntese da investigação da área. Acreditamos que, da posse desta informação e mesmo não sendo mais que percepções dos alunos, a instituição, os seus órgãos e os seus docentes poderão avançar para melhores práticas de ensino, maior incentivo às aprendizagens pelos alunos e à melhoria geral da formação no seio da Universidade.

Parte I

CAPÍTULO 1

Avaliação do ensino no Ensino Superior

"Teaching is being taken more seriously"

(Seldin, 2000)

1.1. Introdução

Actualmente, a Universidade Portuguesa vê-se confrontada com a necessidade de responder ao desafio da avaliação do desempenho dos docentes no quadro institucional da autonomia universitária, na perspectiva da formação permanente, metodologicamente articulada com os resultados da avaliação dos cursos e conceptualizada como instrumento estratégico de mudança (Cachapuz, 2001). A par deste repto, encontra-se uma das missões da Universidade que inclui, entre outras, a promoção e incentivo de acções genericamente orientadas para a qualidade do ensino/formação como um dos seus vectores estratégicos⁷.

Os factores subjacentes ao interesse crescente pelo desenvolvimento de dispositivos de avaliação em contexto universitário podem ser de natureza diversa. O primeiro prende-se com a exigência cada vez maior da qualidade das formações que as universidades fornecem, aliada a uma cada vez maior competitividade entre as Instituições de Ensino Superior (IES) no quadro global da sociedade de conhecimento. Tudo isto implica a necessidade de reinvenção de uma nova relação com o conhecimento em geral, e com o conhecimento pedagógico, em particular (Atkins, 1987 in Cachapuz, 2001).

O ensino continua a constituir o campo de actividade com maior visibilidade social das Universidades, actividade da qual os docentes são directamente responsáveis, desmultiplicando-se em diferentes formas da sua avaliação, que não se esgotam numa avaliação pelos alunos por questionário (Estrela & Simão, 2003). Por sua vez, a actividade docente no contexto do Ensino Superior é uma tarefa complexa e exigente do ponto de vista intelectual e social, envolvendo também competências específicas que podem ser adquiridas e melhoradas (e.g. programas de apoio/consultoria profissional).

A história da avaliação tem-se organizado em torno de duas tendências principais: avaliação processual formativa e avaliação final ou de resultados. Enquanto a avaliação com objectivos formativos se relaciona com o desenvolvimento profissional e a melhoria do desempenho

⁷ Definição de objectivos de Missão Institucional da Universidade do Minho, à luz do Novo Sistema Integrado de Avaliação do Desempenho da Administração Pública - Lei n.º 10/2004, de 22 de Março (cf. Documento de trabalho, 2004).

do docente, a avaliação sumativa encontra-se orientada para a tomada de decisões administrativas.

Uma das formas possíveis de avaliação da qualidade do ensino baseia-se na avaliação das percepções dos alunos. Apesar das críticas frequentemente lançadas pelos docentes (avaliação redutora e unilateral que não contempla a articulação com outras fontes informativas) e pelos alunos (falta de tradução em acções concretas), é aquela que tem sido mais referenciada na investigação disponível.

Seguindo o princípio da diversidade, qualquer dispositivo de avaliação deve ter a possibilidade de triangular e confrontar as diversas informações (auto-avaliação do docente, avaliação por pares ou observadores externos, entre outros), embora se detectem potencialidades e limitações na sua aplicação aos contextos avaliativos. A análise das tendências nacionais e internacionais no domínio da avaliação, destaca a ideia de que, embora se verifiquem lógicas organizativas diferentes nos casos analisados, as instituições apresentam na sua cultura institucional uma elevada preocupação com a problemática da avaliação, a par de uma forte pressão para se melhorarem os padrões de qualidade do ensino.

1.2. Necessidade da avaliação dos docentes e da docência

A sociedade exige cada vez mais a demonstração do valor do trabalho dos docentes do Ensino Superior, pela convicção generalizada de que a qualidade da actividade destes profissionais raras vezes é globalmente avaliada (Abrantes & Valente, 2000). Com efeito, os profissionais da educação sempre foram avaliados no exercício ou nos resultados das suas funções, de um modo mais ou menos informal, por alunos e colegas. Contudo, a consagração de sistemas de avaliação formal dos docentes é um fenómeno relativamente recente, na maior parte dos sistemas educativos. Como refere Hadji (1995), a partir do momento em que a necessidade de avaliação entrou no domínio das práticas sociais e nas “coisas da educação, *“não se vê como é que os docentes possam escapar a essa regra geral”* (p.27). Embora no Ensino Superior os docentes ocupem a maior parte do seu tempo no ensino, tanto a nível graduado como pós-graduado, tem sido mais fácil documentar e apreciar a sua actividade de investigação do que documentar a sua actividade de ensino. Por outro lado, a raridade de estudos sobre as questões do desempenho pedagógico do docente no Ensino Superior em Portugal, reflecte o peso ainda marginal que essas

questões assumem no quadro das avaliações institucionais e na carreira académica dos docentes (Vieira *et al.*, 2002). Contudo, e de acordo com Boyer (1990, in Abrantes & Valente, 2000), é hoje reconhecido por muitos docentes que não é desejável continuar a usar apenas a investigação e as publicações como critério fundamental para a sua promoção profissional. Estas preocupações estão no cerne da discussão em torno da necessidade da revisão do Estatuto da Carreira Docente Universitária (ECDU).

Uma das linhas de estudo que tem apontado para a necessidade da formação pedagógica e da inovação da função docente é aquela que se tem preocupado com o sucesso académico dos alunos (Tavares, 2001). De acordo com o autor, a pedagogia na universidade não pode ser desligada do sucesso académico (assumido aqui em sentido lato) e, por conseguinte, da docência e da inovação. Por isso, é necessário equacionar uma formação do docente em novos moldes em que o peso da preparação pedagógica seja semelhante ao peso atribuído à preparação científica e de investigação. É no quadro do envolvimento em processos de formação pedagógica por parte dos docentes (transmutação psicológica, social e cultural das formas de ensinar e aprender), e por outro, em contextos de aprendizagem que traduzam efectivamente a vontade de aprender dos alunos, ambos numa matriz colaborativa e profissional, que se constitui a chave para alcançar esse objectivo (*ibidem*).

Sabendo que o ensino tem recebido menos atenção do que a investigação, existe a necessidade dos indivíduos e as instituições se posicionarem de modo a especificarem os critérios de qualidade do ensino e da sua eficácia (Biggs, 1990), procurando aumentar a consciência do docente a respeito da sua própria filosofia, objectivos e estilos de ensino, assim como aumentar a sua familiaridade e flexibilidade na utilização de métodos alternativos para alcançar os seus objectivos (Fonte & Vasconcelos, 2001; Knight, 2002; Macedo, 2001). Com efeito, o docente no contexto do Ensino Superior tem, de um modo geral, sofrido a falta de uma formação académica e profissional no campo da docência e da avaliação, e daí a emergência do crescente interesse pelas questões da formação pedagógica dos docentes universitários (programas de apoio profissional).

A actividade docente no Ensino Superior é uma tarefa complexa e exigente quer do ponto de vista intelectual quer social (Marsh, 2001). Ela envolve competências específicas que podem ser adquiridas e melhoradas num cenário de auto-formação (estratégia de formação dominante em Portugal) ou num cenário de frequência de programas de formação (PFPEs)*; desenhados no

* Programas de formação com sucesso variável em vários países, sobretudo da Europa do Norte e que, embora sob formas variadas, exploram a perspectiva do docente reflexivo, pesquisador do seu próprio ensino (aliás conceito cujo principal valor heurístico, é o de articular de forma inovadora, ensino com investigação).

conteúdo e na forma, para promover e apoiar a reflexão crítica de cada docente sobre problemas do seu ensino e a reinvestir posteriormente e de forma inovadora, na sua prática profissional (incentivo à meta-avaliação). Por exemplo, a maioria das universidades britânicas requer que os novos docentes recebam formação em fundamentos de aprendizagem de adultos, avaliação e auxílio na aprendizagem e a certificação dos docentes através da obtenção de um curso de pós-graduação (Cowan, 2004). A par deste movimento de formação impulsionado por factores de ordem externa (e.g. avaliação e financiamento das Universidades; orientações decorrentes da Declaração de Bolonha), decorre a necessidade de encarar o paradigma emergente da incerteza na relação formação/trabalho (Canário, 2001). Vivemos um momento pautado pela crescente mobilidade profissional, pela rápida desactualização da informação e pela tendência actual das organizações de trabalho para uma “cultura de interacção e de resolução de problemas”, o que desaconselha uma visão tradicional das práticas de formação. A constante evolução e incerteza parecem trazer apenas uma certeza – é necessária uma actualização contínua não apenas de conteúdos, mas também de estratégias e metodologias de ensino, e ainda, de ferramentas e suportes, como por exemplo, as novas tecnologias. Aliás, alguns autores como Slotte, Wangel e Lonka (2001) defendem que os docentes estão interessados e receptivos à aprendizagem e utilização das novas tecnologias no seu trabalho, apesar de, neste momento a sua aplicação no trabalho de investigação ser significativamente superior à sua utilização na actividade docente. Por último, e no quadro destes novos desafios, importa atender ao enquadramento da internacionalização do Ensino Superior pela visão de qualidade que propõe (Gomes, 2003). Para este autor, o recrutamento e mobilidade internacional de docentes/investigadores e de alunos tem de ser encarado como factor decisivo à internacionalização das instituições.

Embora a literatura no domínio da Psicologia e das Ciências de Educação reconheça o processo de ensino/aprendizagem como uma das áreas de intervenção com maior potencial “*não há lugar algum onde os resultados positivos da gestão da qualidade total sejam mais necessários do que na sala de aula*” (Helms & Key, 1994, p.97 in Rego, 2003) o que se constata é que “*apesar do extenso discurso sobre a qualidade, a actividade nuclear das universidades - a docência - permanece em grande medida intacta*” (Hansen & Jackson, 1996, p.211-212 in Rego, 2003). Contudo, hoje é reconhecido que para aumentar a “*diferenciação através de um portfólio constituído por professores competentes em saber e em saber comunicar o que sabem*” (Ferreira, 2004) na Universidade, a avaliação deverá fazer parte integrante do processo

ensino/aprendizagem, aliada à necessidade permanente de actualização e contínua reflexão acerca da qualidade no Ensino Superior.

1.3. Uma actividade docente para além da sala de aula

Vários são os trabalhos sobre a definição e classificação do trabalho dos docentes no Ensino Superior (Abrantes & Valente, 2000). Por exemplo, Braskamp e Ory (1994) defendem uma classificação pragmática que permite tipificar os diferentes parâmetros de avaliação docente: ensino, investigação e actividade criativa, prática e serviço profissional; cidadania e participação cívica. No caso do ensino, importa que a conceptualização da sua eficácia não se esgote em termos de competências genéricas aplicáveis a qualquer tema ou disciplina. Importa, pois, a demonstração de uma pedagogia susceptível de não apenas transmitir conhecimentos mas também de transformar e ampliar o conhecimento. Mais uma vez, Braskamp e Ory (1994) consideram na categoria de ensino um conjunto de actividades que classificaram nas seguintes áreas: i) ensino dos alunos em cursos, laboratórios, clínicas, estúdios, *workshops*, seminários e a gestão dos cursos (desde o planeamento das aprendizagens e de experiências, a classificação e o registo de resultados); ii) supervisão e orientação de alunos (em contexto laboratorial, trabalho de campo, aconselhamento sobre carreiras, cursos, supervisão no estudo individual e em trabalhos de teses e dissertações); iii) desenvolvimento de actividades de aprendizagem (consulta bibliográfica, revisão, reorganização de currículos e dos cursos, materiais de ensino, manuais e software didáctico); e iv) actividades promotoras do desenvolvimento profissional enquanto docente (avaliação do ensino dos colegas, realização de investigação sobre o ensino e frequência de cursos para o desenvolvimento profissional).

Por sua vez, no Reino Unido, Elton e Partington (1991 in Abrantes & Valente, 2000) tentaram, relativamente à categoria de ensino, enunciar um conjunto de critérios que deveriam ser considerados num esquema geral para a caracterização do perfil docente no contexto do ensino superior. Esses critérios conduziram a um conjunto de categorias relativas ao ensino na instituição e aos critérios por regulação externa. No que concerne ao ensino na instituição, os autores referem os seguintes critérios: preparação para o ensino, qualidade do acto de ensinar, volume do ensino, comunicação geral com os alunos fora das aulas, procedimentos de avaliação, avaliação do próprio ensino, gestão do ensino, melhoria e investigação do próprio ensino, e ensino e

envolvimento com o mundo do trabalho. Em relação aos critérios por regulação externa, incluem-se os convites para ensinar fora da instituição, serviços prestados noutras Universidades, desenvolvimento de trabalhos para outras instituições, etc. A análise de outros desenvolvimentos sobre esta problemática levou Aylett e Gregory (1996) a considerarem a existência de cinco componentes essenciais para determinar um perfil global de caracterização do ensino: i) componente institucional – descreve os factores externos que condicionam o ensino nessa instituição; ii) abordagem ao ensino – inclui todo o processo de preparação e formas de contacto com os alunos; iii) ensino propriamente dito sob a forma do desempenho nas aulas; iv) processo de regulação e de avaliação dos resultados do ensino; e, v) reconhecimento externo, quer na própria instituição, quer fora dela. Por último, Cachapuz (2001) acrescenta algumas actividades possíveis e desejáveis, e cuja apreciação pode ser feita com a necessária dose de objectividade: elaboração de textos didácticos de suporte ao estudo dos alunos (não confundir com sebatas); elaboração de software didáctico; publicação de artigos de inovação sobre o ensino; auto-avaliação do seu ensino (envolvendo ou não os alunos) e em que medida essa informação é equacionada ao serviço da melhoria do ensino; auto-avaliação do seu ensino com base na avaliação externa; participação em colóquios com apresentação de trabalhos; actividades de internacionalização dos cursos (Sócrates, Erasmus, etc.) e desempenho de funções de tutoria, entre outras. Assim, o argumento da dificuldade de avaliação da vertente pedagógica formalmente não previsto no estatuto actual da carreira docente universitária não nos permite alhear da questão tendo em vista a sua necessária revisão.

1.4. Objectivos prosseguidos na avaliação docente

O conceito de avaliação tem sido objecto de entendimentos diversos em diferentes momentos e contextos (Bloom *et al.*, 1971; Green, 1995, in Ferreira, 2003), podendo ter diferentes finalidades consoante o nível em que se situa (Vroeijerstijn, 1995). Apesar da diversidade de definições, elas podem organizar-se em volta de duas tendências principais. De uma forma sucinta, pode dizer-se que funciona, ou como um sistema regulador (avaliação processual formativa ou *on-going*), ou como um sistema de controlo (avaliação final ou de resultados), desde o nível individual ao nível sistémico. Por sua vez, Bonniol e Vial (2001), numa compilação exhaustiva sobre avaliação, sugerem três formas de conceber o campo da avaliação correspondente a três posturas

epistemológicas: avaliação *como medida* (centração nos resultados); avaliação *como gestão* (ênfase nos procedimentos) e avaliação *como problemática do sentido* (ênfase nos processos).

Tradicionalmente, a avaliação dos docentes foi considerada como uma avaliação do tipo sumativo, no qual as finalidades e os objectivos da organização assumiam maior importância do que o desenvolvimento profissional dos docentes. Concebida como medição, descrição, julgamento ou negociação, no caso específico da avaliação dos docentes, como assinala Rosales (1992 in Simões, 2000) “a avaliação tem consistido na recolha de informação e na formulação de juízos sobre o valor ou o mérito do professor”. Esta é uma das razões pela qual os docentes sempre lhe ofereceram resistência e manifestaram o seu descontentamento. Por sua vez, dedicaram-se escassos recursos e pouca atenção aos resultados, não se assumindo a avaliação como um meio para melhorar a qualidade do ensino. Esta era geralmente percebida como um meio de controlar os docentes, de os motivar e de os tornar responsáveis pelo seu trabalho, ou de os afastar quando o seu desempenho era insatisfatório. A imagem que sobressai da avaliação dos docentes é que esta se apresenta como algo contra os mesmos, em vez de se encontrar ao seu serviço (Nevo, 1994 in Simões, 2000).

Embora se verifique a predominância do carácter judicativo (valor ou mérito) da avaliação, vem-se assistindo à emergência de concepções mais favoráveis a uma avaliação marcadamente formativa do desempenho docente, potenciadora do desenvolvimento profissional e, em última instância, da melhoria da qualidade do ensino. Para autores como Guba e Lincoln (1989 in Simões, 2000), trata-se da emergência da 4ª geração da avaliação que, do ponto de vista ontológico, epistemológico e metodológico se distancia de uma visão positivista (existência de uma realidade em si, independente do avaliador) onde passa a destacar-se uma visão mais formativa de natureza construtivista. Assim, neste enquadramento conceptual, sai valorizada a dimensão formativa da avaliação ou, nas palavras de Hadji (1995) a perspectiva de uma aprendizagem assistida pela avaliação. A avaliação passa a ter como objectivos fundamentais: verificar se o(s) dispositivos pedagógico(s) utilizado(s) pelo docente são adequado(s); verificar se e como se realizam as aprendizagens e introduzir alterações ou ajustamentos quando necessário.

Contudo, num domínio que se orienta mais pela controvérsia de posições e ideologias (Aleamoni, 1981 in Marsh, 2001) é interessante constatar uma relativa unanimidade em torno da avaliação do ensino e, em particular da avaliação docente, com objectivos formativos e de diagnóstico. De facto, um estudo conduzido por Fernandez e colaboradores (1996) na Universidade Complutense e na Universidade de Oviedo, concluiu que os docentes tendem a

manifestar uma atitude mais favorável em relação à avaliação quando os resultados são utilizados de modo formativo, ajudando-os a melhorar o seu desempenho. Pelo contrário, revelam sentimentos mais negativos quando os resultados da avaliação são usados sumativamente para a tomada de decisões relacionadas com o trabalho, como por exemplo a obtenção de promoções (Duke, 1995; Haefele, 1993 in Simões, 2000).

Os adeptos da perspectiva formativa ou de desenvolvimento profissional consideram que a maioria dos docentes são competentes e que, por isso, os recursos devem ser utilizados para ajudar os bons docentes, em vez de serem usados para demitir alguns docentes incompetentes. Além disso, a maior parte dos docentes experientes tirará pouco ou nenhum proveito de avaliações sumativas, baseadas na prestação de contas. Esta asserção apoia o princípio de que dificilmente se obtém mudanças na qualidade docente apenas mediante devolução da informação, pelo que a experiência tem mostrado a necessidade de criar um serviço de ajuda profissional⁹ “*What is important is that action is taken to follow-through on results, and put them to use for teaching improvement as means of supporting this process, including discussion with an external consultant, follow-up support from ‘expert’ teachers and the provision of practical strategies and guidelines for effective teaching*” (Marsh & Roche, 1993, 1994; Murray, 1984; Wilson, 1986 in Ballantyne *et al.*, 2000, p.221). Por sua vez, os estabelecimentos de ensino, para além de lugares de trabalho, constituem *comunidades de aprendizagem institucional* (Santiago, 1997), onde a necessidade de garantir um elevado nível de flexibilidade profissional e organizacional exige a implementação de modelos organizacionais orientados para o auto-desenvolvimento. A este respeito, Mintzberg (1995), nos seus estudos acerca das organizações e dos trajectos possíveis para a mudança, refere que numa organização profissional, como são as universidades, a mudança só é possível através de um lento processo de mudança dos próprios docentes.

Assim, o debate sobre as funções da avaliação passou a ter como referência várias nomenclaturas, sucedendo o mesmo com os propósitos da avaliação dos docentes. As denominações mais comuns, e geralmente equacionadas em termos antinómicos, são: avaliação formativa/sumativa, avaliação para o desenvolvimento profissional/prestação de contas, avaliação não controladora/controladora e avaliação profissional/burocrática. Aliás, esta dualidade de

⁹ Especificamente, na Austrália os projectos designados QUT – *Teaching and Learning Plan* (1993-2003) e o *Improving Teaching Project* (Marsh & Roche, 2001), com base no feedback fornecido pelos alunos, procuram em regime de consultoria, ajudar os docentes a encontrar formas alternativas de resolução de problemas pedagógicos. Por sua vez, em Espanha têm sido desenvolvidos cursos de iniciação à docência universitária na Universidade de Valência (Perales, Sánchez & Chiva, 2002).

funções tem sido equacionada de modos distintos nos diversos modelos que têm presidido à avaliação dos docentes. De resto, a história da avaliação dos docentes pode ser descrita em termos de uma tensão entre dois tipos de propósitos: enquanto a avaliação com objectivos formativos se relaciona com o desenvolvimento profissional e a melhoria do desempenho do docente, a avaliação sumativa encontra-se orientada para a tomada de decisões, entre elas a decisão de contratar, seleccionar, promover, reter ou demitir (Haefele, 1993; Cousins, 1995 in Simões, 2000). Aliás, se por um lado existem modelos em que prevalece exclusivamente uma das funções, noutros constata-se a sua coexistência. Todavia, esta coexistência nem sempre tem sido uma tarefa pacífica e de fácil gestão. De acordo com Macedo (2001), existe um certo grau de controvérsia quanto à possibilidade de utilizar um mesmo sistema de avaliação tanto para fins formativos como para fins sumativos. Por sua vez, os defensores da necessidade de dois sistemas argumentam que é impossível utilizar um único sistema para fins formativos e sumativos, dadas as suas diferenças. Por outro lado, os defensores da utilização de um único sistema, sugerem que na prática estes dois sistemas estarão a ser utilizados com sucesso em alguns contextos (*ibidem*). No sentido de contornar esta dicotomia, o desenvolvimento de um dispositivo de avaliação que procure harmonizar a avaliação sumativa com a avaliação formativa parece adequado, e pode considerar-se útil na tomada de decisões válidas e justificadas, permitindo estabelecer procedimentos para a melhoria do ensino, controlando ao mesmo tempo a sua qualidade e as suas implicações na satisfação dos alunos (Cantero, Deus & Abalde, 2002).

Porém, a razão de ser da avaliação não se esgota neste debate, podendo e devendo incluir uma reflexão sobre o que se avalia e o modo como se avalia (funções e fins da avaliação). Por outras palavras, e tomando como referência a proposta conceptual de Hadji (1994), sempre que nos propomos construir um dispositivo de avaliação pertinente, é fundamental clarificar as variáveis de sentido ou axiológicas e as variáveis técnicas. As primeiras prendem-se com questões ligadas ao *porquê* da avaliação, ao *para quê*, ao *para quem* e *para que tipo de decisões*. As segundas relacionam-se com as interrogações relativas ao *quê*, *quem*, *como* e *quando* da avaliação.

Para além das duas funções já mencionadas - regular e controlar - a avaliação pode igualmente ter uma função social e, ao nível sistémico, pode ser usada como forma de mostrar eficácia e adequação do sistema aos padrões dominantes. De qualquer modo, e seja for a função que norteia um sistema de avaliação de docentes, reconhece-se que os mesmos desempenham um papel fundamental na qualidade do sistema educativo (Thélot, 1993 in Simões, 2000).

Um dos aspectos da investigação em torno da avaliação do ensino que continua por resolver é o tipo de informação que deve ser usado para fins sumativos e, por consequência, para a tomada de decisões institucionais (Harrison, Douglas & Burdsal, 2004). Mais de 2000 estudos (Cashin, 1995), apesar de defenderem a natureza multidimensional da avaliação, sugerem que os itens globais dos questionários apresentam pouco valor informativo para a melhoria da qualidade do ensino, enquanto que itens globais tais como a avaliação global do docente, da disciplina e valor das aprendizagens efectuadas providenciam informação útil e suficiente para tomada de decisões institucionais (Cashin & Downey, 1992; Centra, 1979). Contudo, alguns estudos (Marsh, 1984, 1987, 1991; Marsh & Bailey, 1993; Marsh & Hocevar, 1991a) defendem que a avaliação do desempenho docente decorre de constructo de natureza multidimensional, e nesse sentido, importa considerá-la para informar a tomada de decisões institucionais e do próprio docente.

Contudo, outros autores (Abrami & d'Apollonia, 1990, 1991) referem quatro argumentos a favor da utilização dos itens globais para avaliação com fins sumativos: i) dificuldade em definir quais as dimensões representativas do ensino eficaz e como estas se relacionam entre si; ii) a validade empírica pode estar comprometida quando os questionários são aplicados numa diversidade de contextos (disciplinas, docentes e alunos); iii) estudos de Cohen (1981 in Cashin & Downey, 1992) sugerem que as dimensões específicas dos questionários não tendem a correlacionar-se de forma estatisticamente significativa com a aprendizagem dos alunos comparativamente aos itens globais; e, por último, iv) dados da investigação não permitem assumir a capacidade de generalização dessas dimensões/facetadas, o que não acontece no caso dos itens globais. Por sua vez, essas dimensões explicam pouca da variância encontrada nos resultados comparativamente à variância encontrada nos itens globais. Por sua vez, num artigo mais recente Abrami & d'Apollonia (1997) defendem que na tomada de decisões institucionais, no essencial, importa olhar para a distribuição dos resultados, o que tende a simplificar a apresentação, leitura e a compreensão dos resultados obtidos, por exemplo, analisar a percentagem de alunos que consideram o docente bom ou excelente, ou ainda, a percentagem de alunos insatisfeitos com a docência (Bernard & Normand, 1998).

Outro aspecto que também continua por resolver no âmbito da investigação é os estudos que se baseiam no *feedback* que é recolhido entre os semestres lectivos, e que tende a reduzir os efeitos potenciais de mudanças nas práticas docentes. Com efeito, em alguns estudos o intervalo temporal entre o acesso aos resultados e a avaliação seguinte, é em média, entre 4 a 5 semanas. Por sua vez, o facto dos alunos serem mais influenciados pelo que acontece nas primeiras

semanas de aulas (*efeito de primazia*) comparativamente às últimas semanas, parece explicar a existência de mudanças modestas que não chegam a ter expressão na avaliação realizada (Marsh & Dunkin, 1992). Por último, estas avaliações “intermédias” dificilmente poderão ser generalizáveis para efeitos de uma avaliação sumativa e numa lógica mais global de melhoria da qualidade do ensino (Hommediu *et al.*, 1990 in Marsh, 2001).

1.5. Avaliação pelos alunos

A introdução da temática da avaliação da docência no contexto do Ensino Superior é relativamente recente no contexto europeu e latino-americano, embora tenha uma tradição mais antiga no contexto anglo-saxónico. De referir, que este tipo de avaliação teve a sua origem formal nos anos 20, nos Estados Unidos¹⁰, onde hoje se impõe como prática universalmente institucionalizada (Fernández, Mateo & Muñiz, 1996; Marsh, 2001). Assim, a avaliação da qualidade docente no contexto do ensino universitário, inferida com base nas percepções dos alunos, constitui uma das formas possíveis de avaliação com perspectivas de consolidação, em boa parte das Universidades, constituindo mesmo um dos aspectos fundamentais sobre os quais se apoia uma das dimensões de qualidade nas IES (Cantero, Deus & Abalde, 2002; Simões, 2000).

É vasta a investigação sobre o desempenho dos docentes e variada a literatura internacional sobre a sua avaliação e os respectivos modelos em que se apoia o que atesta a ideia de que nenhum outro método de avaliação tenha sido tão profusamente estudado¹¹. Com efeito, existem mais de 2000 estudos referenciados na base de Dados ERIC (Centra, 2003), repartidos por diferentes autores e países¹² (Abrami & d'Apollonia, 1997; Abrami, d'Apollonia & Cohen, 1990; Cashin, 1995; Centra, 1993; Coffey & Gibbs, 2001; Cohen, 1987; Doyle, 1983; Feldman, 1989ab, 1997, 1998; Fernandez, Mateo & Muniz, 1997; García & Congosto, 2000; Marsh & Dunkin, 1992; Marsh & Roche, 1997; Marsh, 1984, 1987, 2001; Mateo & Sanchez, 1993; McKeachie, 1979; Poissant, 1995; Theall & Franklin, 1990).

¹⁰ Guthrie e Remmers iniciaram a tradição da avaliação em 1920.

¹¹ Cashin constatou que até 1998 existiam cerca de 300 artigos e livros publicados acerca da investigação sobre a avaliação do ensino pelos alunos. Seldin (1993) refere que, entre 1973 e 1993, essas avaliações passaram de 29% para cerca de 86%.

A avaliação dos alunos tem sido assumida pelos governos como indicador válido de desempenho (Ramsden, 1991; Simão, 2003ab) e tem como valia constituir-se como uma medida directa de satisfação face à experiência universitária, e, em particular, da qualidade do ensino ministrado (Marsh, 1987; Marsh & Roche, 1994; Ramsden, 1991; Santiago *et al.*, 2001). Assim, a integração de resultados de natureza atitudinal e afectiva veiculados pelas percepções dos alunos, poderá representar um esforço de *empowerment* (Harvey, 2003) tornando as instituições mais responsivas às suas necessidades e garantindo maiores níveis de satisfação (Almeida *et al.*, 2002).

Contudo, sabemos que a avaliação das percepções dos alunos não constitui, nem deve constituir, o único sistema de avaliação do ensino (Millman, 1981), mas tem sido o que goza na actualidade de maior suporte empírico. Apesar das resistências, Murray (1987 in Marsh, 2001, p.10) defende que a avaliação da qualidade docente inferida com base nas percepções dos alunos, quando positivamente assumida, pode cumprir vários objectivos: “(a) *SET's provide useful feedback for diagnosing strengths and weaknesses in teaching effectiveness, (b) feedback can provide the impetus for professional development aimed at improving teaching, (c) the use of SET's in personnel decisions provides a tangible incentive to working to improve teaching, and (d) the use of SET's in tenure decisions means that good teachers are more likely to be retained*”. A este respeito, Marsh (1987, 2001) refere que as avaliações feitas a partir da recolha de informações junto dos alunos são, no seu conjunto, multidimensionais, fiáveis e estáveis; dependem mais do docente do que da disciplina; são relativamente válidas e pouco afectadas pelos enviesamentos potenciais; assumem-se como ferramentas úteis de melhoria da qualidade do ensino quando combinadas com estratégias concretas de intervenção em domínios/áreas sinalizadas como “fragilidades” da actividade docente.

Claro está que o processo de avaliação, seja ele qual for e em que contexto se realize, provoca reacções díspares de ajustamento por parte dos docentes envolvidos. Apesar da abertura para o exercício da prática da avaliação pedagógica, os intervenientes levantam reservas sobre os critérios que presidem a tal tarefa. A aceitação deste tipo de avaliação não significa a sua validade tácita. Bedggood e Pollard (1999) defendem que esta forma de avaliação não está isenta de críticas, nomeadamente o seu carácter unidireccional, quando é sabido que o processo de ensino/aprendizagem é demasiado abrangente. Neste sentido, não pode ser avaliado em toda a sua complexidade por um único sistema de avaliação, menos ainda por um único ou simples questionário.

A par da progressiva generalização dos procedimentos de avaliação do ensino através das percepções dos alunos, os docentes tendem a ver com uma certa desconfiança essa avaliação. Autores como García e Congosto (2000) afirmam que alguns docentes resistem a ser avaliados pelos alunos, argumentando que a influência subjectiva das suas opiniões tende a influenciar negativamente os resultados das avaliações. Mais concretamente, os docentes temem que esta seja enviesada por diferentes factores, como sejam, a dificuldade da disciplina, a classificação obtida, a motivação do estudante e a dimensão da turma. A propósito desta ideia, McKinney (1997, p.3 in Marsh, 1987) refere que *“Instructors may believe that student course/teacher evaluations are unreliable, invalid, biased and affected by many inappropriate factors. In general, research does not support these beliefs”*. De uma forma geral, consideram que os alunos não avaliarão de forma adequada os objectivos de aprendizagem; a relevância das leituras/trabalhos; o conhecimento equilibrado e actualizado das matérias; e o grau de dificuldade da avaliação (cf. Cashin, 1989).

De acordo com Westerman (2002 in Rowley, 2003) existem dúvidas quanto à capacidade dos alunos avaliarem adequadamente a docência, avançando mesmo com a ideia de que os mesmos podem encontrar-se em melhor posição para avaliar o ensino dois a três anos após o término do curso do que imediatamente a seguir (crítica contra uma avaliação feita a *“quente”*). Para alimentar este debate, alguns autores (Simpson & Sigauw, 2000 in Chonko, Tanner & Davis, 2002) referem algumas razões que tendem a justificar o cepticismo instalado em torno destas avaliações: i) o perigo dos administradores, sem experiência na área, utilizarem as avaliações dos alunos como único instrumento para tecer considerações acerca do desempenho docente; ii) os alunos que não trabalham tendem a penalizar os docentes que os desafiam intelectualmente; iii) os alunos podem *“vingar-se”* dos docentes atribuindo uma avaliação negativa; iv) os docentes encaram as avaliações como uma competição de popularidade, e sentem que as mesmas têm pouco valor para avaliar a qualidade do seu conhecimento científico; e v) os docentes sentem que os alunos não possuem conhecimentos suficientemente credíveis para fornecer uma *“radiografia”* objectiva do ensino ministrado.

Num artigo publicado por Spencer e Schmelkin (2002), que procura explorar as opiniões dos alunos acerca da eficácia do ensino, os autores concluíram que, de uma forma genérica, os alunos não têm receio das repercussões da avaliação que fazem dos docentes, ainda que a desejem fazer em anonimato. Queixam-se, contudo, da inexistência de mecanismos de tradução dos resultados obtidos nessas avaliações em acções concretas (Braskamp & Ory, 1994). Na linha

das conclusões do mesmo artigo, infere-se que, sem prejuízo de uma avaliação assumida anonimamente e a sua tradução em informações fiáveis, regista-se a existência de poucos estudos que procuram investigar como os alunos percebem o processo de avaliação, isto é, quais os níveis de consciência e seriedade que colocam na realização desta tarefa. A este propósito, relembremos as palavras de Harvey ao afirmar que *“it is counterproductive to ask students for information then not use it; students become cynical and uncooperative if they think no one really cares about what they think”* (Harvey, 2001 in Vieira *et al.*, 2002, p.86).

1.5.1. Resultados alcançados

A análise de um corpo de investigação empírica desde os anos 70, embora com resultados contraditórios, tem mostrado a pertinência do desenvolvimento de dispositivos de avaliação da qualidade do ensino percebida pelos alunos, com destaque para a avaliação do desempenho pedagógico dos docentes. Por outro lado, as diferenças verificadas em termos dos resultados obtidos nas várias investigações poderão em larga medida decorrer da representatividade das amostras, do balanceamento ou não das mesmas em função de variáveis-chave, como por exemplo, ano e género, área científica do curso, entre outros. Assim, de um modo necessariamente breve, iremos dar conta dos resultados da investigação, procurando distinguir as variáveis do aluno, do docente e do contexto.

Em relação às variáveis do aluno, a avaliação revela-se contraditória quando se parte para a análise das variáveis pessoais. Se, por um lado, encontramos resultados que indicam que as avaliações do ensino tendem a ser influenciadas pelo ano, idade, género, área científica, desempenho académico anterior, métodos de estudo, inteligência, motivação e interesses dos alunos (Coles, 2002), outros estudos indicam que a avaliação da actividade docente não se correlaciona com tais variáveis. No âmbito dos estudos que indicam a influência das variáveis pessoais do aluno, uma dos elementos que parece ter uma relação directamente proporcional com os resultados da avaliação (Marsh & Dunkin, 1992), é a motivação do aluno. Assim, quanto maior for a motivação e o interesse do aluno, mais favorecida é a aprendizagem e mais o docente é eficaz aos olhos dos alunos (Abrami, d’Apollonia & Cohen, 1990; Braskamp, Branderburg & Ory, 1985; Cashin, 1979; Cohen, 1980; Kulik & Mckeachie, 1975; Marsh & Dunkin, 1992; McMillan, Wergin, Forsyth & Brown, 1986; Murray, 1980, 1983). Por outro lado, dados da investigação

mostram que a nota/classificação e as avaliações se associam ao nível de motivação dos alunos, o que leva a concluir que “*the relationship between grades and students satisfaction might viewed as a welcome result of important causal relationship among other variables rather than a simply as evidence of contamination due to grading leniency*” (p.810). Por sua vez, o interesse prévio na disciplina aliado à frequência de disciplinas mais atractivas em resultado do seu esforço e envolvimento (Coles, 2002; Howard & Maxwell, 1980; Marsh, 1987) podem estar na base de avaliações mais favoráveis. Ainda neste domínio, a investigação tem mostrado que as expectativas dos alunos contribuem de forma significativa para a variância dos resultados obtidos. Com efeito, as expectativas, definidas como aquilo que os alunos esperam encontrar e realizar enquanto alunos, podem ser conceptualizadas como uma dimensão não-cognitiva relevante na compreensão do processo de adaptação ao contexto universitário e na qualidade de ensino percebida (Soares & Almeida, 2001ab). Face à importância e mesmo à variabilidade que as expectativas iniciais dos alunos podem assumir no processo ensino/aprendizagem, alguns estudos indicam que os alunos que tendem a apresentar expectativas positivas do docente e da disciplina, tendem a confirmá-las mais tarde - *efeito pigmaleão* (Marsh & Dunkin, 1992; Perry, Abrami, Leventhal & Check, 1979).

Uma segunda variável a nível dos alunos prende-se com o sexo/género de pertença. Os estudos sugerem que o género dos alunos não parece correlacionar-se com os resultados obtidos nas avaliações do ensino (cf. Feldman, 1983), admitindo-se, porém, que os mesmos tendem a avaliar mais favoravelmente docentes do mesmo sexo (Basow & Silberg, 1987; Bennett, 1982; Bernard, Keefauver, Elsworth & Maylor, 1981; Fernández & Mateo, 1997; Ory, 2001). Adicionalmente, outras investigações indicam que os alunos do sexo feminino avaliam os docentes de forma mais positiva na dimensão que se reporta à sensibilidade face a casos problemáticos (relação individual) do que os alunos do sexo masculino. Por sua vez, os alunos das Ciências tendem a avaliar menos positivamente os docentes comparativamente aos alunos das Humanidades e das Ciências Sociais (Cramer & Alexitch, 2000), não sendo tais diferenças muito significativas, embora consistentes (Cashin, 1990; 1992; Centra, 1993; Ory, 2001). Deberg e Wilson (1990) defendem que esta diferenciação pode ser afectada pela percepção de maior grau de dificuldade associada às disciplinas das áreas das Ciências. Contudo, outros estudos (Rego, 2001) indicam que não foram detectadas diferenças consideráveis entre as áreas de humanidades e de ciências, sendo que o único dado digno de explicitação é o que sugere que a facilitação da comunicação é mais ponderada na área de ciências que nas humanidades. Pode presumir-se que

os docentes destas matérias traduzam a complexidade e/ou cientificidade subjacente em linguagem simples, atractiva e assente em exemplos práticos.

Olhando à tipologia de aula, as aulas práticas tendem a receber melhores avaliações pelos alunos do que as aulas teóricas (Cashin, 1992; Feldman, 1984; Marsh & Dunkin, 1992), ao mesmo tempo, e comparativamente às disciplinas obrigatórias, as disciplinas opcionais, para as quais existe à partida um interesse prévio por parte dos alunos, influenciam de forma positiva a avaliação atribuída aos docentes (Costin, Greenough & Menges, 1971; Feldman, 1978; Kulik & Mckeachie, 1975; Marsh, 1984; Marsh & Dunkin, 1992; Mckeachie, 1979).

Finalmente, os alunos que frequentam disciplinas de anos avançados tendem a atribuir melhores avaliações comparativamente às disciplinas dos primeiros anos, provavelmente pelo facto dos alunos estarem mais motivados nos seus estudos e discriminarem, com maior conhecimento de causa, o processo de ensino/aprendizagem (Langbein, 1994). Autores como Aleamoni e Hexner (1980), e mais recentemente Goldberg e Callahan (1991) ou Langbein (1994), citam duas dezenas de artigos que suportam esta relação de causa e efeito. Outros estudos, no entanto, não têm encontrado correspondência entre o ano do curso e as avaliações dos alunos (Braskamp, Branderburg & Ory, 1985; Centra, 1979; McMillan, Wergin, Forsyth & Brown, 1986; Marsh & Dunkin, 1992).

Os questionários de avaliação do desempenho docente tendem a estabelecer correlações moderadas com o rendimento académico dos alunos (Marsh & Dunkin, 1992). Por exemplo, a literatura tem veiculado a ideia de que os alunos que avaliam de uma forma mais positiva os docentes, tendem a apresentar melhores resultados escolares e uma maior probabilidade de manifestação de interesse e de envolvimento nas matérias das disciplinas (Abrami, d' Apollonia & Cohen, 1990; Braskamp, Branderburg & Ory, 1984; Cohen, 1981; Kulik & Mckeachie, 1975; Marsh & Dunkin, 1992; McMillan, Wergin, Forsyth & Brown, 1986). Não constituirá novidade a existência de uma associação forte entre a expectativa de obtenção de notas/classificações favoráveis e os resultados da avaliação dos docentes, o que em última instância constitui o reflexo da qualidade das aprendizagens. Em acréscimo, McKeachie (1997) defende que os alunos tendem a aprender mais em disciplinas asseguradas por bons docentes. Por sua vez, num estudo levado a cabo por Marsh (1980) concluiu-se que o interesse prévio dos alunos apresenta um impacto muito significativo no conjunto dos resultados obtidos nas avaliações, contribuindo mesmo para cerca de um terço da relação entre a nota esperada e as avaliações nos questionários. Por último, importa ainda fazer referência ao impacto que as atribuições causais podem ter na compreensão destes

fenómenos educacionais, ao indicar que os alunos tendem a atribuir melhores classificações à inteligência e ao esforço enquanto que as notas mais baixas são atribuídas a um ensino pobre. Por se tratar de uma área crítica e susceptível de manipulações, alguns docentes acreditam que a facilidade de obtenção de notas mais elevadas tende a favorecer docentes menos eficazes do ponto de vista científico e dos seus métodos de ensino em favor da sua generosidade e simpatia (*efeito Dr. Fox*¹³). Contudo, este efeito não tem um impacto globalmente significativo, à excepção da dimensão do entusiasmo docente (Marsh, 2001). As críticas mais directas ao mesmo relacionam-se com o facto de que, globalmente, ele é irrelevante para a investigação em torno das avaliações. Primeiro, porque apenas foi leccionada uma aula antes da avaliação, e tal como é sabido, a avaliação baseia-se num volume significativo de aulas por semestre/ano. Por outro lado, os assuntos abordados eram completamente novos para o grupo de alunos, contudo, sabemos que os alunos não ignoram completamente os assuntos leccionados nas aulas. Para corroborar a desvalorização deste efeito no conjunto das avaliações dos alunos, Rego (2003) conclui que um docente apenas cortês, que é simpático, mas não denota orientação prática e não fomenta a participação, acaba por não ser bem cotado do ponto de vista do seu desempenho.

De acrescentar que alguns estudos experimentais procuraram investigar o impacto do *efeito de leniência* junto dos alunos. Por exemplo, num estudo procurou-se atribuir notas erradas/incongruentes com o desempenho dos alunos e verificar as avaliações atribuídas pelos docentes aos seus alunos. Marsh e Roche (1997) colocaram em causa este trabalho pela violação de princípios éticos ao contrariar as expectativas dos alunos. Por exemplo, uma das críticas a estes estudos é a de que não foi tida em devida atenção a factores tais como a área disciplinar, quando é demonstrado pela investigação, que as notas/classificações e as avaliações do ensino pelos alunos são, em larga medida, influenciadas pela área científica em que a disciplina se inscreve (cf. Cashin, 1990; Centra, 1993; Feldman, 1978). Se partirmos do pressuposto da existência de uma correlação entre a nota esperada e o desempenho do aluno, então a atribuição de notas/classificações falsas poderá constituir uma prática contraproducente, à semelhança do que se veio a constatar num estudo realizado por Abrami, Dickens, Perry e Leventhal (1980).

Em suma, os estudos de natureza experimental ou correlacional não conseguem suportar empiricamente que as avaliações dos alunos são contaminadas pelas notas/classificações que

¹³ Definido como influência excessiva da expressividade docente e a possibilidade deste comportamento induzir à produção de avaliações favoráveis. Um indivíduo foi treinado para simular o papel de docente a leccionar uma aula a um grupo de estudantes de medicina. Nessa aula, o “falso docente” apresentou informações incorrectas bem como referências bibliográficas inexistentes, tendo contudo, recebido avaliações favoráveis dos alunos (Naftulin, Ware & Donnelly, 1973 in Kulik, 2001).

recebem. De uma forma genérica, os investigadores estão de acordo que a associação entre a nota esperada e a natureza dos resultados obtidos nas avaliações é mínima (Howard & Maxwell, 1980): “*relatively unimportant when ratings are used to make gross distinctions between teachers*” (Abrami, Dickens, Perry & Leventhal, 1980, p.107) e “*relatively small and inconsistent*” (Centra, 1993, p.74). Estudos levados a cabo por Marsh (1984) sugerem que, de uma forma geral, os docentes que pedem mais trabalho aos alunos são aqueles que apresentam melhores resultados. Contudo, num artigo mais recente publicado em 2001, o mesmo autor aponta para a necessidade de se avaliar a utilidade do trabalho proposto, sugerindo que a aprendizagem pode ser promovida e estimulada aumentando a percepção de trabalho útil por parte dos alunos.

Outro estudo levado a cabo por Centra (2003) concluiu que os alunos que esperam obter notas/classificações elevadas nas disciplinas das áreas das Ciências Naturais tendem a avaliar menos positivamente a qualidade do ensino. Por sua vez, as disciplinas que apresentam extremos de dificuldade ou de facilidade tendem a ser avaliadas mais desfavoravelmente, e, por isso, as disciplinas que apresentam níveis adequados de dificuldade “*just right*” tendem a receber avaliações mais favoráveis.

Considerando agora algumas variáveis associadas aos docentes, a investigação não tem encontrado correlações significativas entre o género do docente e os resultados obtidos nas avaliações (Bennett, 1982; Ferber & Huber, 1975; Lombardo & Tocci, 1979; Strathan, Richardson & Cook, 1991; Wilson & Doyle, 1976 in Ory, 2001). Numa revisão da literatura neste domínio, Feldman (1993) encontrou correlações de .02 entre o género e as avaliações globais, sendo que os docentes do sexo feminino tendiam a receber avaliações mais favoráveis. Contudo, a investigação é inconclusiva. Alguns estudos indicam que os docentes do sexo feminino tendem a apresentar valores mais baixos quando comparados com os docentes do sexo masculino (Andersen & Miller, 1997), sugerindo-se que avaliações mais favoráveis tendem a ocorrer quando o género dos alunos coincide com o género do docente. Outro estudo realizado em duas universidades canadianas com uma amostra de 292 alunos da área das Ciências Económicas, concluiu que os docentes do género masculino, classificados com scores baixos de feminilidade (e scores pouco elevados de masculinidade) e os docentes do género feminino classificados com scores neutros (isto é, indiferenciados em termos do papel de género) tendem a ser percebidos como melhores docentes pelos alunos. Em geral, a masculinidade tende a ser apreciada pelos alunos mais velhos e pelos alunos com estatuto de trabalhador-estudante¹⁴ (Das & Das, 2001).

¹⁴ *Bem's Sex Role Inventory* (BSRI; Bem, 1981 in Das & Das, 2001).

De acordo com outros estudos (Feldman, 1983), é possível concluir pela existência de uma correlação positiva entre o grau académico, anos de experiência e a avaliação global do desempenho docente, sendo que este tende a melhorar com a progressão na carreira. De acordo com o mesmo autor, o docente tende a receber uma avaliação mais fraca no seu primeiro ano de leccionação, aumentando progressivamente nos cinco anos consecutivos, estabilizando entre os sete e os doze anos de carreira. Outros estudos têm mostrado uma associação entre as características de personalidade do docente (ex. energia, entusiasmo, auto-estima positiva) e os resultados obtidos nas avaliações dos alunos, sendo essa associação menos clara em relação ao rendimento académico (Abrami *et al.*, 1982 in Emery, Kramer & Tian, 2003).

A avaliação do docente tende a manter-se estável em termos temporais o que corrobora a existência de um perfil atemporal consistente. Esta ideia decorre de um estudo de natureza longitudinal conduzido por Marsh e Hocevar (1991b), e que procurava avaliar as mudanças do desempenho pedagógico percebido pelos alunos tomando uma amostra de 195 docentes ao longo de 13 anos. O estudo concluiu pela existência de um padrão consistente de avaliação das percepções de competência pedagógica percebido pelos alunos.

A par de variáveis pessoais dos alunos e dos docentes, podemos, por último, associar as avaliações da qualidade do ensino às características dos contextos onde ocorrem. Por exemplo, os alunos de Departamentos percebidos como oferecendo “ensino de maior qualidade” tendem a envolver-se mais em abordagens profundas na sua aprendizagem bem como a perceberem os conteúdos escolares como relevantes e úteis (Almeida, 2002). Por outro lado, a avaliação dos docentes pode ser influenciada pelo tamanho da turma, ao sugerir que turmas mais pequenas tendem a atribuir melhores avaliações comparativamente às turmas numerosas (Koh & Tan, 1997; Liaw & Goh, 2003). Numa revisão de mais de 52 estudos feitos em várias turmas com tamanhos distintos, Feldman (1978) constatou a existência de correlações de .09 entre o tamanho da turma e as avaliações. Em 1992, Cashin conclui que *“Taken alone, class size is not a serious source of bias”* (p.2) e um ano mais tarde, Centra (1993, p.102) refere que as diferenças encontradas nas avaliações devido ao tamanho das turmas *“have little practical significance”*. Por outro lado, um estudo recente levado a cabo por Chiu (1999 in Ory, 2001) conclui que apenas 1% da variância encontrada nos resultados é explicada pelo tamanho da turma. Contudo, nos estudos em torno do SEEQ (Marsh, 2001), apenas as dimensões de interação da turma e da relação docente/aluno tendem a ser afectadas pelo tamanho da turma, verificando-se resultados semelhantes quer nas avaliações dos alunos quer nas auto-avaliações dos docentes. Face à natureza dos resultados

encontrados, existem motivos para supor que estas dimensões do desempenho tendem a ser mais positivamente avaliadas em presença de turmas mais pequenas.

Por último, algumas variáveis de natureza mais administrativa poderão concorrer para a produção de avaliações mais favoráveis tais como a possibilidade do questionário não ser anónimo, eliminando por isso preocupações/receios de eventuais represálias (Ory, 2001), a presença do docente na sala de aula e a explicitação do objectivo que se pretende atingir com a administração do questionário (Cashin, 1990).

1.6. Formas alternativas de avaliação

Seguindo o princípio da diversidade, qualquer dispositivo de avaliação deve ter na sua base fontes e instrumentos de avaliação diversificados que permitam não só obter uma compreensão parcial dos dados, como também conferir a possibilidade de triangular e confrontar as diversas informações (Abrantes & Valente, 2000; Curado, 2000). A qualidade do ensino é um conceito multifacetado e, quando muito, o questionário apreende uma parte do todo que o define. Temos consciência que para avaliar e intervir na qualidade do ensino/aprendizagem não basta avaliar o docente, devendo-se incluir também aspectos relativos à disciplina, ao curso e à instituição. Por exemplo, em Espanha começou-se a introduzir a denominada avaliação *circular* (Fernández & Mateo, 1994), nas Universidades de Oviedo e Complutense de Madrid (Fernández, Mateo & Muñiz, 1995) baseada na ideia de que todos os agentes que participam no contexto educativo têm o direito, e o dever, de se avaliarem reciprocamente (Fernández, 1992). Desta forma, procura-se abarcar as diferentes componentes do processo de ensino/aprendizagem, a par do cruzamento das informações provenientes dos diferentes agentes com o objectivo de promover a riqueza interpretativa dos resultados obtidos, assumindo que “*while instruments designed to measure student evaluations have proliferated in recent years, there is little evidence that using such instruments in isolation has any effect on improving teaching performance*” (Dunkin & Barnes, 1986; Seldin, 1989 in Ballantyne *et al.*, 2000). Por outro lado, a defesa da pluralidade informativa implica, ainda, a abertura do sistema de avaliação da qualidade do ensino à sua própria evolução, na medida em que a mudança da cultura pedagógica da instituição exige a adaptação sistemática das metodologias utilizadas para a sua avaliação. Se tradicionalmente a aprendizagem dos alunos foi considerada o objecto primogénito da avaliação, no presente, e de um modo mais sistemático,

a avaliação tem vindo a incidir sobre outras vertentes da realidade educativa, das quais se salientam as reformas, os currículos, os projectos, as metodologias e os docentes. Esta avaliação recorre aos pares, aos administradores e aos observadores experientes/externos. No âmbito da avaliação do ensino tomando os ambientes educativos, autores como Fraser (1991 in Soares & Almeida, 2002) propõem a existência de três abordagens principais: (i) observação sistemática; (ii) estudo de casos; e (iii) avaliação das percepções de alunos e dos docentes.

De acordo com Fraser (1991), a abordagem perceptiva apresenta algumas vantagens relativamente às restantes metodologias acima mencionadas. No essencial, i) sendo uma avaliação de papel-hápis é mais económica do que o recurso a observações de sala de aula que envolvem o recurso a observadores especializados e experientes; ii) as medidas perceptivas baseiam-se na avaliação da experiência efectiva dos alunos num determinado ambiente enquanto os dados resultantes da observação habitualmente restringem-se a um período de tempo determinado; iii) as medidas perceptivas envolvem um conjunto de julgamentos de todos os alunos da turma enquanto que as técnicas de observação tipicamente envolvem um único observador. Quanto às percepções dos alunos, uma vez que são elas as determinantes dos seus comportamentos mais do que as situações reais, podem assumir-se como mais relevantes do que os próprios comportamentos observáveis. Por último, as medidas perceptivas do ambiente de sala de aula conseguem explicar uma boa parte da variância (Fraser & Walberg, 1991 in Soares & Almeida, 2002). Aspecto que não deve ser alheio a essa constatação, é o facto das percepções dos alunos serem altamente influenciadoras uma vez que elas são a base a partir da qual os alunos interpretam e respondem às experiências em sala de aula. Contudo, como veremos adiante, existem outras formas igualmente importantes para a avaliação da qualidade do ensino.

1.6.1. Auto-avaliação do docente

No essencial, a utilidade deste tipo de avaliação baseia-se na possibilidade de uma auto-reflexão mais substantiva por parte do docente sobre a sua actuação. Por sua vez, a complexidade da actividade docente implica o desenvolvimento sistemático de um processo metacognitivo (avaliação de atitudes e conhecimentos) e de um acto pedagógico de auto-regulação dessa actividade, isto é, planificar, organizar, intervir e avaliar (Tremblay, 1998). Um dos exemplos de questionário disponíveis na literatura é o *Staff Member Self Rating Survey* (SMSRS; Marsh, 1987,

1991) que procura constituir o contraponto da avaliação por questionário que é também utilizado pelos alunos, no sentido de testar a validade convergente e divergente das percepções dos actores envolvidos no processo de ensino/aprendizagem (alunos e docente). Apesar deste tipo de avaliação configurar uma avaliação eficaz da qualidade do ensino, apresenta algumas limitações na sua aplicação. Por exemplo, numa meta-análise de 19 estudos acerca de correlações entre os resultados obtidos nos questionários dos alunos e a auto-avaliação do docente, Feldman (1989b) encontrou uma média de .29 para as avaliações globais e entre .15 a .42 para as dimensões/factores nesses questionários. Por sua vez, Marsh (1987) levou a cabo 2 estudos tendo solicitado aos docentes para fazerem a sua auto-avaliação com base no mesmo instrumento que foi utilizado pelos alunos. Nos dois estudos concluiu que: i) a estrutura factorial do questionário aplicado aos alunos e docentes é idêntica; ii) o grau de concordância entre aluno/docente em cada dimensão do questionário é significativo (média de correlação entre .49 e .45) embora este valor tenha sido superior comparativamente à avaliação global do ensino ($r = .32$); iii) na generalidade dos casos, as diferenças de médias entre os alunos e os docentes revelaram-se pouco significativas e assistemáticas; e iv) a análise multimétodo - multitraço fornece suporte para a validade convergente e discriminante das avaliações. Em suma, e numa lógica de procurar harmonizar as diferentes perspectivas de avaliação, importa desenvolver dispositivos de avaliação que incluam a auto-avaliação docente no sentido de superar uma das limitações mais apontadas nos vários estudos e que se relaciona com o carácter unilateral da avaliação do ensino.

1.6.2. Avaliação pelos pares

A avaliação por pares – colegas, administradores – que não se baseia em visitas às salas de aula, parece correlacionar-se com os resultados obtidos nos questionários de avaliação do ensino, mas provavelmente esta avaliação pelos pares tende a ser influenciada pela informação veiculada pelos alunos (Marsh, 1987; Marsh & Dunkin, 1992; Marsh & Roche, 1994). Por outro lado, as avaliações de colegas e administradores que se baseiam em visitas às salas de aula não parecem ser suficientemente fiáveis (isto é, as avaliações entre pares nem sempre são consensuais) e, de uma forma sistemática, não tendem a correlacionar-se com os resultados dos questionários de avaliação do ensino ou com outros indicadores do ensino eficaz (Centra, 1979; Marsh, 1987; Murray, 1980; Scriven, 1983). Contudo, e de acordo com Poissant (1996), a

avaliação feita por pares é essencial na determinação da qualidade do trabalho do docente (Lachapelle, 1998), adiantando mesmo que os pares estão particularmente qualificados quando se trata de avaliar os materiais de ensino. Estas apreciações têm normalmente por base a observação directa de aulas ou através de vídeo. No entanto, a observação de aulas levanta alguma controvérsia. Por exemplo, os pares não treinados têm tendência para avaliar mais positivamente o docente comparativamente com os alunos (*efeito de leniência*). Outros factores são igualmente apontados como limitadores: amizade muito próxima entre os pares e o avaliado, número reduzido de pares juizes, conhecimento insuficiente da disciplina leccionada pelo avaliado, competição com o avaliado e falta de anonimato dos pares (Abrantes & Valente, 2000). Face às considerações feitas, não parece haver evidência quer para a aceitação incondicional quer para a refutação liminar deste tipo de avaliação, contudo, a evidência empírica sugere que as avaliações de colegas e administradores baseadas em visitas às salas de aula não constituem indicadores suficientemente válidos da avaliação da qualidade do ensino (Murray, 1980).

1.6.3. Avaliadores externos

Para obviar algumas das limitações antes referidas, é apontado o recurso a avaliadores externos (Murray, 1980). Cranton e Hillgarten (1981) procuraram avaliar a relação entre os resultados dos questionários de avaliação e os comportamentos específicos observados em sala de aula. Estes estudos concluem que, na dimensão interacção da turma, valores superiores tendem a ocorrer quando os docentes valorizam o desempenho dos seus alunos, através do questionamento e da clarificação de dúvidas. Por seu lado, e na dimensão organização/clareza, valores superiores ocorrem quando os docentes passam mais tempo a organizar, a estruturar as aulas e a explicar a relação entre as matérias. Num estudo mais ambicioso, Murray (1983) treinou observadores para estimar a ocorrência de comportamentos específicos em 54 docentes que tinham previamente recebido avaliações baixas, médias e altas nos questionários. Os resultados obtidos indicam que os docentes avaliados de forma distinta pelos alunos, também o foram quanto à manifestação de comportamentos sinalizados pelos observadores externos.

Apesar da investigação ser limitada neste domínio, alguns estudos indicam a pertinência de incluir as dimensões de clareza e expressividade dos docentes nos questionários de avaliação (Marsh, 1987; Marsh & Roche, 1997), por constituírem comportamentos facilmente avaliados por

alunos e observadores externos. Os mesmos autores sugerem também que tais comportamentos são componentes fundamentais da qualidade do desempenho docente, gerando assim o necessário consenso entre alunos e docentes. De salientar que estes padrões comportamentais tendem a associar-se ao rendimento académico dos alunos, quer em contextos naturalistas quer em contextos experimentais. Por sua vez, o ensino e a investigação são vistos como produtos da qualidade universitária, contudo os resultados das investigações mostram que estas duas áreas não tendem a correlacionar-se, e que apesar de não haver suporte quer para a sua aceitação ou refutação, os indicadores de produção científica não devem ser utilizados como indicadores da eficácia do ensino (Aleamoni, 1999).

1.7. Práticas de avaliação do ensino e dos docentes universitários em diferentes países

A partir da caracterização das experiências no campo da avaliação dos docentes no Ensino Superior e do estudo de casos de experiências mais significativas em termos internacionais (Abrantes & Valente, 2000) e dos resultados do projecto EVALUE (*Evaluation and Self Evaluation of Universities in Europe*)¹⁵, procura-se caracterizar o *estado de arte* dessa avaliação. No que se refere à avaliação dos docentes, é obvio que, para além da avaliação que se integra no processo de avaliação das Universidades, existe em todos os países uma avaliação que se liga à carreira e se processa no momento de recrutamento e em diferentes etapas da carreira profissional dos docentes (veja-se no Estatuto da Carreira Docente Universitária, ECDU). Os critérios estão essencialmente ligados aos graus académicos e à investigação, mas na progressão na carreira observa-se a emergência de outros processos de avaliação ligados à competência pedagógica, por exemplo, os *portfólios* sobre as actividades de ensino na Finlândia, ou a apresentação de uma aula perante um grupo de professores em Itália ou diante de alunos em algumas universidades na Alemanha.

Na Inglaterra, a avaliação é da responsabilidade de organismos nacionais e centrais (e.g. *Quality Assurance Agency*) que tem como objectivo a avaliação dos procedimentos da Universidade

¹⁵ Projecto EVALUE (*Evaluation and Self Evaluation of Universities in Europe*), subsidiado pela União Europeia e realizado no âmbito do TSER (*Targeted Socio Economic Research Program*) teve como principal objectivo um estudo comparativo dos processos de avaliação em curso nas Universidades Públicas de oito países europeus - Alemanha, Espanha, Finlândia, França, Grã-Bretanha, Itália, Noruega e Portugal, entre 1996 e 1998 (Estrela & Simão, 2003).

para assegurar a qualidade e a avaliação da qualidade do ensino (Harvey, 2003; Cowan, 2004) e o *Higher Educational Funding Council for England* que avalia a qualidade da investigação, determinando a distribuição dos seus fundos. Apesar do elevado grau de autonomia que caracteriza o sistema superior Inglês, foi criado em 1997, o *Institute of Learning and Teaching* com o objectivo de fornecer apoio aos docentes/instituições em matéria de educação. Assim, no Reino Unido a avaliação dos Departamentos é feita por comissões de pares, dentro do *University Funding Council* e dos *Research Councils*, e tem uma profunda relação com os mecanismos de financiamento das Instituições (Amaral, 1998). No *King's College*, o dispositivo de avaliação do ensino consiste num relatório anual elaborado por cada docente. Este relatório contém uma análise crítica das actividades realizadas no último ano, o qual servirá de base numa reunião com o Director de Departamento para uma avaliação do desempenho do docente¹⁶. Nessa entrevista é discutido e avaliado, em conjunto, o desempenho do docente nas diversas actividades e traçado o diagnóstico das necessidades de formação para o seu desenvolvimento profissional. Neste caso, os processos de avaliação são monitorizados pelo *Steering Commite* no qual participam membros da *Association of University Teachers*. Adicionalmente à avaliação docente é feita a avaliação dos alunos através do preenchimento de um questionário anónimo, organizado num conjunto de 10 afirmações que procuram descrever comportamentos-chave para um bom desempenho do docente: ser audível, estimular espírito crítico, prestar apoio fora das aulas, etc. Os resultados são entregues ao Director de Departamento que os comunica ao respectivo docente. Trata-se de uma avaliação obrigatória para todos os docentes, com excepção daqueles que se reformam no prazo de um ano e dos que têm contratos temporários. Os docentes que exercem a sua actividade apenas há um ano apresentam um relatório centrado na recolha de informação e na discussão sobre as finalidades e objectivos para o ano seguinte.

Assim, no Reino Unido existem cinco formas em que se pode realizar a avaliação do ensino. A primeira, e a única que se relaciona directamente e em exclusivo com os docentes, é a confirmação ou negação da sua competência. No caso de queixas pertinentes dos alunos e/ou colegas, estas serão transmitidas ao docente e é dada a este a oportunidade de melhorar o seu desempenho. No caso da sua manutenção, é formado um grupo de consultores de competência reconhecida que avalia o trabalho do docente. O despedimento de um docente só se concretizará

¹⁶ O relatório de actividade contém elementos relativos a cinco dimensões: ensino, actividades administrativas, investigação, publicações, prémios e outras actividades.

se for provada, sem qualquer sombra de dúvida, a sua incompetência, não sendo de todo frequente.

A segunda forma de avaliação dos docentes é realizada no decurso da sua formação como docentes. Actualmente, a maioria das Universidades Britânicas requerem que os novos docentes recebam formação pedagógica e em educação de adultos, por vezes em termos de um curso de pós-graduação. Em qualquer dos casos, o docente é observado e avaliado formativamente e, no final sumativamente, quanto a um satisfatório desempenho docente, disso dependendo a sua contratação definitiva.

A revisão de módulos e programas é a terceira forma de avaliação. Esta revisão é realizada periodicamente pelas próprias instituições com participação externa. Desta avaliação constam amostras de questionários preenchidos pelos alunos, actas de reuniões de comissões de alunos, actas de reuniões da equipa encarregada da revisão, relatórios de examinadores externos, etc. A quarta forma de avaliação é a revisão das estratégias de aprendizagem/ensino pela instituição. Por fim, a quinta decorre da auditoria que é conduzida por pares (académicos seniores com cargos de responsabilidade) que são orientados pela QAA (*Quality Assurance Agency*). A auditoria académica iniciada há quinze anos, apesar de ser muito criticada, especialmente pelo volume de trabalho que exige às instituições, tem sido um factor extremamente importante na melhoria da qualidade do ensino superior britânico. Por sua vez, a prática da auditoria, entre outras, convenceu muitos académicos de que as práticas incorrectas e baixos padrões de qualidade têm vindo a ser identificados e eliminados (Cowan, 2004).

No caso francês, a avaliação das Universidades e da docência está cometida ao *Comité National d'Evaluation* (CNE). Trata-se de um organismo independente, directamente ligado à Presidência da República e não ao Governo (o que o diferencia do modelo inglês). A situação em França parece assegurar uma melhor separação entre o financiamento e a avaliação da qualidade. No sistema francês, os resultados da avaliação estão menos ligados ao financiamento e são aplicados a cada Universidade de forma variada. Por exemplo, em alguns casos, uma má avaliação da qualidade pode, por vezes, dar origem a um reforço do financiamento. Além disso, neste sistema, a avaliação não se faz prioritariamente por áreas científicas, mas por instituições.

O sistema holandês destaca-se pela sua popularidade na Europa, tendo sido seguido com algumas adaptações nos países nórdicos e, também, em Portugal. É de facto o sistema mais difundido na Europa e tem uma característica fundamental que é assumir-se como um sistema essencialmente vocacionado para a melhoria da qualidade (Amaral, 1998). A particularidade deste

modelo é o facto de associar a auto-avaliação com a avaliação externa, cujos resultados configuram perfis reputacionais das Universidades que, de alguma forma, se repercutem na sua capacidade de obtenção de recursos. A avaliação é conduzida pela Associação das Universidades Holandesas (VSNU) que abrange as 14 Universidades do País, sendo que os processos de avaliação dos cursos e da actividade de investigação estão totalmente separados. Por exemplo, na Universidade de Amesterdão, a avaliação do desempenho pedagógico dos docentes surgiu na década de 90 como que em resposta a uma pressão por parte do corpo discente, tendo sido adoptado um procedimento geral para toda a Universidade consistindo no preenchimento pelos alunos, no final do período de leccionação de cada disciplina, de um questionário anónimo (*Uvalon*)¹⁷. Por outro lado, parece visível uma preocupação generalizada com a realidade do trabalho pedagógico apontando para o interesse do desenvolvimento de um *portfólio* que contenha materiais pedagógicos produzidos pelos docentes, assim como a análise das suas práticas pedagógicas.

Embora com uma tradição menos longa, importa destacar a experiência das Universidades Canadianas. Nestas, os alunos são a fonte informativa mais consultada, embora não única. Por exemplo, desde 1989 que a nível institucional a Universidade de Laval, assume uma política de avaliação do ensino decorrente das experiências anteriores e dos recentes desenvolvimentos da investigação no contexto universitário. Esta política segue uma abordagem holística do ensino, que procura implicar todos os intervenientes na formação dos alunos e decorre da necessidade de assumir as suas responsabilidades face à comunidade educativa e da sociedade em geral. Assim, a formação de qualidade integra em boa medida, o trabalho pedagógico dos seus docentes como forma de informar as diferentes categorias da administração de modo a que possam apreciar os esforços e inovações dos seus docentes. A implementação desta política levou à criação, em 1995, do *Bureau d'Animation et de Soutien de l'Enseignement* (BASE) que tem como principais funções a concepção, tratamento e divulgação dos resultados da avaliação, apoio/consultadoria a docentes e alunos na resolução dos obstáculos ou dificuldades emergentes da própria avaliação e, ainda, a formação contínua dos docentes. A avaliação administrativa é baseada na aplicação de um questionário a preencher pelos alunos organizado em três partes: caracterização da disciplina e do estudante, organização e funcionamento pedagógico da disciplina, e recolha de

¹⁷ Trata-se de um instrumento elaborado ao nível da Universidade que contém um conjunto limitado de questões (cerca de 10) às quais são acrescentadas mais algumas dirigidas a aspectos específicos de cada disciplina. Este questionário tem como objectivo conhecer o balanço que os alunos fazem do programa da disciplina, dos métodos de ensino e dos processos de avaliação.

comentários/sugestões sobre a disciplina, o docente e o plano de estudos. A Reitoria recebe os resultados globais do conjunto das disciplinas avaliadas, identificando pontos fortes e fracos, e apresentando decisões com vista à melhoria da qualidade do ensino.

Na Alemanha, as Universidades encontram-se sobre a responsabilidade dos “*Länder*”. Alguns deles introduziram um sistema central e regular da avaliação, mas outros deram plena autonomia às Universidades para procederem à avaliação. De salientar que, de uma forma genérica, a avaliação do ensino, da investigação e da administração é feita separadamente (Estrela & Simão, 2003).

Na Finlândia e na Noruega, onde as Universidades têm menos autonomia, os organismos da tutela procedem à sua avaliação, criando organismos nacionais externos com o objectivo de conduzir as Universidades a um processo de mudança. Por último, o Conselho de Avaliação da Fundação das Universidades em Portugal, o Conselho das Universidades em Espanha e o Observatório de Avaliação em Itália constituem exemplos de organismos, criados por iniciativa das Universidades, que conduzem a avaliação externa da Universidade.

1.7.1 Balanço da situação portuguesa

Em Portugal, a necessidade de avaliação do sistema de Ensino Superior começou a ser reconhecida formalmente em 1986 até ser consagrada na própria Constituição (revisão de 1997). Nesse período, as Universidades Públicas e a Universidade Católica, através do Conselho de Reitores das Universidades Portuguesas (CRUP), decidiram iniciar um debate público sobre a matéria e promoveram um processo de auto-avaliação, no âmbito de uma experiência-piloto acolhida posteriormente na Lei da Avaliação do Ensino Superior (Lei n.º 38/94, de 21 de Novembro), onde está definida a política de avaliação e acompanhamento das instituições do Ensino Superior criando para esse efeito a Fundação das Universidades Portuguesas (FUP).

A implementação da lei da avaliação, com o seu alargamento a todos os subsistemas, ocorreu com a constituição, em 1996, do Grupo de Reflexão e Acompanhamento do Processo de Avaliação das Instituições do Ensino Superior. Este Grupo foi incumbido de propor as regras e os princípios gerais tendentes a definir e assegurar a concretização do sistema de avaliação. Coube-lhe, também, a elaboração do guião para a auto-avaliação, de um primeiro documento estratégico

e de um projecto de diploma legal que está na origem do Decreto-Lei n.º 205/98, de 11 de Julho, o qual criou o Conselho Nacional de Avaliação do Ensino Superior (CNAVES)¹⁸.

De acordo com as propostas da União Europeia, Portugal viria a adoptar a nível nacional o “modelo geral” de avaliação (baseado nas revisões elaboradas por Van Vught & Westerheijden, 1993). Muitas vezes designado por “modelo holandês”, estabelece um equilíbrio dinâmico entre a avaliação interna e avaliação externa (Simão, 2003a). Este modelo assenta em quatro elementos principais: (i) a existência de uma *Agência Nacional* com responsabilidades na coordenação e estabelecimento de procedimentos e métodos a serem usados pelas IES, no sentido de assegurarem a qualidade; ii) um processo de auto-avaliação desenvolvido por cada IES de forma regular e de acordo com os procedimentos e métodos estabelecidos pela Agência Nacional, a ser conduzido pelo próprio grupo académico da instituição; iii) um momento de avaliação externa que deve ter por base a auto-avaliação e incluir contributos do pessoal docente, não docente e alunos; iv) a elaboração e publicação de um relatório acerca dos resultados da avaliação externa, cujo principal objectivo deve ser o de organizar as recomendações para as instituições, no sentido de as ajudar a melhorar a qualidade do seu ensino e investigação.

Contudo, o carácter recente das práticas da avaliação aliada à quase total ausência da informação pública ou da sistematização dessas experiências coloca dificuldades na obtenção da informação para a realização de um balanço do *estado de arte* em matéria de avaliação de desempenho docente. A maioria das Instituições não tem estas experiências, outras referem que a avaliação decorre do estatuto da carreira docente e outras da avaliação dos cursos ou dos estabelecimentos de ensino (Abrantes & Valente, 2000). Do ponto de vista analítico, este dado revela, desde logo, uma realidade *sui generis* quando comparado com outros casos internacionais. Por exemplo, os questionários utilizados encontram-se reduzidos ao questionário estandardizado tendo como única fonte a avaliação dos alunos.

Em conformidade com um estudo apresentado num encontro sobre avaliação pedagógica de docentes concluiu-se que: “os questionários são reveladores do que é que uma Instituição pensa que deve ser a missão do docente. Apesar de muitos aspectos positivos dos modelos, como as questões sobre o próprio aluno, a quem se dá um sinal da sua co-responsabilização no processo ensino/aprendizagem, uma boa parte dos questionários não diferencia claramente a disciplina do

¹⁸ Esta entidade tem por missão: apreciar a coerência global do sistema de avaliação, a partir dos indicadores utilizados nas várias modalidades de ensino, nos níveis de exigência praticados, na relação entre os cursos ministrados e as tendências do mercado de trabalho e na perspectiva da dimensão europeia dos cursos avaliados. Cabe-lhe, igualmente, produzir relatórios prospectivos e recomendações de racionalização e melhoria do sistema de ensino superior.

docente e, em muitos casos não se interroga sobre o conjunto das componentes do curso e não reflecte os diferentes níveis de responsabilidade dentro da Instituição e dos respectivos órgãos, podendo induzir os alunos a atribuir aos docentes encarregados da leccionação, por exemplo, falhas que se verificam ao nível da coordenação da disciplina, às equipas docentes de cada disciplina” (Gonçalves, 2004). No que diz respeito à análise de conteúdo, os modelos existentes suscitam numerosas interrogações em termos da apresentação, das indicações sobre o preenchimento, dos temas incluídos, da concepção em concreto das questões, nomeadamente a adjectivação e o efeito de persuasão sobre os inquiridos, da adequação e do significado das escalas, da adequação ao conhecimento dos inquiridos e da adequação linguística e da clareza e simplicidade das questões, da inclusão de questões sensíveis e da extensão dos questionários (Teixeira, 2004).

Em suma, face à natureza das experiências apresentadas, pode concluir-se que as Instituições do Ensino Superior, a nível internacional, têm utilizado várias fontes de informação sobre os seus docentes e o seu ensino ministrado. As instituições têm a consciência de que a avaliação não deve ser realizada a partir de indicadores únicos e isolados, dando-lhes um carácter participativo, e, por outro lado, procurando preservar a confidencialidade dos dados e a sua utilização no sentido de melhorar a qualidade do ensino. Contudo, a nível nacional começa a emergir uma preocupação mais concertada com estas questões e que decorrem dos desafios recentes colocados às IES (e.g. Declaração de Bolonha, financiamento das Universidades, etc.).

1.8. Em síntese...

A instauração progressiva de mecanismos de avaliação no contexto do Ensino Superior coloca desafios sérios às IES. Com efeito, a avaliação é cada vez mais um conceito de natureza holística, e por isso, autores como Wilkerson, Manatt, Rogers e Maughan (2000) defendem a adopção do paradigma do *feedback* de 360°. Trata-se de um conceito comumente utilizado no contexto organizacional e que pode ser aplicado à avaliação no Ensino Superior. Assim, a avaliação da qualidade docente deve ser uma prática multidimensional envolvendo, no interior da organização, a selecção de múltiplos objectos capazes de corporizarem a complexidade das práticas que nela têm lugar, considerando a análise de trabalho do docente na instituição como um todo (investigação, ensino e serviço à comunidade). Com efeito, importa destacar a necessária

articulação dos dados da avaliação do ensino pelos alunos com outras fontes e instrumentos de avaliação no sentido de promover uma abertura do sistema de avaliação da qualidade do ensino à sua própria evolução (e.g. *peer review*, *portfólios*, observações de aulas e auto-avaliações dos docentes). Esta auto-avaliação, poderia traduzir-se num relatório reflexivo e prospectivo de actividades a elaborar pelo docente, que permitisse incluir outros aspectos não considerados nos questionários (e.g. análise de materiais pedagógicos produzidos e utilizados pelo docente) bem como para destacar o seu investimento no ensino em geral e para perspectivar as acções a desenvolver no futuro.

Todavia, julgar as contribuições de um membro da instituição depende da missão e dos objectivos da instituição e deve ser feita no contexto das responsabilidades e expectativas da instituição. Por exemplo, se o objectivo das IES é a melhoria da qualidade do ensino, então para a prossecução deste objectivo, importa, por um lado, implementar a formação dos docentes universitários para agir nos métodos de estudo/aprendizagem dos alunos, e, por outro, adoptar um instrumento avaliativo de algumas dimensões relevantes do ensino. Por sua vez, à avaliação da qualidade docente deve presidir um conjunto de pressupostos orientadores que a possam legitimar. Em primeiro lugar, importa considerar a pertinência de uma avaliação contextualizada que permita caracterizar os contextos em que ocorre e introduzir mudanças graduais nesses contextos. Nesta perspectiva, a avaliação terá como principal finalidade contribuir para a compreensão e transformação da cultura pedagógica da instituição, reforçando as características que lhe conferem qualidade e estimulando a mudança das características que lha retiram (Vieira *et al.*, 2002). Isto significa que a avaliação tem de ir além da sua função de “prestação de contas” face a um conjunto de atitudes ou procedimentos “a cumprir”, indicando pistas sobre como melhorar os processos de ensino no sentido de elevar a qualidade das aprendizagens. Na senda da avaliação formativa, a par dos questionários, os docentes dispõem ainda de outras formas de recolha de informação sobre a qualidade do ensino/aprendizagem no decurso do semestre/ano. Cita-se ilustrativamente, algumas metodologias de análise qualitativa denominadas “*fast feedback techniques*” (Davis, 1993, p.346-349) procurando responder mais de perto às especificidades do contexto, permitindo captar e aprofundar aspectos mais idiossincráticos e subtis da experiência dos alunos comparativamente à avaliação de natureza mais quantitativa.

¹⁹ Uma das formas possíveis de avaliação consiste em distribuir cartões em branco aos alunos no final da aula, no sentido de destacarem três aspectos das matérias abordadas que compreenderam melhor ou não; aspectos considerados positivos ou menos positivos das condições do ensino/aprendizagem e sugestões.

Em segundo lugar, a avaliação da qualidade docente deve focar-se em factores transdisciplinares, potencialmente válidos num conjunto diversificado de contextos de ensino/aprendizagem (transversais às diferentes áreas disciplinares, embora se reconheça que a sua pertinência varie de situação em situação). Em terceiro lugar, a avaliação da qualidade do ensino através das percepções dos alunos deve ter um sentido retrospectivo e prospectivo, integrando-se num sistema de avaliação da qualidade mais alargado que contemple estratégias de mobilização dos resultados obtidos na definição de planos de acção, principalmente no âmbito dos órgãos de gestão pedagógica e em articulação com a avaliação interna e externa dos cursos. A este propósito, no projecto “Concepções de Pedagogia Universitária” refere-se que “é problemática a inexistência de mecanismos que favoreçam a divulgação e discussão da informação recolhida através da avaliação institucional [...]” (2000, p.49). Também aqui se refere que “os resultados compilados em relatórios seriam mais úteis numa reflexão mais alargada, no âmbito dos departamentos e órgãos de gestão pedagógica dos cursos, no segundo caso com a participação dos alunos” (Melo *et al.*, 2000, p.14). Por outro lado, a avaliação deve articular-se, ainda, com uma política de qualidade que estimule a investigação pedagógica, a inovação das práticas de ensino e a (auto) formação pedagógica dos docentes.

Em acréscimo, e não menos importante, a avaliação deve ser uma prática transparente onde cada sujeito conheça as coordenadas relativamente às quais o desempenho será aferido. De novo, essa avaliação deve ser institucionalmente assumida como motor das transformações pedagógicas do seu ensino, dela devendo decorrer efeitos reconfiguradores das práticas pedagógicas.

Por último, Theall e Franklin (2001) referem a necessidade de seguir algumas “*guidelines*” para a emergência de um dispositivo de avaliação. A saber: (i) estabelecer os objectivos da avaliação e dos avaliadores; (ii) incluir os *stakeholders* nas decisões sobre os processos e as políticas de avaliação; (iii) favorecer o equilíbrio entre as necessidades institucionais e as necessidades individuais; (iv) publicitar a informação sobre a avaliação de uma forma clara (critérios, processos e procedimentos); (v) providenciar recursos para a promoção e melhoria da qualidade do ensino; (vi) promover uma avaliação regular e continuada no tempo; (vii) utilizar e adaptar instrumentos a situações específicas de ensino; (viii) utilizar instrumentos validados para efeitos de avaliação institucional; (ix) manter uma avaliação formativa separada da avaliação sumativa para tomada de decisões, sendo que na avaliação sumativa importa apenas comparar docentes com base em situações de ensino similares.

CAPÍTULO 2

Questionários de avaliação do ensino

“Precisamos de um pensamento apto a captar a multidimensionalidade da realidade.”

Edgar Morin

2.1. Introdução

A necessidade de maior responsabilidade pela melhoria da qualidade do ensino tem sido um tópico conciliador de interesse crescente no contexto do ensino superior (Coaldrake & Stedman, 1998; Ramsden, 1991; Simão, 2003ab; Sobrinho, 2000; Wilson *et al.*, 1997). Em resposta a este desafio, as instituições de ensino superior têm procurado implementar políticas de avaliação e monitorização dessa qualidade. A influência do comportamento docente na qualidade das aprendizagens dos alunos e nos seus comportamentos alerta para a sua importância como indicador de profissionalismo (Bireaud, 1995, p.176; Hargreaves, 1994 in Simões, 2000) ou de facilitador das aprendizagens. Neste capítulo, procuramos destacar a necessidade de atender às percepções dos alunos e os seus contributos ao nível da avaliação e melhoria da qualidade do ensino com base na aplicação de questionários.

Com base na literatura revista, destacamos alguns dos instrumentos disponíveis para avaliar a qualidade do ensino pelos alunos, vulgarmente conhecidos à escala internacional como *Student's Evaluation of Teaching* (SET). Com efeito, encontramos ao longo da história de avaliação, dispositivos de avaliação da docência resultantes das avaliações institucionais das Universidades (e.g. avaliação interna dos cursos), e, por outro, dispositivos de avaliação com recurso a questionários, muitas vezes, elaborados por docentes ou outros agentes do sistema educativo, no sentido da recolha de informação sobre as suas práticas. Nestas posturas encontramos um denominador comum - a necessidade de se avaliar a docência, embora se questione, a existência de instrumentos necessariamente válidos e fiáveis, susceptíveis de fornecerem informação com um nível de confiança significativo (Marsh, 2001).

Na procura de alguma sistematização em torno deste tipo de avaliação, procuramos caracterizar a arquitectura conceptual (dimensões/facetadas) associada ao conceito do ensino e/ou ensino/aprendizagem, encontrada nos vários questionários ao longo de vários anos. Com efeito, admite-se que existe um relativo consenso entre alguns investigadores sobre as características que

definem os bons docentes, independentemente do paradigma em que nos situamos (modelo ensino tradicional *versus* modelo emergente do ensino/aprendizagem). Por último, apresentamos alguns estudos a propósito da precisão e validade dos resultados obtidos em tais questionários, e que justificam a continuidade da sua aplicação nos mais diversos contextos de ensino/aprendizagem, sem prejuízo da sua articulação com outras fontes informativas (auto-avaliação pelos docentes, recurso aos pares, *portfólios*, etc.).

2.2. Papel dos alunos na avaliação dos docentes

Um ponto de viragem que marcou a importância de criar políticas de avaliação, emergiu na década de 60, período então marcado por uma forte contestação dos alunos que “reclamavam” a necessidade de se incluir a avaliação do ensino e, em particular, avaliação do desempenho docente num projecto mais vasto de avaliação institucional no contexto do Ensino Superior. Volvido esse período histórico incontornável para a avaliação, hoje, o mundo universitário tornou-se cada vez mais competitivo, com estudantes mais conscientes dos seus direitos e necessidades, e com maior poder reivindicativo (McKeachie, 1997). Com efeito, o Ensino Superior em Portugal está a entrar numa nova fase, despoletada pela tendência de diminuição do número de candidatos, pelos constrangimentos sucessivos de financiamento (sobretudo os públicos) e pela maior consciencialização da responsabilidade das IES face aos alunos que acolhem.

Por sua vez, na sociedade dita de conhecimento e aprendizagem, em que a qualidade dos serviços oferecidos passa a ser um critério de escolha para os estudantes, as instituições são obrigadas a estabelecer planos de avaliação de qualidade se quiserem sobreviver. Esta mudança, que não se esgota no espaço nacional, tem vindo de forma progressiva a aumentar a competitividade entre as Instituições que, para assegurarem o número de alunos e, portanto, os montantes de financiamento, são confrontadas com a necessidade de alterarem a sua postura, e prosseguirem o “*bem público*” que, no fundo, define a agenda da sua missão (Amaral & Rosa, 2004). Assim, de um isolamento, tanto a montante como a jusante, com uma responsabilidade limitada ao período de frequência, passa a verificar-se, por exemplo, a emergência de serviços educacionais com base no pressuposto que o estudante é um “cliente do Ensino Superior” (Lourtie, 1998). Por outro lado, cada vez mais frequentemente, é pedido aos elementos de cada escola de Ensino Superior que demonstrem a forma como estão comprometidos com as suas

instituições e com a sua profissão, e conseqüentemente, a forma como estão a responder ao investimento que a sociedade realizou para com eles (Braskamp & Ory, 1994). A par destas preocupações emerge a importância de assumir a qualidade do ensino/aprendizagem, e a necessidade de olhar para as percepções dos alunos “*good teaching means seeing learning through the learner's eyes*” (Ramsden, 1998, p.353). De acordo com alguns autores, a legitimidade da avaliação do ensino pelos alunos ganha força ao defender-se que os alunos são os primeiros agentes da aprendizagem e, como clientes, têm o direito de participar, em certa medida, na avaliação do ensino (Felder, 2004).

Segundo Murray (1984), só os alunos são capazes de julgar se os comentários do docente aos seus exames ou trabalhos são úteis ou não e responder pela adequabilidade ou não dos materiais de apoio utilizados em prol da aprendizagem. Aliada a esta preocupação com a centralidade do aluno preconizada na Declaração de Bolonha²⁰, encontra-se a necessidade de criar uma nova visão e paradigma de ensino superior procurando remodelar os *curricula* de forma a ir mais além que a mestria cognitiva das disciplinas (Ferreira, 2003). Por sua vez, Chickering e Gamson, em “*Seven Principles for Effective Undergraduate Education*” (1987), argumentam que o sucesso académico dos alunos se associa à aquisição de competências transversais ou “*soft skills*” (capacidades e aptidões para a comunicação, análise criativa e crítica, pensamento independente, trabalho em equipa e em contextos multiculturais). Ao defender-se estratégias activas de ensino, através da apresentação de casos/problemas, da participação em projectos de investigação, do recurso à avaliação contínua, da discussão dos assuntos nas aulas, do confronto de ideias de teorias ou de vias alternativas de resolução de problemas ou da orientação dos alunos na procura intencional de nova informação, a Universidade torna os alunos mais activos, responsáveis e autónomos no processo de aprendizagem (Almeida, 2002). De facto, é desta forma que os alunos estarão capazes de dar real significado à informação e, a partir desta, construir conhecimento e desenvolver competências.

Teorias recentes na Psicologia e na Educação definem a aprendizagem como um processo pessoal de construção progressiva de conhecimento ou “construção activa do conhecimento”. Aqui o docente é assumido como facilitador e orientador da aprendizagem dos alunos, alguém que

²⁰ Os Ministros da Educação europeus reunidos em Praga, em Maio de 2001, reconheceram a importância e a necessidade de mais três linhas de acção para evoluir no processo de Bolonha: promoção da aprendizagem ao longo da vida; maior envolvimento dos estudantes na gestão das instituições de Ensino Superior e promoção da atractibilidade do Espaço Europeu do Ensino Superior.

coloca problemas, apresenta pistas para a sua fundamentação e vias alternativas da sua resolução (Almeida, 2002; Biggs, 1999; Dewey, 1956). Outros autores como Lobrot (1966 in Sakai & Ukita, 2002) defendem um modelo de auto-gestão pedagógica, em que o docente deve, no par educativo, despertar e incentivar abordagens activas de ensino/aprendizagem “*condição de levar os alunos a uma situação onde não tenham saída, a não ser que pensem que estejam envolvidos e gastem energia na experiência académica*”. Pensando desta forma, alunos e docente assumem logicamente uma quota-parte de responsabilidade na aprendizagem e no rendimento académico.

Por sua vez, a concepção do ensino/aprendizagem passa a ser encarada como “finalidade e o processo de transformação dos alunos em participantes críticos e produtores criativos de saberes (*saber, saber fazer, saber aprender, saber ser*), actores activamente implicados na aprendizagem escolar e ao longo da vida, intervenientes críticos na melhoria da racionalidade e justiça dos contextos em que (inter) agem” (Vieira *et al.*, 2002, p.31). Para responder a este desafio, importará no futuro, perceber se, de facto, os docentes também conseguem promover estas competências, isto é, se para além de conhecimentos técnico-científicos, promovem a aprendizagem das “*soft skills*”.

2.3. A formação dos docentes para a docência

Na procura de um modelo do bom docente, a evolução do conceito e das metodologias utilizadas, salienta percursos com incidências diversas (nos traços de personalidade, nos comportamentos em situação profissional e na influência do comportamento docente na qualidade das aprendizagens dos alunos e nos seus comportamentos), alertando para a importância do comportamento docente (Estrela, 1986). Por sua vez, face à insatisfação com as definições sobre o que é ser-se docente universitário, Fiolhais (2004) defende três tipos possíveis de docente com diferentes enfoques. O primeiro relaciona-se com o docente cujo enfoque reside na transmissão de conhecimentos resultantes da investigação ou de uma área do saber ou da investigação do docente. Este saber tende a basear-se na sua transmissão de forma unidireccional e não questionando o aluno. Um segundo docente é aquele cujo enfoque reside naquilo que faz, e por isso, reflecte sobre as suas práticas e nos resultados produzidos, assumindo-se também como co-responsável por eles (ênfase maior no que é aprendido do que é ensinado). O terceiro é aquele que se baseia no que os novos modelos de ensino/aprendizagem preconizam, não centrando

exclusivamente o processo no que se aprende mas na forma como se aprende. A sua posição é de orientador e coordenador dos processos de aprendizagem, bem como facultar, através de processos diversos, os saberes necessários à aprendizagem. Assim, cabe ao aluno a responsabilidade por aquilo que faz e pela construção do conhecimento que vai aprendendo. Não constitui um objectivo central desta investigação dissecar a diversidade de perspectivas relativas às concepções do que é ser docente e do que é ensinar, que se traduzem numa diversidade de metáforas²¹. Contudo, dadas as implicações que esta questão acarreta para a problemática da avaliação do desempenho docente, não só não nos é possível omiti-la, como é igualmente importante salientar que a falta de consenso nesta matéria constitui um dos entraves ao desenvolvimento de modelos de avaliação. Aliás, esta falta de consenso revê-se na heterogeneidade de instrumentos utilizados para avaliar a qualidade de ensino e do número variável de itens e do respectivo conteúdo que tendem a integrar as sub-escalas definidoras do constructo (Marsh, 2001). Assim, a necessária diversidade de definições obriga-nos a contextualizar o nosso objecto de estudo.

Neste sentido, procuramos avaliar o desempenho do docente por referência ao seu comportamento no trabalho (mais ao que faz do que ao que pode fazer, isto é, quão competente é). Sendo específico à situação de trabalho, o desempenho depende da competência do docente, do contexto em que trabalha e da sua habilidade para aplicar as competências²² em qualquer momento.

Centrando a atenção neste aspecto do desempenho docente e das dimensões que o suportam, Ellington (2000) sugere algumas regras de ouro para um docente eficaz no Ensino Superior, i) descobrir e descrever como os alunos aprendem; ii) fixar objectivos de aprendizagem apropriados; iii) adoptar métodos de ensino/aprendizagem adequados; iv) adoptar métodos de avaliação adequados; v) monitorizar e avaliar o ensino; vi) procurar melhorar o seu desempenho; e vii) manter-se actualizado. Por sua vez, Market (2001 in Ferreira, 2003) organiza, num breve comentário, as suas ideias acerca dos elementos característicos do bom docente: apresenta

²¹ (cf. Modelo multidimensional do papel do docente na Educação Médica de Harden & Crosby, 2000), preconiza a existência de 12 papéis associados ao docente por sua vez agrupados em 6 categorias. A saber: (i) transmissor; (ii) modelo de desempenho; (iii) facilitador ou na qualidade de tutor; (iv) avaliador; (v) planificador; e (v) gestor de recursos (Ferreira, 2003).

²² Segundo Alarcão (1996), verifica-se hoje a adopção de uma atitude holística, preferindo-se falar em competência (associada a qualquer conhecimento específico, aptidão ou posição do valor profissional, cuja posse se crê ser relevante para uma prática de ensino com sucesso) e não competências (coisas específicas que os docentes sabem, fazem ou acreditam mas não aos efeitos destes atributos nos outros. A qualidade é assim referenciada em relação ao conjunto de competências que possui).

motivação intrínseca, privilegia o ensino centrado no aluno, utiliza a parcimónia na comunicação, demonstra entusiasmo e interesse na matéria que lecciona, procura formação contínua, etc.

Num estudo realizado no Brasil com uma amostra de 60 docentes universitários (Sakai & Ukita, 2002) concluiu-se que a característica mais valorizada e susceptível de definir um bom docente é a capacidade de despertar interesse dos alunos pelos assuntos e a menos valorizada é o cuidado com a aparência. Com efeito, as características que foram indicadas para definir a eficácia docente, tomando as percepções dos docentes, estão mais relacionadas com os conhecimentos específicos do docente (e.g. conhecimento das matérias, adequação da relação docente/aluno, entusiasmo no trabalho desenvolvido com os alunos) e com as características mais pessoais (e.g. equilíbrio emocional e capacidade de adaptação a novas situações). Por outro lado, a literatura tem, ainda, associado as condições que conduzem a abordagens profundas e significativas da aprendizagem no Ensino Superior por parte dos alunos aos docentes que imprimem entusiasmo e estímulo à curiosidade, à resolução de problemas e à autonomia dos alunos, como também o recurso a métodos de aprendizagem que recorrem à cooperação, ao trabalho de grupo e à interacção dos pares (Gibbs, 1992 in Almeida, 2002).

No âmbito de estudos sobre o perfil psicológico dos docentes do Ensino Superior, um estudo de Rego (2003) adianta a percepção dos docentes e discentes sobre ser “bom docente” com base num repertório de comportamentos de cidadania (comportamento participativo; orientação prática; conscienciosidade pedagógica; e cortesia), assumidos como “conjunto de acções do docente que ajudam a modelar o contexto organizacional, social e psicológico em que as tarefas formais dos agentes escolares se desenrolam”. Trata-se de comportamentos extra-papel, isto é, comportamentos discricionários que ultrapassam as expectativas de papel e os conteúdos funcionais formalmente estabelecidos, mas que simultaneamente beneficiam a organização. Extrapolando para o contexto do Ensino Superior, é possível hipotetizar que estes comportamentos exercem efeitos (directos ou não) no desempenho académico dos alunos. Assim, é plausível assumir-se que os comportamentos de empenho, dedicação e disponibilidade dos docentes possam suscitar comportamentos similares nos seus alunos (efeito de imitação social) contribuindo desta forma para a melhoria do processo de ensino/aprendizagem. Por sua vez, com as mudanças que se têm vindo a verificar, em particular, com a “revolução cognitiva” em Portugal, assume relevância o paradigma “mediacional” da figura do docente, enquanto entidade possuidora de crenças, preconceitos, representações e estereótipos, forjados ao longo da sua biografia pessoal e profissional (Lopes, 2002, p.37 in Magalhães, 2004). Na tradução desta afirmação para o domínio

da docência, não é possível dissociar a pessoa do papel e estatuto docente, o que no fundo traduz a ideia consubstanciada na afirmação “*transmitimos o que sabemos, ensinamos o que somos*” (Zabalza, 1999).

2.4. Tipologia de questionários

De uma forma genérica, os questionários tipicamente utilizados para avaliar o ensino pelos alunos, são de tipo papel-lápis, obedecendo a um formato de *rating scales*, sendo no entanto, cada vez mais notória a emergência de questionários formatados *on-line*. De facto, a investigação em torno deste domínio tem dado conta do número crescente de instituições americanas que utilizam ou que pensam vir a utilizar este formato de avaliação (Hoffman, 2000 in Sorenson & Reiner, 2003). A discussão sobre as suas vantagens em relação às metodologias tradicionais terá que ser equacionada em termos comparativos. Todavia, a justificação para a emergência do *on-line* decorre em larga medida de factores tais como o tempo, a flexibilidade, a quantidade de comentários qualitativos e quantitativos e os custos associados.

Em termos de vantagens dos questionários *on-line* destaca-se o facto dos alunos disporem de mais tempo para o seu preenchimento, e daí, decorrer um aumento significativo de informações quantitativas e qualitativas. Por sua vez, o intervalo de tempo que medeia a realização da avaliação propriamente dita e a divulgação dos seus resultados é relativamente curto.

A flexibilidade na adaptação e na personalização dos questionários constitui uma mais valia em termos da sua formatação facilitando o desenvolvimento de modelos de questionários mais personalizados face à heterogeneidade de métodos de ensino. Contudo, na prática, a maioria das instituições de ensino superior não permite efectuar alterações ao modelo de questionário proposto. Por sua vez, a investigação indica que os alunos tendem a produzir comentários qualitativos e quantitativos mais elaborados e consistentes comparativamente ao método tradicional (Hardy, 2002; Hmielecki & Champagne, 2000 in Sorenson & Reiner, 2003). De acordo com os autores, esta situação decorre do facto dos alunos não responderem à pressa, e perceberem a possibilidade de anonimato nos comentários escritos às questões de natureza aberta. Por último, importa ainda mencionar a produção de relatórios (*reports*) coerentes e sistemáticos relativos aos resultados de uma forma imediata e relativamente acessível (através de PC em qualquer parte do Campus) permitindo desta forma a redução significativa de custos com papel e de recursos

humanos para o processamento dos dados, sendo apontado, nalguns casos, valores na ordem dos 97% de poupança relativamente ao modelo tradicional²³ (Hmieleski & Champagne, 2000, *ibidem*).

A investigação em torno desta metodologia tem suscitado algumas preocupações/desafios relacionadas com a percentagem de respostas obtida que tende, em alguns casos a ser pouco significativa. Tal situação poderá dever-se à percepção de falta de anonimato, falta de compromisso e apatia dos por parte dos alunos e aos problemas técnicos decorrentes da implementação do sistema. Contudo, outros estudos indicam que esta limitação não se verifica em todas as situações (Ballantyne & Flower, 2001; Goodman & Campbell, 1999 in Sorenson & Reiner, 2003). As respostas enviesadas também constituem objecto destas preocupações. Estudos mostram que os alunos com história de sucesso académico tendem a envolver-se com maior probabilidade no preenchimento de questionários comparativamente aos que possuem uma história menos bem sucedida (Layne, De Cristoforo & Mc Ginty, 1999, *ibidem*). A par desta tipologia de questionário e de respostas de determinados grupos de alunos encontra-se também o viés associado aos comentários escritos devido à baixa frequência de resposta. Por sua vez, os níveis de contacto com as novas tecnologias de informação poderão excluir determinados grupos de alunos do processo avaliativo.

Em suma, e de acordo com a literatura revista, as análises comparativas entre as duas formas de avaliação sugerem a não existência de diferenças consistentes apesar das variações encontradas em alguns estudos.

Autores como Marsh e Dunkin (1992) defendem que a concepção dos questionários tende a resultar de (i) abordagens empíricas, em particular com recurso à análise factorial e análises multimétodo-multitraço; (ii) análise do constructo ensino eficaz com base no *feedback* dos alunos e docentes²⁴ (entrevistas); e (iii) dos contributos das teorias de ensino/aprendizagem. Na prática, a generalidade dos questionários têm sido construídos com base no segundo pressuposto. Por sua vez, os instrumentos utilizados para avaliar a qualidade de ensino apresentam diferentes sub-escalas e diferem no número e no conteúdo dos itens (Marsh, 2001). Partindo de uma perspectiva longitudinal dos estudos realizados neste domínio, pode concluir-se que, com maior ou menor

²³Theall (2002 in Sorenson & Reiner, 2003) contesta estes números uma vez que foram comparadas situações extremadas: apresentação do melhor estudo de caso sobre a avaliação on-line e o pior estudo de caso sobre o método tradicional de avaliação.

²⁴As pessoas possuem teorias implícitas sobre o bom ensino, derivadas das múltiplas experiências tidas nas suas vidas na qualidade de docentes ou de alunos. Em geral, as definições do que é um bom ensino são construídas a partir dessas experiências e são extremamente variáveis e idiossincráticas.

número de itens e/ou factores ligados ao ensino eficaz, regista-se a presença de várias comunalidades entre os questionários disponíveis.

Para a realização deste nosso estudo tivemos acesso a vários instrumentos nomeadamente, o Questionário de Avaliação do Ensino Ministrado (QAEM; Universidade do Minho, 1991); Questionário sobre a Qualidade da Pedagogia na Universidade (QQPU; Vieira *et al.*, 2001); *Evaluación del profesorado de la valoración de los Alumnos* (EPVA; Fernandez *et al.*, 1996); *Student Feedback Questionnaire*, (SFQ; Lee Harvey, 2001); *Course Experience Questionnaire* (CEQ; Ramsden, 1991); *Instructional Development and Effectiveness Assessment* (IDEA; Centra, 1975); *Student Instruction Rating* (SIR II; Linn, Centra & Tucker, 1975); Comportamentos de Cidadania Docente (CIDOCE; Rego, 2000) e o *Student's Evaluation of Educational Quality* (SEEQ; Marsh, 1987, 1991, 2001).

Em seguida, passamos a descrever com maior pormenor alguns modelos de questionário, mais concretamente o utilizado durante vários anos na Universidade do Minho como meio de acesso às percepções dos alunos sobre a qualidade do ensino. Trata-se de um instrumento composto por três partes, a primeira (A) referente à leccionação da disciplina avaliada, a segunda (B) referente à postura do próprio aluno face à disciplina, e a terceira (C) referente ao curso (em anexo juntamos exemplar deste questionário - Anexo A). A primeira parte é composta por 20 itens, a segunda por 10 e a terceira por 9. As questões são formuladas numa escala de 4 pontos em que “1” corresponde a uma avaliação mais desfavorável e “4” a uma avaliação mais favorável. A análise factorial do mesmo reconhece a presença de 3 dimensões: i) Elementos de estudo/Avaliação/Horários/Preparação; ii) Estratégias de Ensino; e iii) Articulação da Disciplina²⁵.

2.5. Dimensões consideradas nos questionários

As dimensões encontradas nos questionários dirigidos aos alunos baseiam-se em concepções cognitivistas, construtivistas e interaccionistas da aprendizagem. Tais concepções decorrem da Psicologia do Desenvolvimento e da Psicologia Cognitiva. Com base nesta matriz conceptual, Mackie (1981 in Marsh 2001) chegou a 10 princípios que devem ser contemplados aquando da formulação dos itens dos questionários: i) o aluno deve estar motivado para aprender; ii) a abordagem da aprendizagem deve atender às diferenças e estilos individuais de aprendizagem

²⁵ Dados presentes em reunião do Conselho Académico da Universidade do Minho em Junho de 2001.

dos alunos; iii) a abordagem da aprendizagem deve atender aos conhecimentos prévios e atitudes dos alunos; iv) a importância de reforçar o que deve ser aprendido; v) a abordagem da aprendizagem deve proporcionar oportunidades para praticar; vi) o aluno deve assumir-se como elemento participante e proactivo através do questionamento em detrimento de uma postura de ouvinte; vii) as matérias devem ser organizadas em unidades de aprendizagem e devem ser abordadas com ritmo adequado; viii) apoio/orientação docente deve ser assegurado no sentido de se desenvolverem novas respostas; ix) capacidade de generalização dos conhecimentos às novas situações de aprendizagem; e x) a abordagem das matérias de uma forma significativa para o aluno.

Contudo, a revisão da literatura sobre a avaliação, confirma a ideia de que as teorias do ensino/aprendizagem não têm sido explicitamente referenciadas para a concepção e desenvolvimento dos instrumentos disponíveis. Isto não impede a progressiva entrada de modelos cognitivistas nas práticas docentes no Ensino Superior. Com efeito, o ensino é cada vez mais concebido como um processo interpessoal, que utiliza essencialmente a comunicação, a situação pedagógica levada a cabo pelo docente como meio de provocar, favorecer e fazer alcançar a aprendizagem de um saber ou de um saber fazer²⁶. A descentração do processo de ensino permitiu ao docente “falar menos, fazer e agir mais e observar durante este tempo” (De Ketéle, 2004), tornando-se assim um mediador dos processos de aprendizagem e um organizador da gestão das aprendizagens, não se assumindo apenas como um distribuidor de conteúdos.

De seguida, procuramos ilustrar a importância de algumas dimensões transversais, presentes numa lógica organizativa do ensino mais tradicional do ensino ou na emergência progressiva do paradigma do ensino/aprendizagem. Com efeito, a literatura tem identificado o carácter imprescindível do *feedback* como elemento moderador dos objectivos dos alunos e promotor da motivação (Rosário, 2003). Assim, os ambientes que definem e transmitem expectativas adequadas às competências dos alunos, disponibilizando o apoio instrumental de que o aluno necessita e favorecendo o desenvolvimento de estratégias baseadas no esforço, parecem contribuir de uma forma significativa para a motivação dos seus alunos. Nem sempre os alunos são suficiente ou claramente informados sobre as aprendizagens que é suposto terem de realizar,

²⁶ A aprendizagem exige a compreensão das matérias pelo aluno. *Comprehendere significa* “aprender em conjunto” (Pimenta & Anastasiou, 2002). Por sua vez, Ausubel (1968 in Fleming, 2003) refere a importância dos esquemas sucessivos de assimilação/acomodação dos novos conhecimentos por referência aos conhecimentos prévios dos alunos, para uma aprendizagem bem sucedida, referindo que: “*The most important single factor influencing learning is what the learner already knows*”.

os objectivos que devem atingir, as tarefas a desenvolver e o caminho a percorrer. Tudo isto acaba por condicionar, na globalidade, todo o processo educativo.

Nos últimos anos, estudos em várias áreas da Psicologia têm enfatizado a importância da relação docente/aluno, considerando-o de real valor educacional (Sakai & Ukita, 2002). Por sua vez, o apoio sócio-afectivo – e, mais genericamente, o “comportamento de consideração²⁷” – pode constituir um dos modos de actuação docente com maior impacto junto dos alunos (Feldman, 1976 in Spencer & Schmelkin, 2002). Por exemplo, Ramsden (2003, p.74) refere a importância dos aspectos emocionais na relação aluno/docente, secundarizando um conjunto de métodos e técnicas específicas do ensino. Desta forma, o autor alerta para a noção de que o respeito e o sentido de justiça constituem elementos motivadores de uma aprendizagem percebida como mais relevante e efectiva por parte dos alunos. Por sua vez, Carson (1996) acrescenta que “*studies of the relationship between emotions and cognition explain in another way the link between how students feel about their professors and how they perform in the classroom*”. Com efeito, os alunos tendem a envolver-se com maior frequência e intensidade nos assuntos das aulas se o docente interagir com eles, de uma forma que promova o envolvimento, auto-confiança, encorajamento e o interesse²⁸. Por sua vez, este envolvimento é mais significativo quando os alunos percebem que os docentes se preocupam com eles e atendem à sua individualidade (e.g. ritmo e capacidade de trabalho e conhecimento dos seus nomes). Vários são os estudos realizados no domínio do ensino superior que têm identificado, de forma recorrente, factores tais como: “*student centeredness*, “*respect for students*” e “*lecturer rapport*”, entre outros (*ibidem*). Por sua vez, Rodabaugh (1996, p.41 in Fleming, 2003) ilustra alguns exemplos de comunicação de respeito na relação aluno/docente, nomeadamente: dar tempo aos alunos para colocarem questões; permitir o desafio de ideias ou pontos de vista diferentes do docente; e encorajamento de discussões/debate de ideias. A mesma autora salienta que “*sarcasm and indifference can be easily conveyed through body language clues and this should be avoided (...) students admire faculty members but they not expected to be best friends with them*” (p.41).

Nesta linha de pensamento, Rego (2003) defende que o docente no Ensino Superior deve ser “cortês, não tratar os alunos como crianças, preparar as aulas, ser metódico na exposição, utilizar exemplos práticos e fomentar a participação nas aulas”. Adianta mesmo que “os professores que não executam estes comportamentos são considerados de menor desempenho”.

²⁷ Fleishman (1957); Yukl (1971, 1998) in Rego, 2003.

²⁸ No fundo, o docente apreciado pelos alunos é “*hardness of head and softness of heart*” (Ory,2001).

Aliás, Tavares (1998 in Santos, 2001), refere que uma das causas do insucesso escolar no ensino superior é a disfunção ao nível da transmissão e compreensão de uma mensagem entre docentes e alunos, que tende a associar-se com a falta de comunicação em sala de aula. A título ilustrativo um estudo desenvolvido por Chickering (1999 in Santos, 2001) mostrou que entre 80 a 90% dos docentes ainda encaram o processo de ensino como a exposição de matéria e a realização de exames, assente numa lógica que coloca o docente como mero transmissor e os alunos como meros receptores.

As investigações conduzidas por Gorham (1988) permitem, também, constatar a existência de efeitos positivos entre o entusiasmo docente, a valorização dos alunos e do seu trabalho, e o envolvimento em contactos fora da sala de aula (activando esquemas de *self disclosure* de exploração de opiniões/perspectivas dos assuntos por parte dos alunos) ao nível da qualidade da sua aprendizagem (e.g. idas aos gabinetes e o uso de e-mail e telefone). Contudo, a interacção no contexto fora das aulas também se encontra condicionada pela iniciativa dos alunos. O mesmo autor avança, mesmo, que a importância destes comportamentos é proporcional ao tamanho das turmas, assinalando-se, por exemplo, um impacto mais significativo junto de turmas numerosas. Outro aspecto que parece importante é a pontualidade e assiduidade docente e que tende a revelar o seu brio e comprometimento profissionais, sugerindo mesmo que não será razoável exigir a pontualidade ou assiduidade se o próprio docente as não cumpre. Assim, os comportamentos dos docentes podem representar um exemplo (positivo ou negativo) para os alunos, para além de que, a inconsistência entre o que é requerido aos alunos e o que é feito pode ter efeitos negativos sobre a imagem que os mesmos projectam (Rego, 2003).

Por sua vez, Terenzini e Pascarella (1994) alertam para a existência de diversos mitos associados ao processo de ensino/aprendizagem, nomeadamente a noção de que “*student’s academic and non academic experiences are separate and unrelated areas of influence on learning*”, o que tem contribuído para uma visão enviesada do ambiente académico em contexto universitário. Face à importância da dimensão relacional no processo de ensino/aprendizagem, autores como Whitman (1987 in Fleming, 2003) alertam para a necessidade de estimular o desenvolvimento de estudos que procurem investigar a natureza e a qualidade da relação aluno/docente, no sentido de “*to indentify the constructive side of professional intimacy, self-disclosure and mentoring (...) determine why and how teacher-students relationships deteriorate and help faculty construct more sucessful relationships.*” Efectivamente, é extensa a lista de

comportamentos que podem ser objecto de estudo de um ensino/aprendizagem de qualidade que, cada vez mais, aparece defendido no contexto de Ensino Superior.

Um olhar atento sobre a literatura permite verificar a existência de um elenco significativo de factores e aspectos associados às práticas e aos comportamentos dos docentes com efeitos positivos ou negativos sobre os alunos. Todos estes factores acabam por afectar o próprio processo de ensino/aprendizagem (Ellington & Ross, 1994; Helms & Key, 1996; Horsburgh, 1999; Hau, Chung & Siu, 1998; Rowley, 1996; Stringer & Irving, 1998 in Rego & Sousa, 2000; Rego, 2003). Citam-se, ilustrativamente: o *feedback* da aprendizagem pontual e justo, entusiasmo e dinamismo; espírito de diálogo e práticas reflexivas partilhadas; actuação amistosa do docente, interesse e disponibilidade para atender os alunos; humor na apresentação das matérias, encorajamento, aulas estimulantes e desafiantes; preparação dos materiais de leccionação e clareza das explicações; pontualidade, credibilidade e fiabilidade nas avaliações; promoção da autonomia e da aprendizagem activa; cortesia e conscienciosidade no exercício do papel docente. Por isso, o sucesso da actividade docente reflecte a interacção entre as qualidades associadas ao ensino organizado e à qualidade da interacção com os alunos, nomeadamente quando se atende à dificuldade das matérias, à carga de trabalho e ao interesse dos alunos pelas matérias. Com efeito, quando esta interacção é positiva, os alunos tendem a avaliar a disciplina e o docente de modo mais positivo e a procurar obter níveis elevados de sucesso. Contudo, ao longo da história tem-se encontrado diferentes entendimentos sobre os factores explicativos de um ensino qualitativamente superior, ou mais adequado ao Ensino Superior. Assim, e de uma forma necessariamente breve, procuramos dar conta de alguns contributos de autores sobre esta matéria.

Em 1980, Cohen sintetizou o trabalho de alguns investigadores e sugeriu seis dimensões associadas ao ensino de qualidade: competência em métodos específicos de ensino; empatia com os alunos; estrutura e organização adequada; grau de dificuldade apropriado; forma de relacionamento com os alunos e o *feedback* aos alunos. Outros autores acrescentaram a esta lista a capacidade de adaptar o ensino ao estilo de aprendizagem do aluno. Por sua vez, Feldman (1986 in Macedo, 2001) identificou catorze dimensões relativas ao ensino efectivo, teórica e empiricamente distintas: características dos docentes (relação empática com os alunos), aspectos do curso como um todo (organização) e acções específicas (forma de atribuição das notas). Centra (1993), assim como Braskamp e Ory (1994), identificaram seis factores comumente encontrados nas avaliações: (i) organização e planeamento da disciplina; (ii) clareza e competências de comunicação; (iii) interacção da turma e relação individual; (iv) grau de dificuldade e carga de

trabalho; (v) avaliação (notas/classificações); e (v) percepções de aprendizagem. Outros estudos, realizados por Trent e Cohen (1973 in Cantero, Deus & Abalde, 2002), estabelecem como factores principais (i) clareza de organização, interpretação e explicação; (ii) fomento da discussão e apresentação de diferentes pontos de vista; (iii) estimulação do interesse e motivação dos alunos; (iv) atenção e interesse pelos alunos e, (v) manifestação de entusiasmo. Outros contributos de Cashin (1995) revelam a presença de seis dimensões: organização e planeamento da disciplina; clareza e capacidade de comunicação; qualidade da interacção docente/aluno; nível de dificuldade da disciplina e carga de trabalho; avaliação (classificação/notas) e a percepção da aprendizagem pelo aluno.

Um estudo comparativo entre os questionários de avaliação do ensino que a Universidade do Minho utiliza e os similares de outras instituições de Ensino Superior em Portugal, permitiu verificar, por um lado, que os questionários das várias Universidades Portuguesas são bastante similares no que diz respeito aos itens que contemplam. Independentemente de algumas diferenças encontradas em cada um dos questionários, foi possível identificar um conjunto de elementos comuns, tais como: conhecimento das matérias; organização/preparação das aulas; clareza com que expõe a matéria; utilização adequada do quadro e dos acetatos; capacidade de se fazer ouvir adequadamente; disponibilização/indicação de elementos de estudo; nível de segurança na exposição da matéria; adequação dos exemplos/exercícios utilizados; capacidade para motivar os alunos; assiduidade e pontualidade; disponibilidade para atendimento dos alunos; relacionamento com os alunos; definição clara e atempada do regime de avaliação e apreciação global do docente.

Mais recentemente, outros estudos desenvolvidos em Portugal, com base em análises factoriais confirmatórias, identificaram três macro-dimensões²⁹. Uma primeira dimensão prende-se com a competência (estrutura, quantidade de informação disponível acerca dos meios que levam à obtenção de resultados desejados, incluindo a clareza de expectativas, consistência e previsibilidade, ajuda instrumental, adequação das estratégias de ensino ao nível do aluno e suas dificuldades). Uma segunda dimensão reporta-se ao apoio à autonomia e oportunidades de auto-determinação incluindo as escolhas, actividades relacionadas com os interesses dos alunos, reconhecimento das opiniões (participação), redução do controlo, vigilância e pressão. Por último, a terceira dimensão integra aspectos de relacionamento, nomeadamente a qualidade da relação

²⁹ Comissão de Avaliação Interna do Curso de Psicologia da Universidade do Porto (2003).

interpessoal, incluindo tempo para comunicar, expressão do afecto positivo nas interacções, sensibilidade à individualidade, dedicação de atenção, apoio e disponibilidade.

Face ao levantamento feito, é possível concluir que as dimensões referidas correspondem às características de docentes exemplares, destacadas na síntese dos resultados de pesquisas sobre a eficácia do ensino, e em particular sobre o desempenho docente. Nesta linha, podemos admitir que existe um relativo consenso entre os investigadores sobre as características que definem os bons docentes, ou que tornam mais eficiente o ensino nas IES.

2.6. Estudos da precisão e validade dos resultados em tais questionários

A validade dos questionários como fonte informativa tem sido objecto de vários estudos. De acordo com Greenwald e Gillmore (1997), as preocupações com a validade dos questionários têm-se organizado em função da sua (i) estrutura e natureza conceptual (desde uma lógica unidimensional a uma lógica multidimensional³⁰); b) validade convergente (grau em que os questionários se correlacionam com outros indicadores do ensino); c) validade discriminante (são os questionários influenciados por outros factores que não estão directamente associados ao ensino eficaz?); e d) validade prática e consequencial (são os resultados obtidos nos questionários utilizados para a tomada de decisões ou para o desenvolvimento e melhoria profissional dos docentes?).

Apesar da investigação publicada nos anos 70 abranger uma variedade de preocupações acerca da validade dos resultados obtidos nesses questionários, emergiu um movimento que questionou fortemente a validade dos mesmos com base na influência das notas/classificações nas avaliações, *“notion that a teacher can get a “good” rating simply assigning “good” grades”*. Assim, acreditava-se que o docente poderia obter avaliações mais favoráveis quando atribuía classificações positivas aos alunos *“it is clear...an instructor (who) inflates grades...will be much*

³⁰ À multidimensionalidade defendida por Marsh e colaboradores contrapõe-se uma visão unidimensional do desempenho docente que tem sido suportada pela evidência dos estudos factoriais e dos estudos sobre as teorias do processamento de informação. No primeiro caso, um estudo realizado por Abrami e d`Apollonia (1997) indica que dos 31 dos 35 itens dos questionários analisados apresentam valores de saturação de .60 no primeiro factor o que traduz a existência de uma dimensão mais global da competência docente. No segundo caso, os autores defendem que os alunos tendem a avaliar o docente com base numa percepção mais global (*efeito de halo*) sendo que, os mesmos tendem a inferir percepções a partir do sentido dos primeiros itens do questionário.

more likely to receive positive evaluations” (Worthington & Wong, 1979, p.774 in Greenwald, 1997). Por outro lado, estes estudos foram posteriormente contestados pelo facto de omitirem o tratamento que é feito à manipulação das notas e o seu impacto nas avaliações, bem como pela constatação de que, esse efeito a existir, não é estatisticamente significativo (Abrami, Dickens, Perry & Leventhal, 1980; Marsh, 1987; Marsh & Dunkin, 1992).

Contudo, a partir da década de 80, os estudos em matéria de avaliação do ensino tomando as percepções dos alunos começaram a ganhar autoridade científica suficiente para se assumirem no quadro da investigação educacional (Greenwald, 1997, p.1182 in Ory, 2001), e daí a sua aplicação exponencial nos diferentes níveis de ensino, incluindo o Ensino Superior. Uma análise das publicações acerca dos questionários aplicados aos alunos (ERIC; 1966-1997; PsyclINFO, 1967-1997) revela que os estudos em torno da defesa da validade dos questionários foi objecto de um número significativo de investigações no início dos anos 80, contudo assistiu-se à emergência de uma postura neutra nos estudos que defendiam a validade (decrécimo verificado de 31 para 16 estudos desde 1976-1980 e 1981-1985). Todavia, a demonstração da validade das avaliações pelos alunos foi amplamente corroborada por estudos correlacionais que incluem, em particular, estudos multisectoriais, *path analysis* e estudos multimétodo-multitraço.

Nos estudos multisectoriais, procurou-se estudar a associação entre as avaliações e o rendimento académico dos alunos nas várias componentes de uma disciplina leccionada por vários docentes. Neste campo, os investigadores correlacionaram a média obtida em cada uma dessas componentes com a média do rendimento académico dos alunos. No essencial, este tipo de estudos tem mostrado correlações estatisticamente robustas entre as duas fontes informativas (Abrami, d’Apollonia & Cohen, 1990; Abrami & d’Apollonia, 1997). Por sua vez, estudos decorrentes da *path analysis* mostram que as notas e as avaliações se encontram associadas aos níveis motivacionais dos alunos nos cursos, e em particular, nas disciplinas “*the relationship between grades and student satisfaction might be viewed as a welcome result of important causal relationship among other variables rather than simply as evidence of contamination due to grading leniency*” (Howard & Maxwell, 1980, p.810). Com efeito, e de acordo com um estudo de Marsh (1980), concluiu-se que o interesse prévio dos alunos apresenta um impacto muito significativo no conjunto dos resultados obtidos nas avaliações, contribuindo mesmo para cerca de um terço da relação entre a nota esperada e as avaliações nos questionários.

Nos estudos multimétodo-multitraço, os investigadores procuram avaliar a validade discriminante das avaliações, correlacionando-o com outros indicadores tais como as percepções

dos antigos alunos, colegas e auto-avaliações em várias disciplinas. Estes estudos tendem a atribuir as diferenças encontradas mais aos atributos dos docentes e menos à influência das características externas (e.g. alunos, disciplinas, contexto). Um exemplo típico destes estudos baseia-se na recolha de várias fontes de informação: colegas, alunos, antigos alunos e autoavaliações do docente, tomando questões de natureza fechada, aberta e entrevistas. Os resultados têm mostrado convergência das medidas obtidas (Howard, Conway & Maxwell, 1985; Marsh, 1982). Os estudos revelam correlações fortes entre as avaliações feitas por alunos actuais e antigos alunos (Howard, Conway & Maxwell, 1985; Overall & Marsh, 1980) correlações moderadas entre as avaliações e o desempenho docente eficaz (Feldman, 1989a) e entre o rendimento académico e o desempenho docente eficaz (Cohen, 1981), e correlações positivas mas com baixo significado estatístico, entre as avaliações feitas pelos alunos e pelos docentes (Blackburn & Clark, 1975; Braskamp, Caulley & Costin, 1979 in Ory, 2001), entre alunos e administradores acerca da eficácia docente (Feldman, 1989a), e entre avaliações dos alunos e a avaliação dos colegas e *portfólios* (Centra, 1993 in Ory, 2001; Seldin, 1993).

Durante a década de 90, a investigação não conheceu um aumento de estudos no domínio, contudo as contribuições de estudos prévios resolveram em boa medida as preocupações mais importantes. Assim, a revisão dos trabalhos feitos tem, de uma forma consistente, mostrado que as avaliações feitas a partir da recolha de informação junto dos alunos são no seu conjunto i) multidimensionais, ii) fiáveis e estáveis, iii) dependem mais do docente do que da disciplina, iv) são relativamente válidas e pouco afectadas pelos enviesamentos potenciais (e.g. *efeito de leniência*, carga de trabalho, tamanho da turma e interesse prévio da disciplina), e v) consideradas úteis pela apresentação de estratégias específicas de intervenção ao serviço da melhoria do ensino (Marsh & Bailey, 1993).

A fiabilidade ou rigor dos questionários de avaliação do ensino constitui uma pré-condição da validade e é determinada por estudos que procuram avaliar o consenso/acordo entre as respostas dadas pelos alunos na mesma disciplina (Gillmore, Kane & Naccarato, 1978; Marsh, 1987). Os estudos em torno desta matéria têm permitido concluir pela existência de uma relação significativa entre o número de alunos que avaliam o docente numa determinada disciplina e os índices de fiabilidade obtidos (Marsh & Roche, 1997). A este propósito, Marsh (2001) indica que os índices de consistência tendem a ser de .95 para 50 alunos, .90 para 25 alunos, .74 para 10 alunos e .60 para 5 alunos. Para Cashin (1995), os índices de consistência tendem a ser de .91 para 40 alunos, .88 para 30 alunos, .83 para 20 alunos e de .69 para 5 alunos, recomendando

por isso, 10 alunos no mínimo, para um nível aceitável de .70. Face à importância deste aspecto importa, no desenvolvimento de estudos de precisão da medida, reunir grupos significativos de alunos em termos amostrais.

No que concerne à estabilidade dos resultados encontrados nos questionários, os estudos multisectoriais apresentam uma correlação negativa entre os resultados nas avaliações do ensino e a idade e anos de experiência do docente (Feldman, 1983). Parece existir a ideia de que essas avaliações tenderão a ser mais favoráveis à medida que existe progressão na carreira, e, em particular, a partir dos primeiros anos de leccionação. Contudo, estes estudos fornecem informação limitada para inferir como as avaliações do docente mudam ao longo do tempo. Marsh e Hocevar (1991b), num estudo longitudinal, procuraram avaliar as mudanças do desempenho de 195 docentes avaliados ao longo de 13 anos. O estudo concluiu pela não existência de diferenças substantivas nas dimensões abarcadas no questionário utilizado *Student's Evaluation of Educational Quality* (SEEQ), na generalidade das sub-amostras estudadas em função do grau de experiência de ensino (pouco, média ou muita) sinalizado no pré e pós-teste. Por sua vez, um estudo conduzido por Rego (2003), a partir das percepções de diferentes elementos: docentes universitários (que foram convidados a avaliarem os seus antigos mestres); pessoas com curso superior - portugueses e brasileiros - exercendo actividade profissional em organizações de diversos tipos (que opinaram sobre os seus antigos docentes) e alunos (que se debruçaram sobre os actuais docentes), constatou a convergência de opiniões e reacções, contribuindo, por um lado, para a evidência empírica de um conjunto de características identificadoras da profissionalidade docente atemporal. Por outro lado, esta evidência inibe a argumentação segundo a qual os alunos perfilham perspectivas enviesadas da vida pedagógica carecendo de maturidade e responsabilidade para ajuizarem as potenciais fontes de melhoria da qualidade do ensino.

Em termos da capacidade de generalização, os estudos multisectoriais têm ainda mostrado um consenso satisfatório entre as respostas dadas pelos alunos actuais e alunos graduados (Centra, 1979, 1993; Marsh, 1987; Rego, 2003). Num estudo longitudinal, Overall e Marsh (1980) avaliaram cerca de 100 turmas e verificaram a existência de correlações de .83 entre os mesmos alunos em períodos temporais distintos: durante o curso e anos mais tarde (pelo menos um ano após o término do curso). Estes estudos permitem concluir pela existência de um padrão semelhante de respostas nos questionários em presença de alunos em graduação ou já graduados, o que coloca em causa o “mito” de que os alunos dos 1º anos não apresentam a

capacidade para fazer uma avaliação adequada do docente, enquanto não terminarem o curso (Aleamoni, 1999).

Por sua vez, os investigadores procuraram averiguar em que medida os questionários se correlacionam em várias disciplinas leccionadas pelo mesmo docente ou por vários docentes na mesma disciplina. Marsh (2001) indica que na avaliação global docente, a correlação entre os docentes da mesma disciplina é de -0.5 (isto é, efeito da disciplina), enquanto que as correlações para o mesmo docente em diferentes disciplinas é de .61, e superior em duas componentes da mesma disciplina (.72). Tais resultados indicam que a avaliação é função do docente que ensina e não da disciplina leccionada, suportando a legitimidade do instrumento como avaliador do desempenho dos docentes. Por último, autores como Gilmore, Kane e Naccarato (1978) sugerem que as avaliações para um determinado docente devem traduzir a média das várias disciplinas leccionadas³¹. Assim, parece claro que para um docente que lecciona há muitos anos, a tomada de decisões sobre a sua progressão na carreira deve basear-se no maior número possível de avaliações de disciplinas leccionadas. Por sua vez, a capacidade de generalização das avaliações em termos institucionais, pressupõe a organização de um arquivo longitudinal das avaliações (e.g. *portfólio*) para efeito de decisões administrativas. Desta forma, estes dados fornecerão base de trabalho útil para a afectação de docentes a determinadas disciplinas, para a redacção de relatórios pedagógicos mais coerentes e para a legitimação de mudanças curriculares ao nível dos cursos, por exemplo. Contudo, constata-se que a maioria das Universidades recolhe sistematicamente estes dados, mas falha na manutenção de um arquivo longitudinal dos mesmos (Marsh, 2001).

Esta capacidade de generalização permite a definição de perfis de desempenho tal como foi avançado por Marsh e Bailey (1993) ao indicarem que cada docente possui um perfil idiossincrático baseado nos resultados obtidos nas várias dimensões/facetas identificadas no SEEQ (e.g. o docente pode obter um score elevado na dimensão organização/clareza e um score baixo na dimensão entusiasmo). Com efeito, num estudo com uma amostra de docentes ao longo dos 13 anos, foi possível determinar as características idiossincráticas do seu perfil, através do cálculo da média em todas as disciplinas leccionadas. Estes resultados têm implicações importantes que permitem suportar a ideia de multidimensionalidade das competências avaliadas nos questionários

³¹ Na perspectiva de Naccarato é de pelo menos 5, sugerindo por sua vez Cashin (1995) pelo menos 2 ou mais disciplinas num período temporal mínimo de 2 anos.

Reservas interpretativas devem ser colocadas às avaliações com amostras inferiores a 15 alunos, ou a um número inferior a 65% dos questionários de uma determinada turma (Cashin, 1995 in Kaufman, 2002).

e a capacidade de generalização da informação obtida na tomada de decisões pessoais e institucionais.

Na linha do reforço da generalização dos questionários, emergem estudos realizados em várias partes do mundo e que suportam a possibilidade da sua transferência para diferentes contextos, embora sem prejuízo da realização de estudos prévios³² (Marsh, 1981). Assim, os resultados a partir da análise multimétodo sugerem que os alunos diferenciam docentes com base nas diferentes dimensões do SEEQ e que essas diferenças se generalizam a outros contextos. Ao mesmo tempo, as respostas obtidas, à excepção na dimensão carga de trabalho e grau de dificuldade, tendem a diferenciar de uma forma consistente melhores e piores docentes (cf. Marsh & Roche, 1992). A aplicabilidade do SEEQ a diversos contextos foi investigada criticamente por Watkins (1994) à luz das teorias da Psicologia Cultural, tendo concluído que este modelo de questionário contempla dimensões associadas ao ensino eficaz universalmente aceites em contextos multiculturais *“the results are certainly generally encouraging regarding the range of university settings for which the questionnaires and the underlying model of teaching effectiveness investigated here maybe more appropriate”*.

Todavia, o constructo ensino dificilmente é validado a partir de um critério único de avaliação. Historicamente, a evidência empírica baseou-se essencialmente na aprendizagem do estudante como critério para aferir a qualidade do ensino, e que aliás terá sido esta visão redutora que contribuiu para a inibição de outras formas mais compreensivas de avaliação. Apesar da aprendizagem do aluno constituir o critério mais aceite, outros critérios emergem: auto-avaliações dos docentes, colegas, administradores, frequência de comportamentos específicos sinalizados por observadores externos, etc. Por sua vez, decorre a necessidade de se utilizar instrumentos de medida que possam integrar de forma parcimoniosa indicadores/dimensões comportamentais que a literatura assume como importantes para a definição da eficácia do ensino. Para além disso, importa que os itens de tais instrumentos sejam susceptíveis de serem facilmente avaliados pelos alunos e possuírem a capacidade de discriminação/tipificação das áreas fortes e menos fortes do ensino (Marsh & Bailey, 1993; Marsh & Hocevar, 1991a). Apesar dos problemas metodológicos inerentes aos estudos multisetoriais, uma meta-análise sobre a validade dos questionários mostrou que as disciplinas avaliadas como tendo um ensino mais efectivo foram também aquelas que apresentaram melhores classificações ou rendimento académico. Por sua vez, estudos neste

³² Estudos desenvolvidos na Austrália (Hayton, 1993); Espanha (Marsh, Tournon & Wheeler, 1985); Nova Guiné, (Clarkson, 1984); Nova Zelândia (Watkins, Marsh & Young, 1987); Índia (Watkins & Akande, 1992); Nepal, (Watkins & Regmi, 1992); Nigéria (Watkins, Marsh & Akand, 1992); Hong Kong (Watkins, 1992) in Marsh, 2001.

domínio apontam para a adequabilidade de aplicar o instrumento no final do semestre, quando os alunos têm noção global do processo de ensino/aprendizagem (Frey, 1976 in Ory, 2001; Marsh & Overall, 1980). A este propósito, Cohen (1987) tomou em linha de conta 41 estudos, e concluiu que a média das correlações entre o rendimento e as diferentes dimensões do SEEQ era significativa para as seguintes dimensões: estrutura (.55), interação (.52), competência (.50), avaliação global da disciplina (.49), avaliação global docente (.45), aprendizagem (.39), relacionamento (.32), avaliação (.30) e feedback (.28), à excepção do interesse/motivação (.15) e do grau de dificuldade (-.04). Por sua vez, Feldman (1989b) alargou esta pesquisa procurando mostrar que muitas das dimensões teorizadas por Cohen resultam de dimensões mais específicas e com um grau de associação distinto com o rendimento académico. Por exemplo, na análise feita por Feldman, a dimensão competência é representada por 3 sub-dimensões que se correlacionavam com o rendimento: .34 (conhecimento das matérias), .56 (compreensão e clareza) e .30 (consideração do nível e progresso da turma). Contudo, Cohen (1987 in Marsh, 2001) refere que as correlações eram superiores quando as dimensões específicas dos questionários eram medidas com escalas multidimensionais do que através de escalas com itens isolados. Este conjunto de resultados confirma que os questionários em uso na avaliação dos docentes tendem a correlacionar-se com a aprendizagem dos alunos (Kaufman, 2002). Com efeito, a evidência da estabilidade e a generalização dos resultados obtidos do docente em várias disciplinas e o consenso encontrado entre alunos mais novos e mais velhos nos questionários podem ser lidos como suportes de validade do próprio constructo. Contudo, a sua validade tem sido criticada pelos estudos sobre as potenciais fontes de enviesamento (*bias*). Estes estudos procuram identificar a influência de factores externos nas avaliações dos alunos, mas têm-se mostrado inconsistentes e pouco fiáveis (Marsh, 2001; McKeachie, 1979) tendo convergido para a produção e perpetuação de duras críticas à validade dos questionários (Aleamoni, 1999)³³, nomeadamente quando apontam a existência de erros inferenciais (e.g. efeito de halo, efeito de leniência e erro de tendência central). Importa, neste domínio, compreender a natureza do viés e a forma como este pode ser controlado (Marsh, 2001). Com efeito, a investigação tem indicado que a variável tamanho da turma é aquela que tende a revelar uma associação com maior significado estatístico com as dimensões interação da turma e relacionamento individual, quer nas avaliações dos alunos quer nas auto-avaliações dos docentes.

³³ Artigo que compila a história de pelo menos 16 mitos associados à avaliação do ensino pelos alunos desde 1924 a 1988.

Por sua vez, e como resposta às críticas lançadas ao *efeito Dr. Fox*, Marsh e Ware (1982) defendem que, nas organizações educacionais que habitualmente incentivam o sucesso académico dos alunos, este efeito não é significativo. Assim, (i) a expressividade docente apenas tende a afectar a dimensão entusiasmo do docente (aliás aspecto mais associado a essa manipulação); (ii) e que a abrangência dos conteúdos das matérias tende a ser mais afectada pela dimensão do conhecimento científico do docente e pela dimensão organização/clareza dos aspectos curriculares (mais uma vez, aspectos mais associados a essa manipulação). Assim, e na observância das variáveis que podem constituir factores de risco da validade dos resultados obtidos, importa mais uma vez considerar a multidimensionalidade na apreciação dos resultados decorrentes das avaliações. Desta forma, um efeito que supostamente pode ser lido como viés deve ser analisado e interpretado como um dos aspectos que pode fazer parte de uma avaliação mais abrangente do ensino. Finalmente, McKeachie (1979) conclui que, de facto, existem variáveis que podem potenciar o viés, contudo esse efeito é pouco significativo, e por isso, a utilização de análises estatísticas (e.g. regressão múltipla e equações estruturais) para controlo desses enviesamentos não tem sido considerada.

2.7. Em síntese...

As mudanças ocorridas nos últimos anos no Ensino Superior em Portugal têm marcado de uma forma incontornável, as formas e contextos da avaliação. Com efeito, a tendência de diminuição do número de candidatos num momento em que a qualidade dos serviços oferecidos passa a ser um critério de escolha para os estudantes tem justificado as preocupações havidas com a avaliação da satisfação dos alunos sobre a qualidade do ensino e do desempenho docente em particular. Uma das formas mais documentadas na literatura especializada é a avaliação do ensino pelos alunos. Aliás, esta avaliação tem granjeado suporte empírico suficiente e daí o seu aumento exponencial em muitas universidades nacionais e estrangeiras. Assim, as avaliações feitas a partir da recolha de informação junto dos alunos são no seu conjunto, multidimensionais, fiáveis e estáveis, dependem mais do docente do que da disciplina, são relativamente válidas e pouco afectadas pelos enviesamentos potenciais e consideradas úteis pela apresentação de estratégias específicas de intervenção ao serviço da melhoria do ensino. Por sua vez, a legitimidade da avaliação do ensino pelos alunos ganha força ao defender-se que os alunos são os primeiros

agentes da aprendizagem e, como clientes, têm o direito de participar, em certa medida, na avaliação do ensino e avaliar o desempenho do docente por referência ao seu comportamento no trabalho (mais ao que faz do que ao que pode fazer, isto é, o quão competente é).

Teorias recentes na Psicologia e na Educação definem a aprendizagem como um processo pessoal de construção progressiva de conhecimento em que o docente é cada vez mais assumido como facilitador e orientador da aprendizagem dos alunos. Na procura de um modelo do docente eficaz e competente, a evolução do conceito e das metodologias utilizadas (heterogeneidade de instrumentos utilizados para a avaliação da qualidade do ensino), salientam percursos com incidências diversas. Um olhar atento sobre a literatura permite verificar a existência de um elenco significativo de factores e aspectos associados às práticas e aos comportamentos dos docentes com efeitos positivos ou negativos sobre os alunos numa lógica baseada em concepções cognitivistas, construtivistas e interaccionistas da aprendizagem.

De uma forma genérica, os questionários tipicamente utilizados para avaliar o ensino pelos alunos na maioria das universidades, são de tipo papel-lápis e, obedecendo a um formato de *rating scales*, sendo porém cada vez mais notória a emergência de questionários em formato *on-line*. Como é possível constatar, os questionários aplicados aos alunos têm sido objecto de vários entendimentos por diversos autores. Com efeito, convém considerar as dimensões que procurem avaliar a actividade docente assente numa visão de natureza transdisciplinar e as dimensões de avaliação da qualidade da aprendizagem dos alunos. Assim, a organização dos itens pode ser variada, por exemplo, veja-se a este propósito os questionários nacionais ou internacionais em uso nesta matéria. Para além das dimensões/facetas que já foram identificadas de forma recorrente na literatura, importa a adopção, em regra, do mesmo questionário para todos os docentes de forma a garantir a equidade do tratamento da informação veiculada e o sentido comum do que se entende por ensino de qualidade (Abrantes & Valente, 2000). No entanto, existe um certo consenso em considerar que um questionário universal e único não responde às necessidades de todas as unidades de ensino (Bernard, 1992). Assim, no caso de disciplinas que pela sua especificidade o exijam, deverão ser introduzidos ajustamentos ao questionário. Para além destes aspectos, os alunos deverão ser igualmente inquiridos sobre algumas variáveis de contexto – biográficas (e.g. estatuto de trabalhador-estudante, estatuto de inscrição na disciplina, percentagem de frequência às aulas, sexo, etc.). Da teoria revista, sobressai a ideia de que a par do questionário institucional standardizado, de natureza fechada, importa progressivamente, a nível mais local (e.g. escolas/departamentos/faculdades) desenvolver dispositivos de avaliação que contemplem

uma noção mais abrangente das condições de ensino/aprendizagem, procurando para o efeito, incluir itens de natureza mais qualitativa e de resposta aberta (e.g. descrição de comentários positivos e negativos).

Em termos da sua administração, importa que esta possa ocorrer no período anterior ao final do semestre/ano (por exemplo, duas semanas antes do final das aulas), bem como através da sensibilização dos docentes e informando os alunos dos propósitos da avaliação. Aliás, estudos têm mostrado que, quando existe a explicitação da importância da avaliação bem como a associação dos questionários a objectivos de natureza formativa para efeitos de promoção, os resultados são tendencialmente mais favoráveis (Centra, 1976; Feldman, 1989a; Overall & Marsh, 1979 in Ory, 2001). Por sua vez, a identificação do aluno tende a produzir avaliações mais favoráveis comparativamente à avaliação feita de forma anónima (*ibidem*), contudo, tem-se julgado necessário, manter o anonimato do aluno pelo receio de repercussões negativas decorrentes dessa avaliação. Por último, importa na interpretação dos resultados, procurar colocar reservas a avaliações que apresentam um número reduzido de questionários.

Parte II

CAPÍTULO 3

Metodología de estudio empírico

“The significant problems we face cannot be solved at the same level of thinking we were at when we created them.”

Albert Einstein

3.1. Introdução

Neste capítulo, é apresentada a metodologia geral do estudo de adaptação e validação de um instrumento multidimensional da qualidade do ensino tomando as percepções dos alunos no Ensino Superior. O instrumento designado *Percepções do Ensino pelos Alunos (PEA)*, inspirado na teoria e investigação anteriormente revistas nos capítulos anteriores, procura ultrapassar algumas das limitações apontadas à investigação precedente, e nele incluir um conjunto de dimensões/facetas associadas a um ensino eficaz e de qualidade.

O trabalho empírico da tese, foi desenvolvido rentabilizando as actividades no âmbito da avaliação do ensino desenvolvidas pelo Gabinete de Avaliação e Qualidade do Ensino (GAQE) da Universidade do Minho³⁴. Dada a inexistência de instrumentos válidos em Portugal e das críticas apontadas aos instrumentos que existem com o propósito de avaliar o ensino (visão fragmentada dos pressupostos que integram uma visão coerente do ensino), procurou-se com a realização deste trabalho, e partindo de uma concepção alargada das questões do ensino/aprendizagem, propor um novo questionário de avaliação dos docentes no Ensino Superior e, neste caso em concreto, da Universidade do Minho.

Neste contexto, começamos por apresentar os objectivos e as hipóteses que nortearam a investigação bem como os contornos metodológicos e as opções havidas na prossecução do estudo empírico. Apresentam-se, ainda, nos pontos seguintes deste capítulo, a caracterização da amostra e as suas variáveis mais significativas, a descrição do instrumento utilizado e os procedimentos considerados na recolha dos dados.

³⁴ Assumido como vector estratégico na consolidação do prestígio e credibilidade da Universidade do Minho, compete a este Gabinete, entre outras funções, garantir a promoção, coordenação e execução de todos os procedimentos associados à avaliação do desempenho institucional na vertente do ensino e formação (Resolução no Diário da República, II Série, n.º 56/2004, de 21 de Maio).

3.2. Objectivos, questões e hipóteses de estudo

A presente investigação pretende contribuir para a adaptação e validação de um instrumento de avaliação do ensino pelos alunos e justificar a sua utilização no contexto do Ensino Superior em Portugal, em geral, e na Universidade do Minho, em particular. Mais especificamente, procura-se disponibilizar informações consideradas relevantes de um dispositivo de avaliação do desempenho docente, que procura ir mais além que cumprir exclusivamente objectivos sumativos, para passar a incluir também uma visão de natureza mais formativa e de desenvolvimento profissional dos docentes, susceptível de ser integrada num processo mais vasto de avaliação institucional da Universidade do Minho. Com efeito, o instrumento apresentado traduz, a nosso ver, as dimensões fundamentais do exercício da profissão docente em função das quais o próprio docente se poderá auto-avaliar e auto-desenvolver. Partindo de uma concepção mais alargada da natureza do ensino mais tradicional e da emergência de um conceito ensino/aprendizagem mais centrado no aluno à luz dos novos desafios colocados pela Declaração de Bolonha, entendemos que o instrumento em estudo, representa um compromisso sério entre estes dois paradigmas, que actualmente, coexistem no contexto do Ensino Superior.

Assim, procuramos que a validação do instrumento na sua versão para os alunos, possa contribuir para o desenvolvimento, posterior, de uma versão paralela do instrumento de auto-avaliação do docente. Acreditamos que desta forma mais plural e abrangente será possível ultrapassar algumas das críticas mais apontadas à avaliação do ensino e que tendem a relacionar-se com uma visão unilateral e reducionista do processo avaliativo. O trabalho de adaptação e validação do PEA vai-nos permitir disponibilizar uma fonte informativa importante do trabalho docente, que importará no futuro rentabilizar em complemento com outras fontes informativas sobre as práticas do ensino, quer os próprios docentes quer os órgãos de gestão pedagógica. Mais ainda, o contributo deste trabalho poderá situar-se ao nível das potencialidades que o próprio instrumento pode oferecer ao serviço da avaliação e do planeamento estratégico de medidas/programas de intervenção que promovam o aperfeiçoamento e a melhoria do ensino. Por último, e dada a similitude do instrumento agora estudado com outros utilizados internacionalmente, pretendemos contribuir para o desenvolvimento de estudos de caracterização do ensino universitário que possam permitir algumas análises comparativas das percepções dos alunos sobre as práticas docentes, em diversos países onde os instrumentos similares são

utilizados. Esta ideia parece-nos fundamental num momento em que caminhamos para a convergência do Espaço Europeu em termos de estrutura e configuração dos graus académicos e, quiçá, de procedimentos de avaliação do ensino.

Com efeito, a investigação revista nos capítulos anteriores deste trabalho oferece motivos suficientes para se considerar a importância da avaliação das percepções do ensino pelos alunos, e em particular, o desempenho docente. Em função do estado de desenvolvimento de cada instituição, essa informação pode apoiar a monitorização da qualidade do ensino, tanto na vertente de identificação de situações de excelência como de situações que aconselham uma melhoria das práticas.

Neste quadro são objectivos específicos deste estudo: (i) apreciar as dimensões do questionário usado e a forma como os itens se agrupam nas mesmas dimensões; (ii) sendo apontado que a avaliação dos docentes oscila de acordo com o ano do curso, iremos apreciar o impacto do ano escolar das disciplinas avaliadas nos resultados obtidos (1 e 2º ano, 3º ano, e 4º e 5º ano); (iii) sugerindo a literatura que os resultados obtidos na avaliação oscilam em função da tipologia de aulas leccionadas, procura-se apreciar o eventual impacto das aulas teóricas e das práticas nos resultados; e (iv) dada a sensibilidade diferente de alunos das humanidades e de ciências e tecnologias às questões do ensino e da pessoa do docente, associadas ao próprio género do aluno e docente, este estudo procura, por último, apreciar eventuais diferenças nos resultados tomando 4 áreas de cursos da Universidade do Minho (Ciências e Engenharia; Ciências Económicas; Ciências Sociais e Humanas e Ciências da Educação).

3.3. Amostra

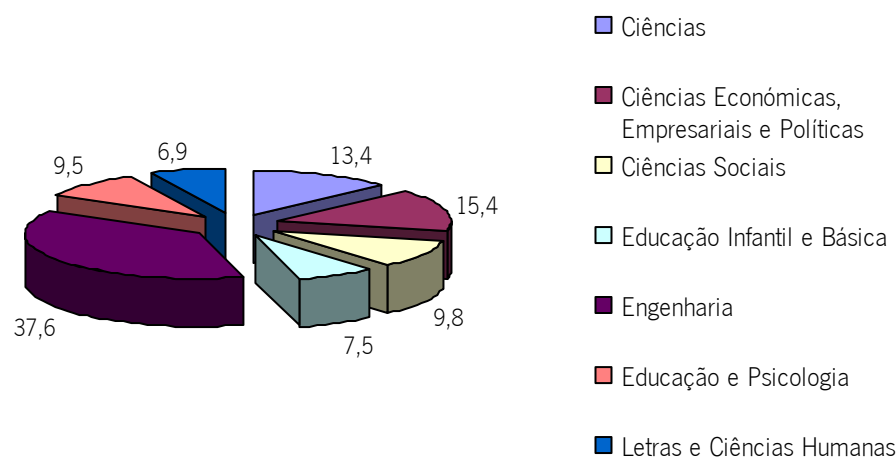
A amostra deste estudo corresponde a um total de 41 962 questionários preenchidos de forma anónima, pelos alunos de 49 cursos de licenciatura e de 6 cursos de Complemento e de Qualificação, num universo de 14 903 inscritos³⁵, no ano lectivo 2003-04. Este número de respondentes representa cerca de 47% da amostra possível de alunos da Universidade do Minho.

³⁵ No ano lectivo 2003-04, a Universidade do Minho possuía 14 751 alunos inscritos dos cursos de licenciatura (número que em termos reais não corresponde ao número de alunos que frequenta assiduamente as aulas), cada um a frequentar simultaneamente cinco ou seis disciplinas com cerca de 1121 docentes envolvidos. Pretendendo abranger todos os alunos e uma discriminação por disciplinas, a informação a processar diria respeito a cerca de 88 506 questionários (cf. Relatório de Actividades da Universidade do Minho, 2003).

Os questionários recolhidos para efeito deste estudo, integram a avaliação das disciplinas anuais e semestrais do 2º semestre dos planos de estudo dos cursos. De salientar que, em muitos dos casos, os docentes foram avaliados mais do que uma vez dado poderem ser docentes em mais do que uma disciplina, e cada aluno avaliou todos os docentes das disciplinas que frequenta. No sentido de contornar algumas das limitações encontradas em estudos prévios apontadas na literatura, a metodologia adoptada na construção da amostra procurou, por um lado, responder à preocupação em obter uma representatividade robusta de alunos no sentido de viabilizar as análises estatísticas, e, por outro, explorar eventuais diferenças encontradas nas disciplinas de várias áreas de cursos.

Os questionários dos alunos nos diferentes cursos foram organizados em função dos Conselhos de Cursos³⁶ a que estão afectos (Gráfico 3.1).

Gráfico 3.1- Conselhos de Cursos



Como poderemos observar (Gráfico 3.1), a amostra encontra-se relativamente equilibrada quanto às áreas de cursos a que os questionários recolhidos dizem respeito. De referir, contudo, o predomínio de questionários dos alunos pertencentes ao Conselho de Cursos de Engenharia (37.6%), ao que se seguem os questionários do Conselho de Cursos de Ciências Económicas

³⁶ A Universidade do Minho organiza as suas actividades, desde a sua criação, de acordo com o chamado “modelo de gestão matricial”. Este modelo baseia-se na interacção entre os projectos e as unidades orgânicas e de recursos. Os projectos são: o ensino, a investigação e os serviços especializados à comunidade. Esta estrutura organizacional proporciona um enquadramento muito flexível para a interdisciplinaridade científica e, funcionando, aumenta o uso efectivo dos recursos disponíveis.

Empresariais e Políticas (15.4%) e do Conselho de Cursos de Ciências (13.4%). Com valores muito próximos encontram-se o Conselho de Cursos de Ciências Sociais (9.8%), Conselho de Cursos de Letras e Ciências Humanas (6.9%) e o Conselho de Cursos de Educação Infantil e Básica (7.5%). Por último, dos 41 962 questionários, (9.5%) foram recolhidos junto dos alunos de dois cursos integrados no Conselho de Cursos de Educação e Psicologia.

Em face da organização da amostra por Conselhos de Cursos, consideraram-se os questionários dos alunos dos seguintes cursos (Quadro 3.1).

Quadro 3.1 – Questionários recolhidos por licenciaturas agrupadas nos Conselhos de Cursos (N=41 962)

Conselho de Cursos	Curso	N
<i>Ciências</i>	Biologia Aplicada	574
	Biologia e Geologia (Ensino)	487
	Física	169
	Física e Química (Ensino)	901
	Geologia - Recursos e Planeamento	195
	Matemática	925
	Matemática Aplicada	55
	Matemática e Ciências da Computação	498
	Optometria e Ciências da Visão	1122
	Química	47
	Química - Ramo Qualidade de Materiais Têxteis	287
Química - Ramo Qualidade de Materiais Plásticos	346	
<i>Ciências Económicas, Empresariais e Políticas</i>	Administração Pública	757
	Direito	2128
	Economia	1039
	Gestão	1287
	Relações Internacionais	991
	Relações Internacionais Económicas e Políticas	92
	Negócios Internacionais	156
<i>Ciências Sociais</i>	Comunicação Social	1068
	Geografia e Planeamento	704
	História (Ensino)	292
	História – Ramo Científico	303
	Arqueologia	456
	Sociologia	1281
<i>Educação e Psicologia</i>	Educação	1249
	Psicologia	2744
<i>Engenharia</i>	Engenharia Biológica	1281
	Engenharia Biomédica	399
	Engenharia Civil	3794
	Engenharia de Comunicações	214
	Engenharia Electrónica e Computadores	1804
	Engenharia de Materiais	870
	Engenharia Mecânica	226
	Engenharia de Polímeros	990
	Engenharia de Gestão Industrial	411
	Engenharia de Sistemas e Informática	2566
	Engenharia Têxtil	387
	Engenharia do Vestuário	101
	Informática de Gestão	1889
	Arquitectura	859
<i>Letras e Ciências Humanas</i>	Inglês e Alemão (Ensino)	378
	Português (Ensino)	729
	Português e Alemão (Ensino)	185
	Português e Francês (Ensino)	481
	Português e Inglês (Ensino)	549
	Línguas Estrangeiras Aplicadas	567
<i>Educação Infantil e Básica</i>	Educação de Infância	1320
	Ensino Básico do 1º Ciclo	1095
	*Educação de Infância - Domínio de Educação para a 1ª Infância	133
	*Administração Escolar e Administração Educacional	83
	*Educação de Infância - Domínio de Expressão e Educação	26
	*Educação Especial - Problemas Graves	95
	*Educação Especial - Problemáticas de Risco	91
	*Ensino Básico (1º Ciclo), domínio de Língua Portuguesa	86

* Cursos de Complemento e Qualificação do Instituto de Estudos da Criança (IEC).

No quadro 3.2 apresentamos uma análise da amostra de questionários recolhidos considerando conjuntamente o ano em que se inscrevem as disciplinas avaliadas e os Conselhos de Cursos a que pertencem os alunos respondentes. De salientar que 54.3% dos questionários foram recolhidos junto dos alunos do 1º e 2º ano; 22.9% junto dos alunos do 3º ano e 22.8% junto dos alunos do 4º e 5º ano.

Quadro 3.2 – Questionários por Conselhos de Curso e anos das disciplinas (N=41 962)

Conselho de Cursos	Ano	N	%
<i>Ciências</i>	1º e 2º ano	3000	53,5
	3º ano	1347	24,0
	4 e 5º ano	1259	22,5
	Total	5606	100,0
<i>Ciências Económicas, Empresariais e Políticas</i>	1º e 2º ano	3577	55,5
	3º ano	1346	20,9
	4 e 5º ano	1527	23,7
	Total	6450	100,0
<i>Ciências Sociais</i>	1º e 2º ano	2302	56,1
	3º ano	1033	25,2
	4 e 5º ano	769	18,7
	Total	4104	100,0
<i>Educação Infantil e Básica</i>	1º e 2º ano	1963	62,7
	3º ano	830	26,5
	4 e 5º ano	336	10,7
	Total	3129	100,0
<i>Engenharia</i>	1º e 2º ano	8389	53,1
	3º ano	3384	21,4
	4 e 5º ano	4018	25,4
	Total	15791	100,0
<i>Educação e Psicologia</i>	1º e 2º ano	1878	47,0
	3º ano	1041	26,1
	4 e 5º ano	1074	26,9
	Total	3993	100,0
<i>Letras e Ciências Humanas</i>	1º e 2º ano	1673	57,9
	3º ano	639	22,1
	4 e 5º ano	577	20,0
	Total	2889	100,0

3.4. Instrumento

O questionário denominado *Student's Evaluation of Educational Quality* (SEEQ; Marsh, 1987, 1991, 2001) tem sido objecto de vários estudos ao longo de mais de duas décadas e em contextos académicos distintos, tendo-se revelado fiável, válido e com uma capacidade significativa de generalização dos resultados obtidos (Marsh & Roche, 1995). Aliás, a evidência teórica e empírica, acumulada em torno deste instrumento, revela que ele apresenta qualidades métricas satisfatórias, quer no que se refere à sua dimensionalidade (replicação de numerosas análises factoriais, em particular a utilização da análise factorial confirmatória (Feldman, 1976; Marsh & Hocevar, 1991b), quer aos valores de consistência interna.

O instrumento encontra-se disponível em duas versões: alunos (A) e docentes (D)³⁷. Trata-se de um questionário de auto-relato que procura avaliar a qualidade do ensino com base nas percepções dos alunos e docentes nas instituições de Ensino Superior. De salientar que, na fase de adaptação, tivemos a preocupação de traduzir os itens no sentido de replicar o modelo original, tendo-se apenas procedido à reformulação de alguns itens³⁸ bem como à inclusão de novos elementos justificados, quer pela especificidade da informação que acrescentam, quer pela possibilidade de cruzamento de informação e de validade dos resultados que tais itens podem assegurar³⁹. Por exemplo, não contendo uma dimensão reportada aos aspectos logísticos/administrativos (assiduidade, pontualidade, cumprimento de prazos, entrega de sumários...), julgamos que a mesma deve existir num questionário para a Universidade do Minho do Minho. Com efeito, atitudes e comportamentos dos docentes “latinos” tendem a não se verificar, por exemplo, em países de tradição anglo-saxónica.

O questionário foi traduzido para a sua utilização em Portugal em íntima colaboração com um grupo de trabalho criado para proceder ao exercício de adaptação e validação de um instrumento de avaliação da qualidade do ensino a utilizar na Universidade do Minho. Refira-se que, neste trabalho de adaptação foram auscultados especialistas na área e analisados questionários congéneres utilizados em Universidades estrangeiras bem como a bibliografia disponível neste domínio. Adicionalmente, procedeu-se à reflexão falada do questionário junto de

³⁷ Nesta fase do nosso trabalho apenas trabalhamos uma versão do questionário para alunos.

³⁸ O Item 3 da dimensão Empenhamento Docente: “*O docente utiliza o humor nas aulas*” foi substituído por: “*A forma como o docente organizou as aulas cativou o seu interesse*”.

³⁹ Percentagem de frequência às aulas; tipologia das aulas da disciplina, cumprimento de horários e outras actividades programadas por parte do docente e a percepção da importância da disciplina no curso.

um pequeno grupo de alunos. Deste processo foram retiradas algumas sugestões consideradas pertinentes para a construção de uma versão mais adequada à realidade do contexto do ensino superior em Portugal.

Constituído na versão experimental, por 37 itens numa escala *likert* de 9 pontos, o questionário passou a designar-se *Percepções do Ensino pelos Alunos* (PEA) e assenta numa conceptualização multidimensional da qualidade do desempenho docente num conjunto de 9 sub-dimensões (i) relevância da aprendizagem (4 itens); (ii) empenhamento docente (4 itens); (iii) organização/clareza (4 itens); (iv) interacção da turma (4 itens); (v) relação docente/aluno (4 itens); (vi), profundidade na abordagem dos assuntos (4 itens); (vii) avaliação/classificações (3 itens); e (viii) trabalhos/leituras (2 itens); (ix) carga de trabalho/dificuldade da disciplina em comparação com outras disciplinas do curso⁴⁰ (4 itens). Este instrumento apresenta, ainda, um conjunto de outros itens diferentes relativos aos docentes e disciplinas, como: a classificação global da respectiva disciplina em comparação com as outras disciplinas do curso; a classificação global do docente em comparação com os outros docentes do curso (Anexo B).

À semelhança do conceito de ensino proposto por Marsh e colaboradores (1987, 1991, 1994), a sua avaliação adequada envolve várias dimensões/facetas. Refira-se que na resposta ao questionário cada aluno deveria avaliar o grau de concordância ou discordância relativamente às afirmações apresentadas, utilizando uma escala de resposta de 9 pontos (1 e 2 – *Discordo completamente*; 3 e 4 – *Discordo*; 5 – *Nem concordo nem discordo*; 6 e 7 – *Concordo*; 8 e 9 – *Concordo completamente*).

A opção pela adaptação e validação deste instrumento decorreu da ênfase numa abordagem comportamental que procura enfatizar o registo de comportamentos observados; e da formulação dos itens numa linguagem clara e objectiva assentando em dimensões educacionais transdisciplinares potencialmente válidas num conjunto diversificado de contextos de ensino/aprendizagem. Considerando o processo sequencial da leitura, entendemos ainda que existe neste questionário a vantagem do colocar a situação do aluno em primeiro lugar, parecendo ser preferível ao começo com a focalização do docente por insinuar uma preponderância do modelo tradicional de transmissão unidireccional de conteúdos em aulas presenciais (cf. Comissão Técnica relativa à avaliação do ensino ministrado da Universidade do Minho, 2001).

⁴⁰ Inclui itens tais como a percepção da exigência/carga de trabalho, o grau de dificuldade, o ritmo das aulas e o número de horas de estudo.

No quadro 3.3 descreve-se, de forma mais qualitativa, as 9 dimensões do PEA bem como o número de itens associados. Com efeito, a proposta do autor passa pela identificação das dimensões que constituem o modelo do ensino/aprendizagem, e, em particular, sugere um conjunto de padrões comportamentais que definem o papel do docente no Ensino Superior. A identificação destas dimensões/facetas baseia-se na análise de um repertório de comportamentos/atitudes facilmente observáveis pelos alunos e que é esperado que o docente desenvolva. Esta lista de dimensões não deve ser considerada como uma hierarquia, mas antes como um conjunto de componentes de um todo funcional que se articulam e que favorecem uma avaliação mais completa da qualidade do ensino.

Quadro 3.3 - Descrição das dimensões do PEA e respectivo número de itens

FACTORES OU DIMENSÕES		ITENS
<i>Relevância da Aprendizagem</i>	Define as percepções de sucesso dos alunos relativas à sua participação e envolvimento na disciplina. Esta dimensão revela o interesse pelas matérias e consciência de as terem compreendido.	4
<i>Empenhamento Docente</i>	Esta dimensão é relevante e baseia-se no princípio de que os alunos devem estar motivados para aprender. Espera-se que o docente que consegue cativar o interesse dos alunos através do entusiasmo, dinamismo e energia que imprime nas suas aulas, obtenha pontuações altas nesta dimensão.	4
<i>Organização/Clareza</i>	Define a organização e clareza como dois elementos fundamentais na organização do estudo, na realização de exercícios e nos trabalhos dos alunos. Alunos que percebem o ensino como organizado tendem a envolver-se mais numa aprendizagem activa.	4
<i>Interação da Turma</i>	Define a ideia de que o ensino/aprendizagem traduz um acto social do qual emergem os processos comunicacionais: interacção verbal entre alunos e docente e que se reflecte na forma como as matérias são abordadas e discutidas.	4
<i>Relação Docente/Aluno</i>	Define a tomada de consciência das necessidades e dos ritmos de aprendizagem diferenciados dos alunos e que depende em boa medida da qualidade do contacto entre alunos/docente. Alunos que percebem como relevante esta dimensão tendem a estabelecer com maior frequência e intensidade contacto com o docente dentro e fora da sala de aula.	4
<i>Profundidade na Abordagem dos Assuntos</i>	Define a variedade de perspectivas sobre a realidade e a importância da contextualização das teorias/conceitos que o docente aborda nas aulas. Esta dimensão traduz o potencial de maior clarificação e apropriação do significado das matérias.	4
<i>Avaliação/Classificações</i>	Define a qualidade do feedback recebido dos trabalhos sujeitos à avaliação e o estímulo contingente para orientação no estudo. As percepções de justiça e relevância da avaliação por parte dos alunos encontram-se provavelmente associadas à motivação para aprender.	3
<i>Trabalhos/Leituras</i>	Define a percepção da utilidade/relevância das leituras/trabalhos, incluindo as leituras aconselhadas constituindo oportunidades válidas para o desenvolvimento de competências de assimilação e integração de novos conhecimentos.	2
<i>Carga de Trabalho/Dificuldade</i>	Define a percepção da carga de trabalho o grau de dificuldade e a importância percebidos pelos alunos. Indica que em actividades que os alunos percebem obter facilmente sucesso com um menor esforço, tenderão a desmotivar-se bem como a desvalorizar as matérias. Contudo, não parece existir uma relação linear entre esta dimensão e outros indicadores do ensino eficaz, tais como a aprendizagem do aluno.	4

3.5. Procedimentos

Considerando o objectivo geral do trabalho – validação de um instrumento de natureza multidimensional da qualidade do ensino com base nas percepções dos alunos da Universidade do Minho – optou-se pela realização de um estudo exploratório (Almeida & Freire, 2001), tomando para isso, todos os alunos dos cursos de licenciatura e de complemento/qualificação da referida Universidade. Através de uma informação prévia dirigida a todos os docentes, os alunos, foram informados sobre os objectivos da investigação em curso (reforço junto dos docentes e alunos da validação do questionário ser mais rigorosa quando aplicada ao universo total de alunos) sendo garantida a confidencialidade dos resultados (Anexo C). A administração teve lugar juntamente com a versão institucional que vigorava na Universidade do Minho - Questionário de Avaliação do Ensino Ministrado (QAEM) no final do 2º semestre do ano lectivo 2003-04 e em ambiente de sala de aula. A justificação da sua aplicação neste *timing* relacionou-se essencialmente com a preocupação de, por um lado, otimizar recursos (tempo, pessoas) na sua aplicação em larga escala, e, por outro, atender a um período suficientemente grande de contacto entre docente e alunos, para que os alunos pudessem dispor de um período significativo de convivência pedagógica e possuir um conhecimento razoável do desempenho docente. Por sua vez, e apontando a literatura para os potenciais enviesamentos produzidos pelo efeito das notas/classificações, houve a preocupação dos alunos estarem suficientemente distanciados do período de avaliação, de tal modo que os seus julgamentos não fossem influenciados pelas classificações auferidas. Por último, foi nossa preocupação abarcar uma amostra representativa de alunos no sentido de viabilizar as análises estatísticas e de criar condições mais adequadas para validação do PEA. De salientar que o questionário foi desenhado e processado através do programa informático de leitura óptica *TELEform Designer*, compatível com o programa *Statistical Package for the Social Sciences* (SPSS; versão 12.5 para Windows).

CAPÍTULO 4

Resultados obtidos

“The research on student evaluations, like than on other teacher evaluation methods, shows significant tendencies, but no certainties.”

(Centra, 1993 in Ory, 2001)

4.1. Introdução

O presente capítulo apresenta e analisa os resultados do estudo empírico acerca das percepções do ensino pelos alunos dos cursos da Universidade do Minho no 2º semestre do ano lectivo 2003-04, com base na administração de um questionário para o efeito. Embora o instrumento seja usado à escala mundial para efeitos de avaliação da qualidade do ensino, nunca foi objecto de estudos sistemáticos de validação em Portugal (sem prejuízo de ter sido utilizado para efeitos de Avaliação Interna do Curso de Psicologia na Universidade do Minho e no âmbito da investigação noutras Universidades do País). Assim, este estudo teve como preocupação avaliar o seu funcionamento no contexto mais global da Universidade do Minho.

Ao longo do capítulo, procuramos fazer uma análise qualitativa e quantitativa do questionário - *Percepções do Ensino pelos Alunos* (PEA). A análise quantitativa considera a variabilidade dos itens, a sua consistência interna e os resultados em função da natureza prática ou teórica das aulas e do ano curricular a que pertencem as disciplinas cuja docência foi avaliada. Adicionalmente, apresentamos os resultados das qualidades psicométricas do instrumento de medida utilizado, fornecendo elementos para apreciação da sua validade (análise factorial). A par das análises tomando a amostra global dos questionários recolhidos consideramos também os resultados obtidos em função dos cursos (Conselhos de Cursos), da tipologia de aula (teórica e prática) e do ano curricular (1º e 2º ano, 3ºano, e 4º e 5º ano) a que pertencem as disciplinas avaliadas. A análise do impacto destas variáveis de agrupamento dos alunos será retomada, depois, através da análise de variância dos resultados por dimensões/factores considerando o ano e a tipologia de aulas.

4.2. Análise dos itens

A análise dos itens seguiu uma metodologia qualitativa e quantitativa. Em termos da análise qualitativa importa referir os cuidados com a tradução, a auscultação de especialistas da avaliação e docentes universitários com responsabilidades pedagógicas nos cursos. Também se ouviu a opinião de alunos quanto à compreensão e relevância dos itens.

Os resultados da análise da distribuição das respostas aos 37 itens que integraram a versão experimental do PEA⁴¹ apresentam-se no Quadro 4.1. Os resultados da análise descritiva reportam-se à variabilidade e às medidas de tendência central da distribuição das respostas aos itens. Dada a natureza ordinal da escala de resposta utilizada no instrumento, os resultados da análise descritiva incluem os seguintes parâmetros estatísticos: valores mínimos e máximos, 1º e 3º quartis e mediana.

Como podemos constatar (cf. Quadro 4.1), os resultados da análise descritiva dos 37 itens do PEA evidenciaram uma distribuição pouco distinta no padrão de respostas do questionário. Com efeito, embora os valores mínimos e máximos nos itens tenham oscilado entre 1 e 9, ou seja os dois valores extremos da escala usada (excepto, o item relativo ao número de horas de estudo que oscilou entre 1 e 4 valores⁴²). Com efeito, os valores de mediana situaram-se em 7 valores na quase totalidade dos itens, à excepção de 3 itens com valor 5 (posição neutra na escala) e que se reportam à percepção do grau de dificuldade, carga de trabalho e ritmo das aulas. Por outro lado, encontramos um valor inferior de mediana no item relativo ao número de horas de estudo (mediana=2). Os valores para o 1º quartil oscilaram entre 1 e 6 valores, com maior prevalência do valor 6. De salientar, ainda que, cerca de 25% dos treze itens (percepções associadas aos trabalhos/leituras, avaliação/classificações, avaliação global do docente e disciplina, assiduidade do docente, exigência/carga de trabalho, dificuldade, ritmo das aulas, e importância da disciplina no curso se situaram no ponto médio da escala (posição avaliativa mais neutra).

⁴¹ Nesta análise foram considerados os 2 itens globais que procuram avaliar a qualidade da disciplina e da docência e 2 itens que foram acrescentados ao questionário original (assiduidade docente e a importância da disciplina no curso). Contudo, quando se procede à análise factorial e à descrição dos resultados das médias e desvio-padrão das dimensões/facetos do PEA, estes itens não são considerados.

⁴² Especificamente a formulação do item relativo ao envolvimento no estudo é a seguinte: (*“Para além das aulas, qual o nº médio de horas semanais de estudo nesta disciplina ao longo do ano/semestre?”*). De referir a existência de 4 níveis de resposta (1=Menos de 1 hora; 2=De 1 a 2 horas; 3=De 3 a 5 horas; 4=6 horas ou mais) (cf. Anexo B).

Quadro 4.1 Análise da distribuição das respostas aos itens do PEA

PEA						
Itens	N	Valor mínimo	Valor máximo	1º quartil	mediana	3º quartil
Aprendizagem 1	41208	1	9	5	7	8
Aprendizagem 2	41181	1	9	6	7	8
Aprendizagem 3	40976	1	9	5	7	8
Aprendizagem 4	40978	1	9	6	7	7
Empenhamento 1	41019	1	9	7	7	8
Empenhamento 2	41033	1	9	6	7	8
Empenhamento 3	40994	1	9	5	7	8
Empenhamento 4	40884	1	9	5	7	8
Organização 1	40883	1	9	6	7	8
Organização 2	40788	1	9	6	7	8
Organização 3	40679	1	9	6	7	8
Organização 4	40799	1	9	6	7	8
Interacção 1	40852	1	9	6	7	8
Interacção 2	40901	1	9	6	7	8
Interacção 3	40871	1	9	6	7	8
Interacção 4	40646	1	9	6	7	8
Relação 1	40457	1	9	7	7	9
Relação 2	40397	1	9	6	7	8
Relação 3	40237	1	9	6	7	8
Relação 4	39170	1	9	6	7	8
Profundidade 1	39472	1	9	6	7	8
Profundidade 2	39463	1	9	6	7	8
Profundidade 3	39463	1	9	6	7	8
Profundidade 4	39564	1	9	6	7	8
Avaliação 1	38490	1	9	6	7	8
Avaliação 2	39861	1	9	6	7	8
Avaliação 3	38490	1	9	5	7	8
Trabalhos 1	38490	1	9	5	7	8
Trabalhos 2	38070	1	9	5	7	8
Avaliação Disciplina*	49382	1	9	5	7	8
Avaliação Docente*	40314	1	9	5	7	8
Assiduidade Docente	39867	1	9	5	7	8
Dificuldade	40335	1	9	4	5	6
Carga de Trabalho	40190	1	9	5	5	7
Ritmo das Aulas	39989	1	9	5	5	7
Nº Horas de Estudo**	39237	1	4	1	2	3
Importância Disciplina	40453	1	9	5	7	8

* Itens de avaliação global; ** Amplitude de variação total de 1 a 4

Valores mais extremados encontram-se no item relativo ao número de horas de estudo (1º quartil situado em 1). Os valores para o 3º quartil oscilaram entre 3 e 8, mas de uma forma genérica nos 8 pontos. De salientar que o item relativo à relação docente/aluno situou-se no ponto máximo da escala. Este conjunto de valores evidencia uma maior concentração das respostas dos alunos junto dos valores mais elevados da distribuição, o que reflecte globalmente um padrão positivo nas suas percepções do ensino e do desempenho dos docentes. Contudo, e apesar das precauções que estes resultados sugerem (menor variabilidade de alguns itens), cremos que tal situação pode suscitar a introdução de mudanças no questionário, nomeadamente a redução do número de pontos da escala original (1 a 9) e incluir, por exemplo, com base nos resultados obtidos, 6 níveis de resposta (3 níveis no sentido negativo e 3 no sentido positivo), sugerindo-se ainda a possibilidade do aluno não emitir opinião, para distinguir esta tomada de decisão dum acto menos ou não consciente de resposta na situação efectiva de falta de informação ou dificuldade efectiva em se posicionar no sentido positivo ou negativo em relação a determinadas questões.

4.3. Consistência interna

De seguida, apresentamos os resultados da análise da consistência interna dos itens, a correlação do item com o total da escala (coeficiente corrigido) e o contributo do item para o *alpha* da dimensão. Mesmo sabendo da multidimensionalidade do questionário, foi nossa opção trabalhar com a globalidade dos itens numa mesma análise (incluindo os 2 itens que foram acrescentados à versão original do questionário).

Quadro 4.2 Análise da consistência dos itens do PEA

	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Cronbach's Alpha if Item Deleted
Aprendizagem 1	222,3	1539,6	,77	,97
Aprendizagem 2	221,7	1560,4	,72	,97
Aprendizagem 3	222,3	1543,1	,74	,97
Aprendizagem 4	222,1	1565,7	,71	,97
Empenhamento 1	221,5	1548,9	,80	,97
Empenhamento 2	222,0	1531,5	,81	,97
Empenhamento 3	222,2	1523,2	,82	,97
Empenhamento 4	222,4	1520,9	,83	,97
Organização 1	222,0	1533,1	,82	,97
Organização 2	221,9	1557,7	,70	,97
Organização 3	221,8	1562,2	,72	,97
Organização 4	222,1	1559,7	,67	,97
Interacção 1	221,9	1548,5	,78	,97
Interacção 2	222,0	1547,3	,79	,97
Interacção 3	221,9	1543,6	,81	,97
Interacção 4	221,9	1546,5	,80	,97
Relação 1	221,5	1561,1	,72	,97
Relação 2	221,7	1548,5	,78	,97
Relação 3	221,7	1547,9	,75	,97
Relação 4	221,6	1565,4	,70	,97
Profundidade 1	221,9	1560,9	,76	,97
Profundidade 2	221,9	1559,5	,78	,97
Profundidade 3	222,0	1556,8	,77	,97
Profundidade 4	221,9	1557,9	,75	,97
Avaliação 1	222,1	1555,1	,73	,97
Avaliação 2	222,2	1557,5	,65	,97
Avaliação 3	222,2	1561,2	,62	,97
Trabalhos 1	222,1	1565,4	,67	,97
Trabalhos 2	222,2	1562,5	,69	,97
Dificuldade *	223,9	1619,2	,23	,97
Carga de Trabalho *	222,9	1642,5	,07	,97
Ritmo das Aulas *	223,0	1635,3	,16	,97
Nº Horas de Estudo *	226,7	1645,3	,12	,97
Assiduidade Docente	221,5	1583,3	,55	,97
Importância Disciplina *	222,3	1580,3	,44	,97

* Itens reportados aos comportamentos do aluno

Verifica-se que os valores obtidos, através do procedimento *Scale Reliability do SPSS*, sugerem um bom índice de consistência interna do questionário, assim como os níveis elevados de correlação entre os itens e o total da escala na generalidade dos itens. Estes dados são consistentes com os valores relativos à consistência obtidos noutros estudos (Marsh, 2001). Os valores obtidos a propósito da análise da consistência interna do questionário mostram-se adequados face aos objectivos da avaliação, sugerindo a possibilidade de assumirmos uma nota global no mesmo (retirando os itens mais reportados aos comportamentos do aluno; cf. 5 itens assinalados com asterisco no quadro 4.2). Por sua vez, os coeficientes de correlação item x total sugerem bons índices de validade interna (poder discriminativo), uma vez que os valores são superiores ao índice crítico de .20 (Almeida & Freire, 2001). Este valor apenas não se obtém nos 3 itens mais associados às percepções de envolvimento pessoal na disciplina (carga de trabalho, ritmo das aulas e o número de horas de estudo). Este facto decorre da avaliação feita através destes itens se reportarem mais aos próprios alunos que ao docente ou à disciplina, sendo por isso interessante salientar esta discrepância nos coeficientes de correlação (validade diferencial dos resultados). Estes itens poder-se-iam agrupar numa nota sobre o envolvimento do aluno na aprendizagem da disciplina em apreço.

Uma das possíveis explicações para a verificação de níveis elevados de consistência obtidos (.97) decorre em larga medida do facto de estarmos perante um questionário com um número significativo de itens ($n=37$). Aliás, e de acordo com Almeida e Freire (2001) estes “artefactos” produzem necessariamente um aumento dos coeficientes de fidelidade, no entanto, não podemos considerar esta a melhor forma de os obter, sob pena de provocar reacções de aborrecimento e estereotipia nas respostas dos alunos. Referem ainda, que o aumento da consistência interna através deste procedimento não irá, necessariamente, provar a validade do questionário. Os níveis de consistência encontrados neste estudo (.97) encontram-se dentro de intervalos de valores muito aceitáveis (De Vellis, 1991, *ibidem*), devendo, por isso, ser considerada a diminuição do tamanho do questionário, ou pensar se necessário, tornar os itens mais heterogéneos na sua formulação e aspectos avaliados. Este procedimento é particularmente válido face à necessidade de validar um modelo de questionário de tipo “*screening*” de formato mais curto, por exemplo, como resposta à preocupação de avaliação institucional das instituições de ensino superior. Associada a esta ideia encontra-se, o pressuposto de que seria possível avaliar com o necessário rigor e acuidade as principais vertentes comportamentais associadas a cada dimensão/faceta com um menor número de itens.

Em seguida, apresentamos uma análise que considera as correlações entre os itens por dimensão, no sentido de se proceder ao levantamento dos itens com coeficientes de correlação significativos que se encontrariam mais sobreponíveis em termos de conteúdo, tomando para efeitos dessa análise, o ano da disciplina e a tipologia de aula. Esta análise devidamente justificada por dimensões/facetadas do PEA, procura salientar eventuais alterações à versão experimental do questionário, com base na preocupação em proceder à diminuição do número de itens por dimensão/faceta, sem prejuízo, de perda de informação relevante. Por último, procura-se ao longo desta análise, atenuar o efeito de estereotipia de resposta, fenómeno mais frequente em questionários relativamente longos e cujos itens apresentem marcada sobreposição de conteúdo. Assim, reportando-nos à primeira dimensão Relevância da Aprendizagem, no quadro 4.3, indicamos os coeficientes de correlação para os seus quatro itens:

- (1.1) Achou as aulas intelectualmente estimulantes
- (1.2) Aprendeu algo que considera útil
- (1.3) A frequência das aulas aumentou o seu interesse pela matéria
- (1.4) Compreendeu e assimilou os conteúdos das aulas

Quadro 4.3 Correlações dos itens na dimensão Relevância da Aprendizagem*

Relevância da Aprendizagem							
Item		it1 x it2	it1 x it3	it1 x it4	it2 x it3	it2 x it4	it3 x it4
Ano	1 e 2ºano	.740	.781	.676	.743	6.83	.717
	3º ano	.754	.802	.686	.755	.678	.719
	4 e 5ºano	.774	.804	.697	.769	.713	.734
Tipologia	T	.747	.795	.680	.748	.693	.720
	P	.752	.788	.681	.752	.684	.721

* Neste e nos quadros seguintes todos os coeficientes de correlação são significativos a $p < 0.01$

Como é possível visualizar a partir da dimensão em análise, cruzando o ano e a tipologia de aula, os itens 1 e 3 encontram-se bastante correlacionados entre si. No caso de se perspectivar uma diminuição do número de itens associados ao questionário, seria de equacionar a possibilidade de eliminar o item 1, dado que os itens 2 e 3 parecem cobrir a avaliação da percepção do carácter estimulante associado às aulas (percepção de utilidade das aprendizagens efectuadas e o aumento de interesse pelas matérias devido à frequência das aulas).

Reportando-nos à dimensão do Empenhamento Docente, no quadro 4.4 indicamos os coeficientes de correlação para os seus quatro itens:

- (2.1) O docente mostrou-se empenhado na leccionação das aulas
- (2.2) O docente imprimiu dinamismo às aulas
- (2.3) O docente expôs a matéria de forma atractiva
- (2.4) O modo como o docente organizou as aulas cativou o seu interesse

Quadro 4.4 Correlações dos itens na dimensão Empenhamento Docente

Empenhamento Docente							
Item		it1 x it2	it1 x it3	it1 x it4	it2 x it3	it2 x it4	it3 x it4
Ano	1 e 2º ano	.807	.764	.738	.884	.849	.908
	3º ano	.819	.778	.759	.892	.859	.914
	4 e 5º ano	.834	.792	.768	.891	.861	.921
Tipologia	T	.789	.754	.732	.888	.850	.911
	P	.828	.782	.756	.886	.855	.913

Nesta dimensão, a generalidade dos itens apresentam correlações significativas do ponto de vista estatístico, no entanto, os itens 3 e 4 conhecem uma correlação mais acentuada. Neste caso, seria de equacionar a possibilidade de retirar o item 3, cuja formulação inclui o termo “*expôs*”, conceito que pode ser contestado à luz dos novos paradigmas de ensino/aprendizagem. Desta forma, a dimensão poderia ficar reduzida para 3 itens, uma vez que o item 4 parece abranger o carácter atractivo na abordagem das matérias, pela forma como organiza as aulas e o seu impacto ao nível do interesse dos alunos.

Relativamente à dimensão Organização/Clareza, no quadro 4.5 indicamos os coeficientes de correlação para os seus quatro itens:

- (3.1) O docente explicou as matérias com clareza
- (3.2) Os materiais de apoio estavam bem organizados e foram disponibilizados atempadamente
- (3.3) Os conteúdos do programa correspondem aos efectivamente leccionados, de modo que os alunos conheçam o rumo das aulas
- (3.4) O ritmo das aulas facilitou o registo de apontamentos

Quadro 4.5 Correlações dos itens na dimensão Organização/Clareza

Organização/Clareza							
Item		it1 x it2	it1 x it3	it1 x it4	it2 x it3	it2 x it4	it3 x it4
Ano	1 e 2º ano	.630	.623	.612	.687	.572	.622
	3º ano	.656	.643	.617	.705	.586	.632
	4 e 5º ano	.726	.685	.641	.761	.642	.656
Tipologia	T	.641	.633	.616	.694	.552	.602
	P	.662	.642	.618	.711	.604	.643

Mais uma vez, nesta dimensão a generalidade dos itens apresenta correlações significativas entre si, porém, os itens 2 e 3 conhecem uma maior correlação. Neste caso, será de

equacionar a possibilidade de eliminar o item 3, dado que, associada à percepção da correspondência entre os conteúdos pré-definidos do programa e os conteúdos leccionados, encontra-se, ainda que mais indirectamente, a organização e a disponibilização atempada dos materiais de estudo; a clareza na explicação das matérias e o registo adequado de apontamentos (assumidos como suportes válidos susceptíveis de garantir a percepção de uma sequencialidade favorável ao ensino/aprendizagem).

Relativamente à dimensão da Interação da Turma, no quadro 4.6 indicamos os coeficientes de correlação para os seus quatro itens:

- (4.1) Os alunos foram encorajados a participar na discussão das matérias
- (4.2) Os alunos foram estimulados a partilhar as suas ideias e conhecimentos
- (4.3) Os alunos foram encorajados a fazer perguntas e obtiveram respostas adequadas
- (4.4) Os alunos foram encorajados a exprimir as suas ideias e /ou a questionar o docente

Quadro 4.6 Correlações dos itens na dimensão Interação da Turma

Interação da Turma							
Item		it1 x it2	it1 x it3	it1 x it4	it2 x it3	it2 x it4	it3 x it4
Ano	1 e 2ºano	.898	.816	.814	.839	.846	.901
	3º ano	.902	.823	.825	.845	.861	.904
	4 e 5ºano	.919	.841	.855	.860	.886	.908
Tipologia	T	.910	.833	.837	.852	.868	.910
	P	.899	.817	.818	.841	.852	.900

Nesta dimensão, os itens revelam correlações bastante consistentes entre si, embora a correlação seja mais saliente entre os itens 3 e 4 e 1 e 2 quando cruzamos o ano da disciplina. Assim, parece não haver diferenciação entre os descritores comportamentais tais como a participação na discussão das matérias e a partilha de ideias e conhecimentos por parte dos alunos dos últimos anos dos cursos. Por sua vez, seria possível equacionar a pertinência de manter 2 ou 3 itens que possam traduzir a existência de uma atmosfera favorável à emergência dos princípios pedagógicos de democraticidade imbuída numa perspectiva construtivista da aprendizagem mais centrada no aluno (partilha de ideias e auto-questionamento). Com base nas considerações feitas, sugerimos a eliminação do item 3, dado que o estímulo à expressão de ideias e de conhecimentos parece conter em si mesmo o estímulo à colocação de perguntas/questões sobre os assuntos abordados nas aulas.

Relativamente à dimensão Relação Docente/Aluno, no quadro 4.7 indicamos os coeficientes de correlação para os seus quatro itens:

- (5.1) O docente mostrou-se cordial na sua relação com os alunos
- (5.2) O docente mostrou preocupação/interesse pelos alunos
- (5.3) O docente fez com que os alunos se sentissem à-vontade para lhe pedirem ajuda/conselhos dentro e fora das aulas
- (5.4) O docente mostrou-se disponível para atender os alunos durante o horário de atendimento ou no final das aulas.

Quadro 4.7 Correlações dos itens na dimensão Relação Docente/Aluno

Relação Docente/Aluno							
Item		it1 x it 2	it1 x it3	it1 x it4	it2 x it3	it2 x it4	it3 x it4
Ano	1 e 2ºano	.833	.781	.653	.834	.707	.743
	3º ano	.858	.810	.683	.863	.742	.772
	4 e 5ºano	.860	.815	.691	.864	.748	.771
Tipologia	T	.837	.782	.669	.841	.719	.755
	P	.848	.801	.668	.849	.725	.756

Nesta dimensão os itens 2 e 3 e 1 e 2 conhecem uma correlação significativa entre si. Assim, seria de equacionar, a possibilidade de reduzir a dimensão para 2 ou 3 itens, procurando manter a ideia da preocupação com os alunos que encerra em si o respeito e a cordialidade (item 2) e a disponibilidade durante o horário de atendimento ou no final das aulas (item 4) que se encontra de alguma forma presente no item 3 (pedido de ajuda/conselhos dentro e fora das aulas). Desta forma, é possível daqui extrair informação acerca do profissionalismo do docente, e em particular, da sua preocupação com os alunos, a partir dos itens 2 e 4.

No que diz respeito à dimensão Profundidade na Abordagem dos Assuntos, no quadro 4.8 indicamos os coeficientes de correlação para os seus quatro itens:

- (6.1) O docente comparou as implicações das diversas teorias/modelos existentes
- (6.2) O docente apresentou o contexto em que apareceram as ideias/conceitos desenvolvidos nas aulas
- (6.3) O docente apresentou outros pontos de vista além do seu, sempre que foi pertinente
- (6.4) O docente abordou desenvolvimentos recentes nas matérias leccionadas

Quadro 4.8 Correlações dos itens na dimensão Profundidade na Abordagem dos Assuntos

Profundidade na Abordagem dos Assuntos							
Item		it1 x it2	it1 x it3	it1 x it4	it2 x it3	it2 x it4	it3 x it4
Ano	1 e 2º ano	.854	.788	.749	.813	.782	.814
	3º ano	.862	.794	.761	.809	.790	.821
	4 e 5º ano	.897	.833	.797	.856	.823	.842
Tipologia	T	.871	.800	.758	.822	.788	.818
	P	.864	.798	.764	.822	.795	.823

Nesta dimensão os itens 1 e 2, 2 e 3 e 3 e 4 apresentam correlações com significado estatístico. Especificamente, neste caso seria de equacionar a possibilidade da redução da dimensão para 2 itens, assumindo apenas os itens 1 e 3, dado que o contexto em que apareceram as ideias/conceitos desenvolvidos nas aulas (item 2), mas sobretudo, o item 4 reportado ao conhecimento actualizado das matérias por parte do docente pode ser de mais difícil aferição por parte dos alunos em algumas áreas e anos dos cursos em análise.

No que diz respeito à dimensão Avaliação/Classificações, no quadro 4.9 indicamos os coeficientes de correlação para os seus três itens:

- (7.1) O feedback/informação sobre os trabalhos sujeitos à avaliação foi útil
- (7.2) Os métodos de avaliação são justos e apropriados
- (7.3) Os diferentes elementos da avaliação (testes, trabalhos, ...) são equilibradamente ponderados na avaliação final

Quadro 4.9 Correlações dos itens na dimensão Avaliação/Classificações

Avaliação/ Classificações				
Item		it1 x it2	it1 x it3	it2 x it3
Ano	1 e 2ºano	.648	.635	.791
	3º ano	.651	.633	.798
	4 e 5º ano	.708	.678	.825
Tipologia	T	.653	.644	.795
	P	.666	.641	.802

Nesta dimensão os itens 2 e 3 conhecem uma correlação significativa. Assim, no sentido, e na lógica de manter 2 a 3 itens por dimensão, seria de equacionar, a possibilidade de manter os itens 1 e 2 e eliminar o item 3, uma vez que para responder com maior objectividade ao mesmo, o aluno poderá não dispor de informação suficiente dado que a aplicação do questionário é anterior à época de exames (dado que pode ser reforçado pela constatação no item em causa, de cerca 25% da amostra se ter-se situado numa posição avaliativa neutra). Por sua vez, esta inferência ganha legitimidade se pensarmos que, na maioria dos cursos de licenciatura do ensino superior em Portugal, e no caso particular da Universidade do Minho, ainda predomina uma visão de avaliação sazonal e localizada em período temporal formalmente definido e não a presença efectiva de lógica avaliativa contínua.

No que diz respeito à dimensão Trabalhos/Leituras, no quadro 4.10 indicamos os coeficientes de correlação para os seus dois itens:

(8.1) Os trabalhos e/ou leituras foram úteis

(8.2) Os trabalhos e/ou leituras complementares, aconselhados pelo docente, contribuíram para aumentar a compreensão e o interesse pela disciplina

Quadro 4.10 Correlações dos itens na dimensão Trabalhos/Leituras

Trabalhos/ Leituras		
Item		it1 x it2
Ano	1 e 2ºano	.813
	3º ano	.842
	4 e 5ºano	.856
Tipologia	T	.826
	P	.829

Nesta dimensão, os itens 1 e 2 encontram-se significativamente correlacionados. Procurando obedecer ao princípio da parcimónia (evitar um número discrepante de itens nas diferentes dimensões), julgamos que a informação veiculada pelos 2 itens se justifica e complementa. Por um lado, o reflexo da utilidade dos trabalhos/leituras e por outro, o seu contributo para o aumento da compreensão e o interesse pela disciplina.

No que diz respeito à dimensão Carga de Trabalho/Dificuldade, no quadro 4.11 indicamos os coeficientes de correlação para os seus quatro itens:

(9.1) A dificuldade desta disciplina relativamente às outras disciplinas do curso foi...

(9.2) A exigência/carga de trabalho desta disciplina relativamente às outras disciplinas do curso foi...

(9.3) O ritmo da disciplina/aulas desta disciplina relativamente às outras disciplinas do curso foi...

(9.4) Para além das aulas, qual o nº médio de horas semanais de estudo nesta disciplina ao longo do ano/semestre?

Quadro 4.11 Correlações dos itens na dimensão Carga de Trabalho/Dificuldade

Carga de Trabalho /Dificuldade							
Item		dif x carga	dif x ritmo	dif x horas	carga x ritmo	carga x horas	ritmo x horas
Ano	1 e 2ºano	.148**	-.041**	-.089**	.378**	.271**	.158**
	3º ano	-.107**	-.026*	-.102**	.352**	.315**	-.164**
	4 e 5ºano	-.097**	-.020	-.137**	.418**	.376**	.193**
Tipologia	T	-.153**	-0.58**	-.112**	.371**	.284**	.169**
	P	-.122**	-.028**	-.097**	.383**	.322**	.169**

* = coeficiente cuja probabilidade para $p < 0.05$ é significativa

** = coeficiente cuja probabilidade para $p < 0.01$ é significativa

Nesta dimensão, apenas os itens associados ao ritmo da aula e à carga de trabalho surgem com maior correlação, ou seja, estes dados revelam grande especificidade dos itens o que valida uma resposta consciente e não estereotipada dos alunos aos próprios questionários. Assim,

e a julgar pela análise, importa considerar a pertinência de manter todos itens, embora o item relativo ao ritmo das aulas apresente alguma sobreposição em termos comportamentais (percepção da adequação do ritmo das aulas) com o item 4 na dimensão Organização/Clareza⁴³.

4.4. Dimensionalidade

A dimensionalidade do PEA foi avaliada a partir do modelo de Análise Factorial Exploratória (ACF) e do método de análise das componentes principais (ACP), tendo-se analisado diversas soluções ortogonais (rotação varimax – com normalização de Kaiser) sem especificação de factores. No quadro 4.4, descrevemos os resultados da análise factorial com base nos 33 itens que constituem a versão experimental do questionário, tendo-se assumido todos os valores com valor próprio igual ou superior à unidade e os coeficientes de saturação dos itens com valores iguais ou superiores a .40. Por razões de maior clareza na apresentação, os itens encontram-se organizados não pela sequência de apresentação no questionário, mas pela sua vinculação aos factores identificados. De sublinhar que, os 2 itens reportados à avaliação global da disciplina e do docente e os 2 itens à assiduidade docente e à percepção da importância da disciplina no curso (itens acrescentados à versão original do questionário) não foram considerados nesta análise.

Os resultados obtidos na solução final parecem bastante consistentes e claros, embora estrutura factorial não tenha replicado na totalidade a esperada (9 dimensões/factores; Marsh, 2001). A estrutura obtida organiza-se de forma consistente, em 5 factores, com valor-próprio superior à unidade, explicando 70.64% da variância dos resultados nos itens.

⁴³ Especificamente a formulação do item é a seguinte: (*“O ritmo das aulas facilitou o registo de apontamentos”*) (cf. Anexo B).

Quadro 4.12 Análise factorial

Factores	PEA				
	I	II	III	IV	V
Aprendizagem 1	,791				
Aprendizagem 3	,784				
Empenhamento 4	,755				
Empenhamento 3	,751	,408			
Aprendizagem 2	,733				
Aprendizagem 4	,717				
Empenhamento 2	,710	,448			
Organização 1	,706				
Empenhamento 1	,623	,475			
Organização 3	,526				
Organização 2	,516				
Organização 4	,500	,409			
Interacção 4	,470	,726			
Interacção 2	,483	,704			
Interacção 3	,498	,700			
Interacção 1	,492	,689			
Relação 3		,680	,461		
Relação 2		,661	,480		
Relação 1		,644	,457		
Relação 4		,572	,493		
Profundidade 2			,720		
Profundidade 1			,718		
Profundidade 3			,703		
Profundidade 4			,684		
Avaliação 3				,755	
Avaliação 2				,719	
Trabalho 1				,702	
Trabalho 2				,661	
Avaliação 1				,597	
Dificuldade				,411	
Carga de Trabalho					,794
Nº Horas de Estudo					,659
Ritmo das Aulas					,622
% Variância	52.44	5.53	5.16	4.44	3.05
Valor-próprio	17.30	1.82	1.70	1.46	1.00

Como seria de esperar em face à natureza das dimensões identificadas e que passamos a assumir como categorias do PEA, alguns itens saturam em mais do que um factor. Explícite-se, no entanto, que face ao critério de saturação definido, .4, apenas 8 dos 33 itens do questionário se repartam por mais que um factor.

De seguida, apresentam-se as dimensões do PEA em termos de uma análise mais qualitativa dos itens e da sua justificação de posicionamento em determinada dimensão⁴⁴. Itens como o Empenhamento Docente (1, 2 e 3); Organização/Clareza (item 4) e Interação da Turma (1, 2, 3 e 4) pontuam em mais do que um factor, contudo foi opção neste trabalho, integrá-los apenas num factor, tendo como razão subjacente a análise qualitativa dos itens em causa.

Mais especificamente o **factor I** traduz um vector mais geral de apreciação da qualidade da aprendizagem (relevância e utilidade) isto é, reflecte um ponto de convergência de uma pluralidade de percepções que decorre da confluência de atributos relacionados com o docente, nomeadamente o grau do seu empenhamento (traduzido no dinamismo e interesse suscitados pelas aulas) e a organização/clareza das matérias abordadas (materiais de apoio ao estudo e a facilidade no registo de apontamentos) num conjunto de 12 itens. Embora o item 4 da organização/clareza (ritmo adequado ao registo de apontamentos) sature no 2º factor, foi nossa opção, mantê-lo associado ao 1º factor pela partilha de indicadores comportamentais associados tendencialmente à organização e gestão da sala de aula. Por sua vez, este factor é muito próximo do conceito de conscienciosidade pedagógica, defendido nos estudos sobre a cidadania docente no ensino superior (Rego, 2003).

O **factor II** reúne 5 itens que, no seu conjunto, traduzem aspectos sócio-relacionais que reflectem o grau de satisfação do aluno com as relações interpessoais estabelecidas com o docente (esclarecimento de dúvidas, pedido de ajuda/conselhos dentro e fora da sala de aula) e com o clima relacional suscitado pela interação da turma (criação de uma atmosfera favorável à estimulação da participação dos alunos, através da partilha de ideias e conhecimentos e de questionamento ao docente). De salientar que, os itens relativos à interação da turma também saturam no 1º factor, contudo, entendemos que a sua análise qualitativa remete para uma dimensão mais relacional do processo ensino/aprendizagem mais congruente com o factor em análise. Por último, os 3 itens relativos à dimensão do empenhamento docente (dinamismo e o carácter atractivo da abordagem das matérias) parecem também concorrer para o estabelecimento de uma matriz relacional conducente à satisfação académica. Por sua vez, estes itens são compagináveis com o conceito de cidadania docente (Rego, 2003) e com os princípios pedagógicos de qualidade relativos à necessária “*democraticidade*” no ensino/aprendizagem (cf. Vieira *et al.*, 2002).

⁴⁴ De referir que, no presente trabalho, procedemos à atribuição de designações às dimensões identificadas, que tem como objectivo, por um lado, a simplificação a informação obtida, e por outro, nelas ver reflectidas o constructo avaliado (percepções do ensino/aprendizagem).

O **factor III** traduz uma dimensão que agrupa itens associados ao grau de transversalidade e abrangência na abordagem dos assuntos (e.g. contexto do aparecimento das ideias/conceitos desenvolvidos nas aulas; comparação das implicações das diversas teorias/modelos variedade e flexibilidade de perspectivas utilizadas pelo docente e a percepção dos desenvolvimentos recentes nas matérias leccionadas). Mais uma vez, este factor é consistente com os princípios pedagógicos de qualidade relativos à necessária “*reflexividade*” no ensino/aprendizagem (cf. Vieira *et al.*, 2002).

O **factor IV** integra 6 itens que se reportam claramente à avaliação e aos trabalhos/leituras (percepções de adequação e justiça dos elementos de avaliação) bem como a utilidade dos trabalhos/leituras aconselhados e os seus contributos ao nível da compreensão e do interesse pela disciplina. Por sua vez, o grau de dificuldade tende a associar-se com esta dimensão.

O **factor V** traduz o confronto pessoal do aluno com as características da disciplina (carga de trabalho, horas de estudo, ritmo das aulas e a importância da disciplina no curso) num conjunto de 3 itens.

De referir que esta análise factorial foi repetida pelos 7 Conselhos de Cursos, replicando-se na generalidade dos casos a estrutura dos cinco factores e o mesmo significado para a amostra global. No entanto, tomando os alunos do Conselho de Cursos de Ciências Económicas, Empresariais e Políticas, a estrutura factorial obtida é constituída por seis factores explicando no seu conjunto cerca de 71.5% da variância dos resultados nos 33 itens (o primeiro factor explica 48.2%). No quadro 4.13 apresentamos os valores da análise.

Quadro 4.13 Análise factorial do PEA no Conselho de Cursos de Ciências Económicas, Empresariais e Políticas

Factores	PEA					
	I	II	III	IV	V	VI
Empenhamento 4	,807					
Aprendizagem 1	,806					
Empenhamento 3	,805					
Aprendizagem 3	,794					
Organização 1	,766					
Empenhamento 2	,757					
Aprendizagem 2	,739					
Aprendizagem 4	,732					
Empenhamento 1	,711					
Organização 2	,555					
Organização 3	,550					
Organização 4	,521					
Profundidade 2		,756				
Profundidade 1		,753				
Profundidade 3		,746				
Profundidade 4		,717				
Avaliação 3			,800			
Avaliação 2			,757			
Trabalho 1			,662			
Trabalho 2		,411	,624			
Avaliação 1			,623			
Relação 3				,783		
Relação 2				,771		
Relação 1				,765		
Relação 4				,689		
Interacção 2	,404				,789	
Interacção 1	,410				,772	
Interacção 4					,772	
Interacção 3	,428				,746	
Carga de Trabalho						,772
Ritmo das Aulas						,738
Nº Horas de Estudo						,503
Dificuldade						-,407
% Variância	48.20	6.76	5.55	4.49	3.45	3.11
Valor-próprio	15.90	2.23	1.83	1.48	1.13	1.02

Conforme podemos constatar a estrutura replica a verificada na amostra global e nos demais Conselhos de Cursos, contudo, verifica-se agora uma melhor clarificação ou diferenciação entre os itens da interacção da turma no contexto de sala de aula (Factor V) e os itens da relação do docente com os alunos (Factor IV).

O quadro 4.14 apresenta os valores da média e desvio-padrão da amostra para os factores/dimensões em função do ano em que se inscreve a disciplina no curso, tendo-se considerado o 1º ano de forma isolada, o 2º ano e 3ºano de forma agrupada tal como o 4º e 5º ano).

Quadro 4.14 Médias e desvio-padrão dos resultados nas dimensões do PEA em função dos Conselhos de Cursos e do ano da disciplina

Conselho de Cursos	Ano da disciplina	Relevância da Aprendizagem		Empenhamento Docente		Organização Clareza		Interacção da Turma		Relação Docente Aluno		Profundidade na abordagem dos assuntos		Avaliação Classificações		Trabalhos Leituras		Carga Trabalho Dificuldade	
		M	DP	M	DP	M	DP	M	DP	M	DP	M	DP	M	DP	M	DP	M	DP
<i>Ciências</i>	1 e 2ºano	6.41	1.49	6.60	1.66	6.77	1.36	6.68	1.55	7.11	1.39	6.56	1.38	6.49	1.52	6.40	1.47	4.58	.776
	3ºano	6.45	1.58	6.50	1.78	6.58	1.47	6.61	1.62	6.98	1.56	6.57	1.45	6.42	1.58	6.36	1.51	4.94	.809
	4º e 5ºano	6.62	1.61	6.65	1.77	6.76	1.55	6.95	1.58	7.19	1.51	6.85	1.48	6.50	1.73	6.52	1.63	5.02	.850
<i>Ciências Económicas e Políticas</i>	1 e 2ºano	6.55	1.57	6.64	1.81	6.76	1.45	6.76	1.60	7.17	1.41	6.83	1.37	6.49	1.56	6.62	1.49	4.92	.775
	3ºano	6.57	1.42	6.58	1.72	6.74	1.37	6.82	1.57	7.03	1.42	7.03	1.21	6.49	1.52	6.56	1.46	4.99	.727
	4º e 5ºano	6.69	1.50	6.71	1.73	6.83	1.45	6.89	1.52	7.18	1.45	7.03	1.33	6.78	1.43	6.75	1.40	5.04	.791
<i>Ciências Sociais</i>	1 e 2ºano	6.67	1.52	6.73	1.78	6.86	1.50	6.83	1.63	7.22	1.44	6.96	1.42	6.76	1.48	6.86	1.45	4.97	.775
	3ºano	6.55	1.57	6.60	1.84	6.95	1.58	6.95	1.58	7.21	1.50	6.91	1.37	6.74	1.56	6.80	1.61	4.92	.769
	4º e 5ºano	6.98	1.44	6.94	1.74	7.17	1.57	7.17	1.57	7.46	1.45	7.23	1.44	6.92	1.59	6.94	1.53	5.19	.835
<i>Educação Infantil e Básica</i>	1 e 2ºano	7.09	1.50	7.11	1.64	7.02	1.51	7.16	1.52	7.28	1.49	7.08	1.43	6.92	1.55	6.94	1.56	5.15	.914
	3ºano	6.45	1.69	6.10	1.88	6.38	1.57	6.61	1.63	6.57	1.66	6.45	1.48	6.18	1.64	6.43	1.56	4.96	.814
	4º e 5ºano	7.10	1.31	6.89	1.47	7.00	1.33	7.14	1.63	7.19	1.56	6.82	1.38	6.86	1.41	7.02	1.31	5.35	.897
<i>Engenharia</i>	1 e 2ºano	6.35	1.53	6.54	1.74	6.59	1.46	6.51	1.60	6.88	1.59	6.45	1.47	6.30	1.59	6.17	1.59	4.87	.842
	3ºano	6.61	1.45	6.80	1.67	6.80	1.40	6.84	1.46	7.22	1.48	6.74	1.41	6.29	1.60	6.31	1.57	4.95	.875
	4º e 5ºano	6.61	1.52	6.67	1.72	6.70	1.47	6.80	1.54	7.12	1.43	6.74	1.46	6.57	1.56	6.55	1.51	5.09	.946
<i>Educação e Psicologia</i>	1 e 2ºano	6.94	1.47	6.96	1.75	7.12	1.35	7.23	1.47	7.45	1.37	7.29	1.25	7.13	1.41	7.04	1.42	5.08	.786
	3ºano	6.95	1.46	6.84	1.82	7.10	1.40	7.04	1.74	7.38	1.56	7.35	1.26	6.66	1.64	6.99	1.52	5.02	.743
	4º e 5ºano	7.01	1.54	6.84	1.90	7.01	1.58	7.18	1.59	7.24	1.69	7.12	1.58	6.79	1.84	6.94	1.63	5.19	.850
<i>Letras e Ciências Humanas</i>	1 e 2ºano	6.75	1.66	6.82	1.81	7.00	1.49	7.14	1.60	7.31	1.45	7.31	1.45	6.78	1.65	6.93	1.57	4.94	.818
	3ºano	6.84	1.58	6.85	1.81	6.98	1.57	7.15	1.72	7.34	1.64	7.34	1.64	7.01	1.53	7.00	1.58	5.08	.765
	4º e 5ºano	6.60	1.54	6.63	1.85	7.13	1.37	6.89	1.71	7.31	1.63	7.31	1.63	6.65	1.65	7.00	1.46	4.91	.746
<i>Total</i>	1 e 2ºano	6.56	1.55	6.69	1.75	6.78	1.46	6.77	1.60	7.11	1.50	6.75	1.44	6.56	1.58	6.54	1.56	4.92	.822
	3ºano	6.61	1.52	6.65	1.77	6.77	1.46	6.84	1.59	7.13	1.54	6.84	1.40	6.50	1.61	6.54	1.57	4.97	.808
	4º e 5ºano	6.72	1.53	6.72	1.75	6.83	1.48	6.93	1.57	7.19	1.50	6.93	1.45	6.66	1.61	6.70	1.53	5.09	.880

O quadro 4.15 apresenta os valores da média e desvio-padrão da amostra para os factores/dimensões em função da natureza Teórica ou Prática das aulas⁴⁵.

Quadro 4.15 Médias e desvio-padrão dos resultados nas dimensões do PEA em função dos Conselhos de Cursos e da tipologia de aula

Conselho de Cursos	Tipologia	Relevância da Aprendizagem		Empenhamento Docente		Organização Clareza		Interação da Turma		Relação Docente Aluno		Profundidade na abordagem dos assuntos		Avaliação Classificações		Trabalhos Leituras		Carga Trabalho Dificuldade	
		M	DP	M	DP	M	DP	M	DP	M	DP	M	DP	M	DP	M	DP	M	DP
<i>Ciências</i>	T	6.33	1.55	6.39	1.73	6.56	1.46	6.57	1.59	6.98	1.44	6.68	1.42	6.48	1.57	6.46	1.47	4.90	.752
	P	6.55	1.52	6.72	1.69	6.83	1.39	6.83	1.56	7.18	1.46	6.58	1.43	6.48	1.59	6.39	1.55	4.85	.842
<i>Ciências Económicas e Políticas</i>	T	6.51	1.59	6.35	1.92	6.68	1.46	6.35	1.76	6.78	1.56	7.04	1.33	6.13	1.63	6.59	1.66	5.01	.803
	P	6.61	1.51	6.71	1.73	6.79	1.42	6.89	1.52	7.22	1.38	6.89	1.33	6.64	1.49	6.64	1.43	4.95	.759
<i>Ciências Sociais</i>	T	6.60	1.55	6.56	1.87	6.84	1.45	6.69	1.72	7.07	1.53	7.01	1.45	6.52	1.50	6.79	1.48	4.94	.798
	P	6.73	1.51	6.79	1.76	6.89	1.51	7.00	1.56	7.32	1.43	7.01	1.39	6.85	1.52	6.89	1.51	5.02	.788
<i>Educação Infantil e Básica</i>	T	7.03	1.52	6.97	1.71	7.01	1.54	6.84	1.65	7.14	1.42	7.29	1.25	6.72	1.65	6.95	1.54	5.12	.974
	P	6.90	1.56	6.79	1.74	6.82	1.53	7.03	1.57	7.07	1.59	6.83	1.47	6.70	1.58	6.79	1.54	5.11	.878
<i>Engenharia</i>	T	6.36	1.59	6.57	1.74	6.64	1.50	6.63	1.57	6.97	1.56	6.68	1.44	6.43	1.60	6.37	1.55	4.88	.828
	P	6.51	1.48	6.65	1.73	6.68	1.43	6.66	1.56	7.02	1.53	6.54	1.46	6.34	1.58	6.26	1.58	4.97	.898
<i>Educação e Psicologia</i>	T	6.88	1.52	6.73	1.88	7.03	1.44	7.01	1.61	7.32	1.51	7.39	1.24	6.92	1.56	7.01	1.49	5.03	.771
	P	7.02	1.46	7.03	1.74	7.13	1.41	7.30	1.54	7.42	1.51	7.16	1.42	7.07	1.62	7.00	1.52	5.14	.811
<i>Letras e Ciências Humanas</i>	T	6.60	1.71	6.68	1.86	6.97	1.48	6.69	1.85	7.34	1.40	7.18	1.47	6.58	1.63	6.86	1.54	4.73	.821
	P	6.76	1.60	6.81	1.81	7.03	1.49	7.16	1.60	7.32	1.54	7.09	1.39	6.85	1.62	6.98	1.55	5.01	.779
<i>Total</i>	T	6.50	1.59	6.55	1.79	6.73	1.48	6.67	1.64	7.04	1.52	6.90	1.42	6.52	1.60	6.58	1.55	4.93	.806
	P	6.65	1.52	6.74	1.74	6.81	1.45	6.88	1.57	7.16	1.50	6.77	1.44	6.59	1.59	6.57	1.56	4.99	.843

⁴⁵ As aulas de natureza mais prática incluem as aulas Teórico-Práticas, Práticas e Seminários. Dos 41 962 questionários, 1089 (cerca de 2.6%) não apresentavam referência à tipologia de aula, e por isso, as análises que incidem sobre a variável tipologia consideram apenas 40 873 questionários válidos.

Globalmente, os resultados relativos à dimensão Relevância da Aprendizagem indicam uma avaliação positiva em todos os itens considerados, isto é, valores médios próximos de 7 numa escala de 9 pontos.

Quando se analisam os resultados em função do ano do curso dos respondentes destaca-se que a estimulação intelectual promovida pelas aulas, o aumento do interesse pela frequência das aulas, a avaliação da utilidade da aprendizagem e a percepção de compreensão e assimilação dos conteúdos não é muito diferente em função dos anos nos vários Conselhos de Cursos. No entanto, esta percepção é mais elevada junto dos últimos anos do curso, com excepção do Conselho de Cursos de Letras e Ciências Humanas, em que valores superiores se encontram no 3º ano (6.84), seguido do 1 e 2º ano (6.75). Os resultados em função da natureza das aulas revelam que as aulas práticas tendem a ser ligeiramente melhor avaliadas por comparação às aulas teóricas para todos os itens considerados, à excepção do Conselho de Cursos de Educação Infantil e Básica.

No que diz respeito à avaliação do Empenhamento Docente, os resultados desta dimensão mostram uma avaliação positiva dos itens considerados, com valores médios próximos de 7. Uma análise em função do ano do curso permite-nos concluir que, os alunos de anos mais avançados tendem a avaliar de forma mais positiva a totalidade dos itens desta dimensão, situação que não se verifica nos Conselho de Cursos de Engenharia e de Letras e Ciências Humanas (anos intermédios) e no Conselho de Cursos de Educação Infantil e Básica (primeiros anos). À excepção do Conselhos de Cursos de Educação Infantil e Básica, verificamos que de uma forma genérica as aulas práticas são contextos que tendem a promover uma avaliação mais positiva em relação a aspectos como, o dinamismo das aulas, a forma atractiva de exposição da matéria e a forma como o docente cativa o interesse dos alunos.

Mais uma vez, os resultados da dimensão Organização/Clareza revelam uma avaliação positiva dos itens considerados em função dos anos, atendendo às médias obtidas (valores próximos de 7). Globalmente, as médias das respostas mostram que as aulas práticas são ligeiramente melhor avaliadas do que as aulas teóricas (mais uma vez com excepção do Conselhos de Cursos de Educação Infantil e Básica), seguindo o mesmo padrão de resultados das dimensões anteriores.

Os resultados obtidos na Interação da Turma mostram uma avaliação bastante positiva dos itens considerados, atendendo às médias obtidas (próximos do nível 7 de uma escala de 9). A análise em função dos agrupamentos dos anos, revela que também em relação a todos estes itens, o ano é um aspecto diferenciador da avaliação nestes itens, ressaltando que a avaliação é mais positiva junto dos últimos anos. Quanto à distinção entre as aulas práticas e as aulas teóricas, mais uma vez se

verifica a superioridade das primeiras (à excepção do Conselho de Cursos de Educação Infantil e Básica).

Os resultados gerais obtidos na dimensão Relação Docente/Aluno, mostram uma avaliação bastante positiva dos itens considerados, atendendo às médias obtidas (valores superiores a 7) comparativamente às restantes dimensões do PEA. A análise destes resultados mostra que esta dimensão relacional é diferentemente avaliada em função dos anos, conhecendo um aumento nos últimos anos. Quanto à distinção entre as aulas práticas e aulas teóricas, mais uma vez se verifica a superioridade na avaliação das primeiras.

Os resultados obtidos em relação à dimensão Profundidade na Abordagem dos Assuntos situam-se, novamente num nível de avaliação bastante positivo, atendendo às médias obtidas que se apresentam como superiores a 6 e próximas de 7, em muitos dos casos. Quanto à distinção entre as aulas teóricas e aulas práticas na avaliação efectuada, emerge um dado que contraria a tendência verificada nas restantes dimensões do PEA, atribuindo valores superiores às aulas teóricas nos últimos anos do curso. Situação excepcional verifica-se no Conselho de Cursos de Educação Infantil e Básica (1 e 2º ano) e no Conselho de Cursos de Educação e Psicologia (3º ano).

Os resultados obtidos em relação à Avaliação/Classificações, e atendendo às médias dos resultados pelos vários itens (valores próximos 7), mostram que a avaliação desta dimensão é mais positivamente apreciada nos últimos anos do curso e nas aulas práticas, excepto mais uma vez no Conselho de Cursos de Educação Infantil e Básica e no Conselho de Cursos de Engenharia.

A análise das respostas aos itens da dimensão Trabalhos/Leituras mostra que, de uma maneira geral, os alunos tendem a manifestar percepções mais positivas nos últimos anos do curso e sem diferenciação entre as aulas teóricas e práticas, apresentando ambas uma avaliação positiva.

A análise das respostas aos itens da dimensão Carga de Trabalho/Dificuldade indica que, de uma maneira geral, os alunos tendem a avaliar as percepções de maior exigência/carga de trabalho, dificuldade, ritmo, horas de estudo, de forma positiva nos últimos anos, embora sem diferenças significativas entre a avaliação nas aulas práticas e teóricas. De salientar ainda que, em relação às primeiras a avaliação é um ligeiramente menos positiva junto dos Conselhos de Cursos de Ciências, Ciências Económicas e Educação Infantil e Básica.

Em síntese, podemos realçar alguns dos aspectos mais consistentes e relevantes deste conjunto de resultados relativos ao PEA. Assim, é de destacar a avaliação sempre positiva de todos os itens. A título ilustrativo, a dimensão que obteve melhor classificação no conjunto de todas as dimensões/facetos, cruzando o ano da disciplina e a tipologia é a que se refere à Relação Docente/Aluno, já a dimensão que obteve pontuações menos positivas (ainda que apresentando um

valor médio positivo no conjunto da escala de resposta) foi a que se refere às dimensões Avaliação/Classificações e Carga de Trabalho/Dificuldade. De realçar também, e em termos gerais, a avaliação é claramente mais positiva junto dos alunos dos últimos anos e das aulas práticas.

No sentido de proceder a uma avaliação mais fina dos resultados obtidos, avançamos na segunda fase para a análise de variância dos mesmos em tais dimensões considerando duas variáveis independentes (ano da disciplina avaliada e tipologia de aula). Em relação ao ano da disciplina consideramos 3 níveis: 1º e 2º anos, 3º ano, e 4º e 5º anos. Em relação à tipologia de aulas consideramos dois níveis: teóricas e práticas. Esta análise de variância tomou os efeitos principais e secundários recorrendo-se ao procedimento F-Manova (3x2).

Quadro 4.16 Análise de variância (MANOVA) dos resultados nas dimensões do PEA em função da tipologia de aula e ano da disciplina

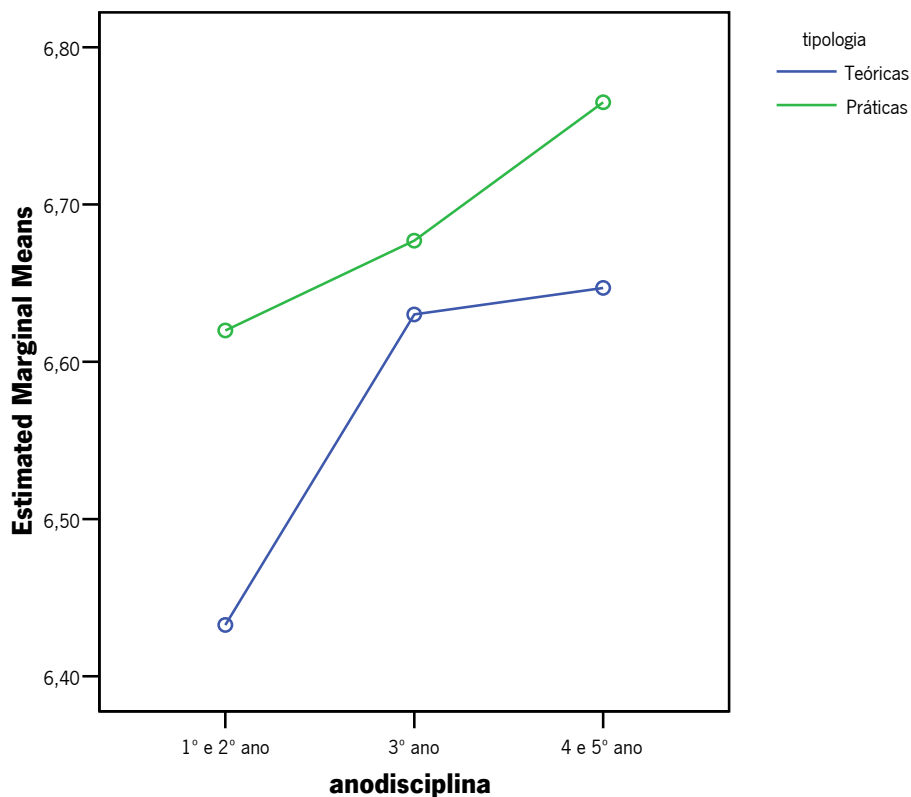
Dimensões	Fontes de Variação	Soma dos quadrados	G.L	F	p
<i>Relevância da Aprendizagem</i>	Ano	147.000	2	31.454	.000
	Tipologia de aula	67.594	1	28.926	.000
	Ano x Tipologia de aula	18.869	2	4.038	.018
<i>Empenhamento Docente</i>	Ano	41.998	2	6.794	.001
	Tipologia de aula	91.000	1	29.443	.000
	Ano x Tipologia de aula	52.327	2	8.465	.000
<i>Organização/Clareza</i>	Ano	28.009	2	6.549	.001
	Tipologia de aula	15.329	1	7.168	.007
	Ano x Tipologia de aula	6.723	2	1.572	.208
<i>Interacção da Turma</i>	Ano	199.849	2	39.521	.000
	Tipologia de aula	104.913	1	41.494	.000
	Ano x Tipologia de aula	85.354	2	16.879	.000
<i>Relação Docente/Aluno</i>	Ano	87.998	2	19.398	.000
	Tipologia de aula	28.603	1	12.611	.000
	Ano x Tipologia de aula	47.731	2	10.522	.000
<i>Profundidade na Abordagem dos Assuntos</i>	Ano	185.318	2	45.195	.000
	Tipologia de aula	85.290	1	41.600	.000
	Ano x Tipologia de aula	5.736	2	1.399	.247
<i>Avaliação/Classificações</i>	Ano	63.053	2	45.195	.000
	Tipologia de aula	6.452	1	2.552	.110
	Ano x Tipologia de aula	41.456	2	8.198	.000
<i>Trabalhos/Leituras</i>	Ano	101.689	2	12.470	.000
	Tipologia de aula	3.208	1	1.325	.250
	Ano x Tipologia de aula	8.022	2	1.657	.191
<i>Carga Trabalho/Dificuldade</i>	Ano	97.739	2	71.519	.000
	Tipologia de aula	32.583	1	47.684	.000
	Ano x Tipologia de aula	16.558	2	12.116	.000

Atendendo à tipologia da aula (Teórica ou Prática) e ao ano da disciplina avaliada (1º e 2º ano, 3º ano e 4º e 5º ano), enquanto efeitos principais, os resultados apresentam-se diferenciados de forma estatisticamente significativa em 3 dimensões do PEA. Os alunos avaliam melhor os seus docentes na dimensão Organização/Clareza quando reportados a aulas mais avançadas no currículo do curso e quando reportados às aulas práticas. Melhores avaliações ocorrem na dimensão Profundidade na

Abordagem dos Assuntos, sendo a diferença aqui de novo favorável aos alunos de anos curriculares mais avançados, mas agora às aulas teóricas (o que se pode entender em face da maior tradição de aprofundamento dos assuntos nas aulas teóricas). Finalmente, foram ainda encontrados efeitos principais apenas do ano na dimensão Trabalhos/Leituras, cuja diferença acompanha a tendência favorável aos alunos dos últimos anos dos cursos. Esta situação apenas não ocorre junto dos alunos do Conselho de Cursos de Educação e Psicologia (este dado é congruente com algumas das conclusões encontradas na Avaliação Interna da licenciatura em Psicologia - cf. Relatório de Avaliação Interna do Curso de Psicologia no ano 2003/04 – podendo isso ficar-se a dever às novas metodologias de ensino-aprendizagem, e em particular à maior carga de trabalho dos alunos, em implementação progressiva na Licenciatura em Psicologia à medida que se avança com o seu novo plano de estudos).

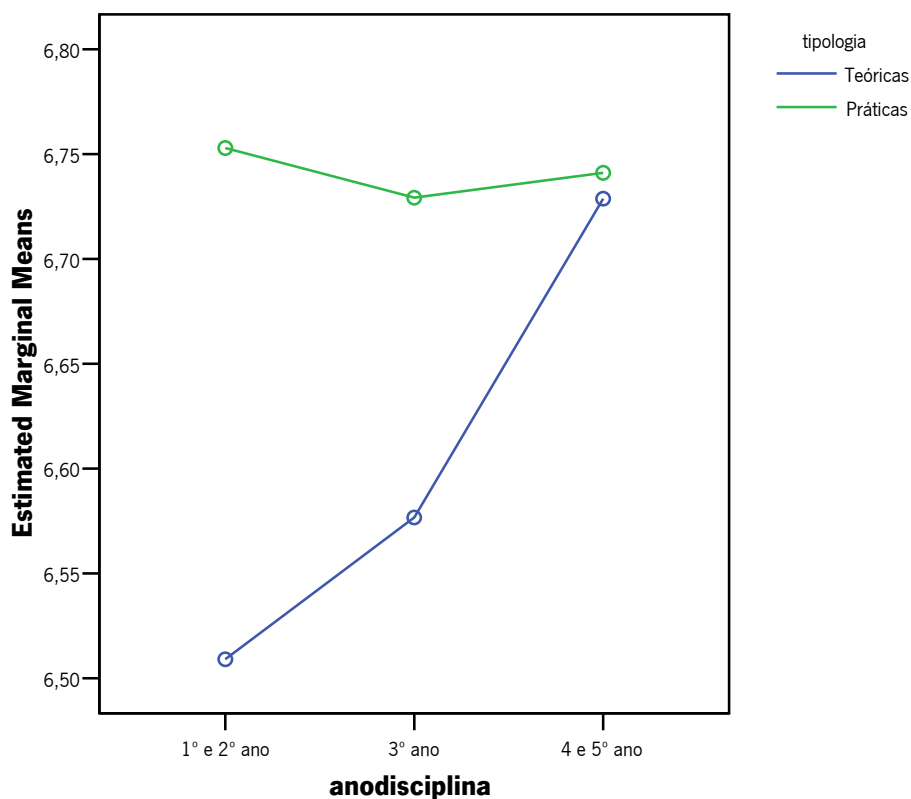
Complementarmente foram encontrados efeitos secundários de interação tomando o ano da disciplina e a tipologia das aulas nas demais seis dimensões do PEA: Relevância da Aprendizagem; Empenhamento Docente; Interação da Turma; Relação Docente/Aluno; Avaliação/Classificações e Carga Trabalho/Dificuldade. Tendo em vista uma melhor análise e compreensão destes efeitos, apresentamos graficamente tais interações. Assim, no Gráfico 4.1 indicamos o sentido da interação das duas variáveis na variância da dimensão Relevância da Aprendizagem.

Gráfico 4.1- Efeitos de Interação na dimensão Relevância da Aprendizagem



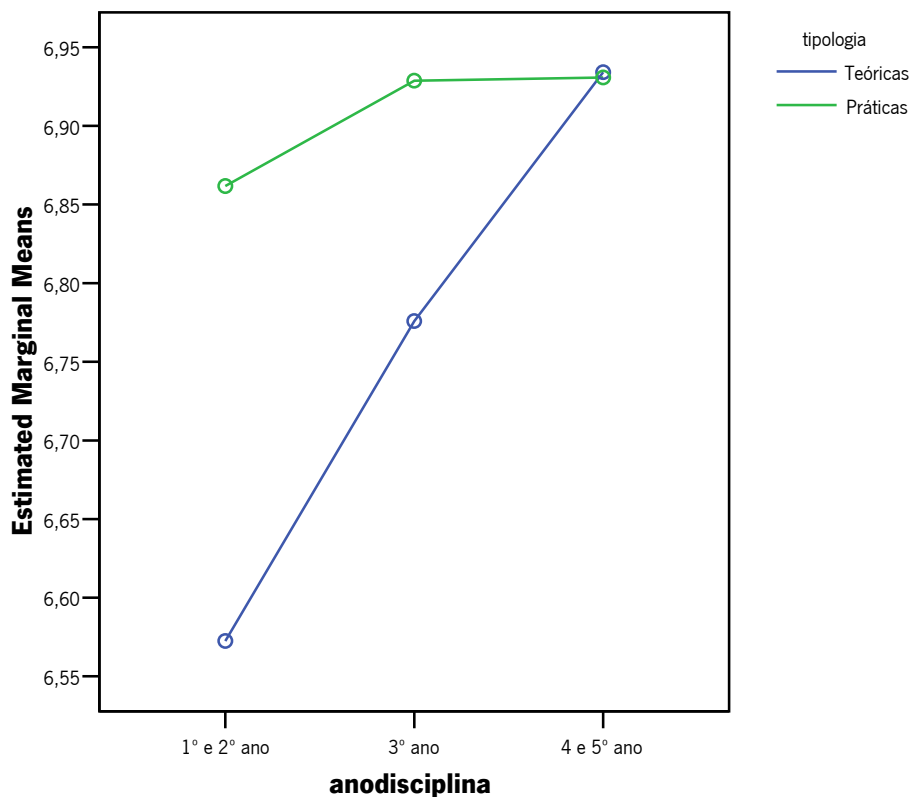
Tal como é possível visualizar, as percepções de Relevância da Aprendizagem tendem a ser mais favoráveis nas aulas práticas e nas disciplinas de anos intermédios e finais do curso. Contudo, a diferenciação entre aulas teóricas e práticas não é comum ao longo dos anos. Tais diferenças são maiores nos anos iniciais e finais das licenciaturas. Este dado vai ao encontro dos resultados obtidos nas investigações, indicando que as avaliações do ensino tendem a ser mais positivas em anos intermédios e finais do curso, provavelmente pelo facto dos alunos estarem mais motivados nos seus estudos e discriminarem, com maior conhecimento de causa as condições de ensino/aprendizagem (Langbein, 1994; Aleamoni & Hexner, 1980; Goldberg & Callahan, 1991).

Gráfico 4.2- Efeitos de Interação na dimensão Empenhamento Docente



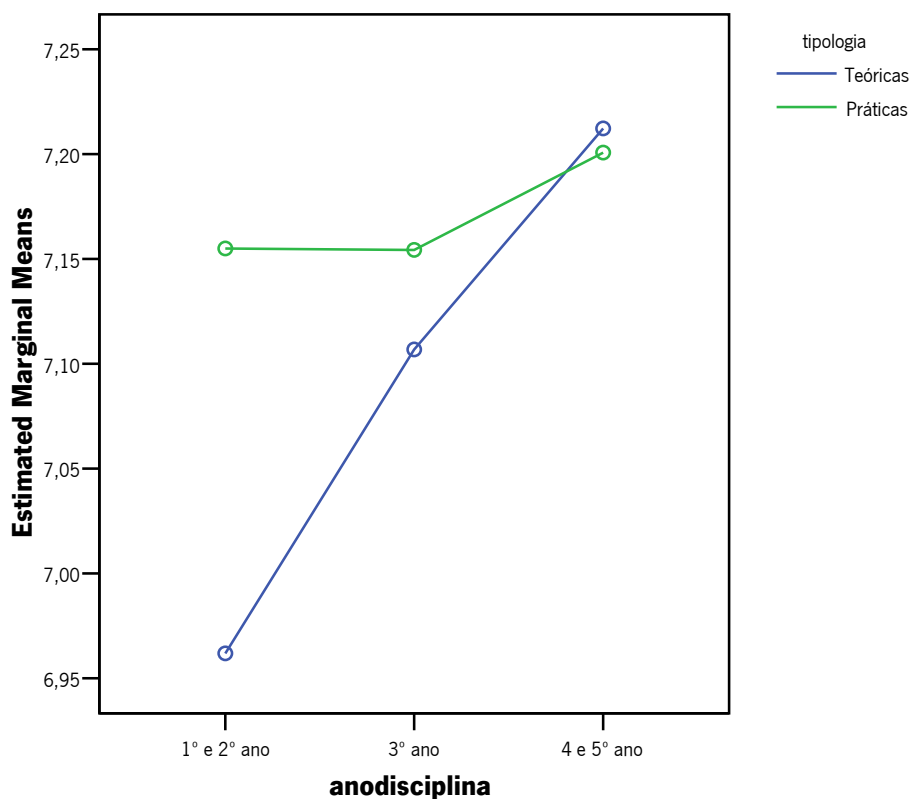
Tal como é possível visualizar no Gráfico 4.2, as percepções de Empenhamento Docente tendem a ser mais positivamente percebidas por parte das aulas práticas, mas isso apenas ocorre nos anos iniciais e no 3º ano das licenciaturas. A diferença vai-se atenuando à medida que se avança dos anos iniciais para os anos finais das licenciaturas. Este dado permite-nos concluir que as aulas práticas constituem contextos em que os estudantes percebem maior empenhamento/compromisso profissional por parte dos docentes, consubstanciado na maior probabilidade de manifestação de comportamentos comunicacionais (e.g. linguagem, tom de voz, meios audiovisuais utilizados) com impacto significativo na abordagem atractiva das matérias.

Gráfico 4.3- Efeitos de Interação na dimensão Interação da Turma



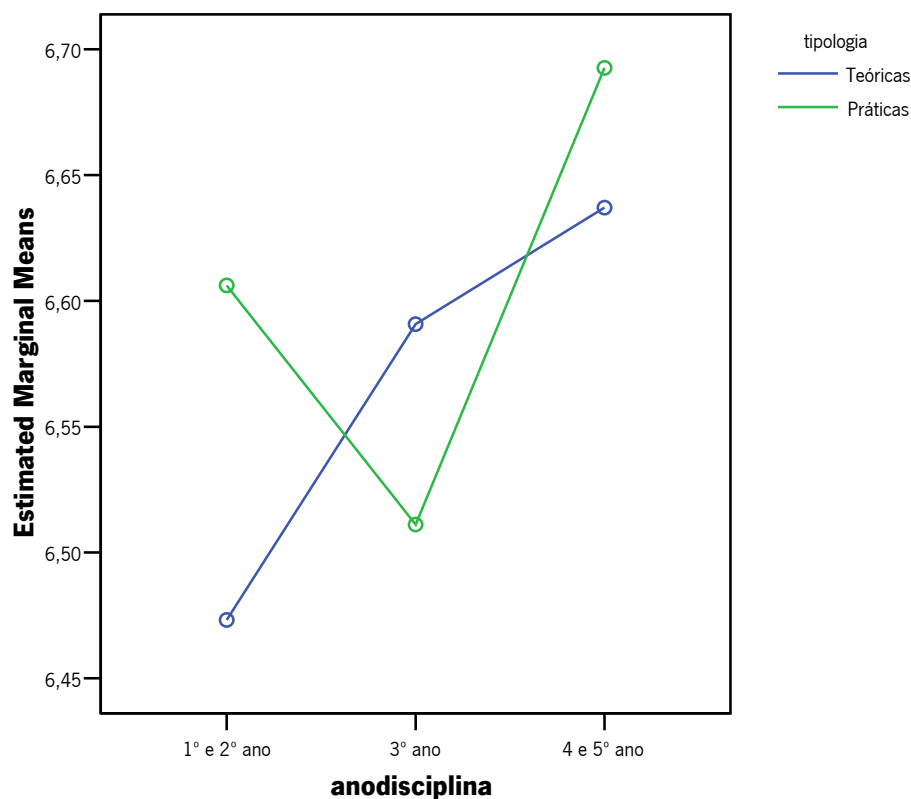
A análise do Gráfico 4.3 permite concluir que, as percepções da Interação da Turma tendem a evoluir positivamente nas aulas teóricas ao longo do curso enquanto que nas aulas práticas essas percepções revelam um padrão mais ou menos consistente nos vários anos do curso. Mais uma vez, este resultado permite-nos concluir, que as aulas práticas constituem contextos percebidos, pelos alunos dos primeiros anos do curso, como promotores de uma maior interação entre alunos e entre docente/alunos comparativamente às aulas teóricas (Cashin, 1992; Feldman, 1984; Marsh & Dunkin, 1992) nomeadamente, pela possibilidade que é dada aos alunos para colocar questões; permitir o desafio de ideias ou pontos de vista diferentes do docente e o encorajamento de discussões/debate de ideias (Rodabaugh, 1996, p.4 in Fleming, 2003).

Gráfico 4.4- Efeitos de Interação na dimensão Relação Docente/Aluno



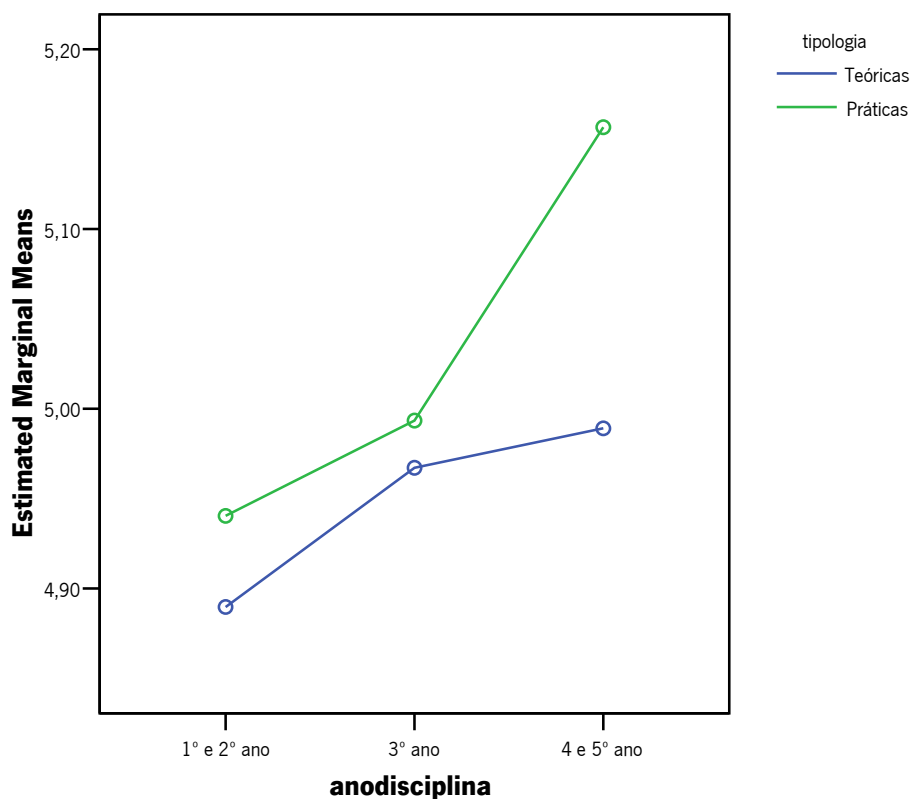
De acordo com a informação veiculada no Gráfico 4.4, as percepções do relacionamento Docente/Aluno tendem a evoluir positivamente ao longo do curso nas aulas teóricas enquanto que nas aulas práticas essas percepções revelam valores mais ou menos consistentes nos vários anos, embora se detectem valores francamente superiores nos últimos anos. Mais uma vez, este dado indica que as aulas práticas constituem contextos que promovem relações interpessoais de maior envolvimento e proximidade entre docentes e alunos comparativamente às aulas teóricas, apenas nos anos iniciais e intermédios (aliás o gráfico apresenta uma inversão da apreciação positiva tomando a natureza das aulas nos anos finais).

Gráfico 4.5- Efeitos de Interação na dimensão Avaliação/Classificações



Tal como é possível visualizar com base na leitura do Gráfico 4.5, as percepções dos alunos oscilam bastante na dimensão Avaliação/Classificações (adequação e justiça) quando consideramos simultaneamente a natureza das aulas e o ano curricular a que pertence a disciplina em avaliação. Nos primeiros e últimos anos as aulas teóricas são avaliadas de forma menos positiva, contudo tomando as disciplinas do 3º ano a situação inverte-se (as aulas práticas são agora menos positivamente avaliadas a propósito da adequação e justiça da “avaliação/classificação”). Este aspecto poderá merecer uma maior preocupação dos docentes e órgãos de gestão pedagógica, olhando às especificidades de cada curso ou Conselho de Cursos. Por sua vez, tratando-se de uma dimensão do questionário em que os estudantes globalmente avaliam menos positivamente os seus docentes, importaria estudar alternativas de maior envolvimento e compromisso dos alunos na definição e negociação das formas de avaliação bem como na ponderação do *feedback* dos trabalhos sujeitos à avaliação (e.g. exame escrito, apresentações nas aulas, exames orais, seminários e relatórios).

Gráfico 4.6- Efeitos de Interacção na dimensão Carga de Trabalho/Dificuldade



Tal como é possível visualizar no Gráfico 4.6, as percepções associadas à Carga de Trabalho/Dificuldade, evoluem positivamente ao longo do curso enquanto nas aulas práticas essas percepções, traduzindo o confronto pessoal do aluno com as características da disciplina (e.g. exigência/carga de trabalho, dificuldade, ritmo das aulas e horas de estudo) tendem a assumir valores superiores a partir dos anos intermédios e finais do curso (provavelmente decorrentes de uma maior concentração de trabalhos práticos por disciplina; maior aproximação ao ano de estágio, etc.). Este dado permite-nos concluir que esta dimensão deverá ser ponderada, por um lado, procurando atender à satisfação académica deste grupo de alunos, e por outro, à necessária reestruturação dos currículos no âmbito das licenciaturas, procurando em última instância, promover a distribuição mais equilibrada da carga de trabalho pelos diferentes anos.

Considerando agora a área do curso a que pertencem as disciplinas avaliadas, procedemos à análise de variância dos resultados nas várias dimensões tomando, por facilidades de análise, quatro grupos de cursos com o respectivo número de questionários: Ciências e Engenharia (grupo 1= 21397); Ciências Económicas (grupo 2=6450); Ciências Sociais e Humanas (grupo 3=6993); Ciências da Educação (grupo 4=7122). Recorreu-se para efeitos desta análise de variância ao procedimento F

Oneway. No quadro 4.17 apresentamos os resultados nas 9 dimensões do PEA tomando os questionários destas quatro áreas de curso.

Quadro 4.17 Análise de variância One Way (ANOVA) dos resultados nas dimensões do PEA em função da área do curso

Dimensão	Grupos	Soma dos Quadrados	GL	F	P	Contrastes
<i>Relevância da Aprendizagem</i>	Entre Grupos	1241.502	3	175.717	.000	G1 <G2, G3 e G4 G2> G1; G2 <G3 e G4
	Intra Grupos	95092.751	40377			G3 <G4; G3> G1 e G2 G4 <G1, G2 e G3
<i>Empenhamento Docente</i>	Entre Grupos	353.760	3	38.104	.000	G1 <G2, G3 e G4 G2> G1; G2 <G3 e G4
	Intra Grupos	124554.84	40248			G3> G1 e G2; G3 <G4 G4> G1, G2 e G3
<i>Organização/Clareza</i>	Entre Grupos	632.618	3	98.329	.000	G1 <G2, G3 e G4 G2> G1; G2> G3 e G4
	Intra Grupos	85106.941	39685			G3> G1; G2 <G4 G4> G1, G2 e G3
<i>Interação da Turma</i>	Entre Grupos	1167.620	3	154.672	.000	G1 <G2, G3 e G4 G2> G1; G2 <G3 e G4
	Intra Grupos	100504.66	39941			G3> G1; G2 <G4 G4> G1, G2 e G3
<i>Relação Docente/Aluno</i>	Entre Grupos	424.682	3	62.337	.000	G1 <G2, G3 e G4 G2> G1; G2 <G3 e G4
	Intra Grupos	87240.858	38417			G3> G1, G2 e G4 G4> G1 e G2; G4 <G3
<i>Profundidade na Abordagem dos Assuntos</i>	Entre Grupos	1872.710	3	175.717	.000	G1 <G2, G3 e G4 G2> G1; G2 <G3 e G4
	Intra Grupos	77372.683	38189			G3> G1 e G2; G3 <G4 G4> G1, G2 e G3
<i>Avaliação/Classificações</i>	Entre Grupos	1500.423	3	199.214	.000	G1 <G2, G3 e G4 G2> G1; G2 <G3 e G4
	Intra Grupos	94010.907	37446			G3> G1 e G2; G3 <G4 G4> G1, G2 e G3
<i>Trabalhos/Leituras</i>	Entre Grupos	2623.782	3	368.938	.000	G1 <G2, G3 e G4 G2> G1; G2 <G3 e G4
	Intra Grupos	88818.296	37467			G3> G1 e G2; G3 <G4 G4> G1, G2 e G3
<i>Carga Trabalho/Dificuldade</i>	Entre Grupos	154.610	3	74.332	.000	G1 <G2, G3 e G4 G2> G1; G2 <G3 e G4
	Intra Grupos	26031.958	37546			G3> G1 e G2; G3 <G4 G4> G1, G2 e G3

Na análise dos resultados por áreas de curso, os resultados obtidos apresentam-se diferenciados do ponto de vista estatístico em todas as dimensões do PEA. Para entendermos melhor o sentido das diferenças encontradas nas dimensões recorreremos ao teste de contrastes (*scheffee*). Estas diferenças apontam para contrastes a favor dos alunos do grupo das Ciências da Educação por comparação com cada um dos restantes grupos. Por sua vez, em relação à generalidade das dimensões do PEA, os alunos da área das Ciências e Engenharia tendem a apresentar percepções tendencialmente mais desfavoráveis acerca do desempenho docente e do seu confronto pessoal com os desafios colocados pelas condições do ensino/aprendizagem comparativamente aos restantes grupos. Mais uma vez este dado é consistente com os resultados da investigação, sugerindo que alunos das Ciências e Engenharias tendem a avaliar menos positivamente os docentes comparativamente aos alunos das Humanidades e das Ciências Sociais (Cramer & Alexitch, 2000), embora tais diferenças não sejam muito significativas (Cashin, 1990; 1992; Centra, 1993; Ory, 2001). Outros autores como Deberg e Wilson (1990) defendem que esta diferenciação pode ser condicionada pela percepção de maior grau de dificuldade associada às disciplinas das áreas das Ciências. De salientar que, é apenas na dimensão da Relação Docente/Aluno que os alunos da área das Ciências Sociais e Humanas avaliam mais positivamente comparativamente ao grupo das Ciências da Educação e aos restantes grupos. Este resultado parece evidenciar que nestas áreas científicas, o relacionamento é provavelmente uma dimensão mais apreciada e valorizada pelos próprios alunos quando são convidados a avaliar a qualidade do ensino, e em particular o desempenho docente.

“A ciência? Ao fim ao cabo, o que é ela senão uma longa e sistemática curiosidade?”

André Maurois (1885-1967)

Conclusão

A expansão e a diversificação do Ensino Superior em Portugal, aquando da entrada na União Europeia, conduziram à evolução algo “descontrolada” da oferta formativa, de professores e de alunos. Com a heterogeneidade instalada, a que os autores denominaram “multiplicidade anárquica” ou “estrangulamento e ruptura” (Braga da Cruz *et al.*, 1995), aliadas ao crescimento desenfreado, não houve tempo e espaço para a monitorização da qualidade. É neste quadro de ponderação da quantidade-qualidade que tem granjeado força o movimento de questionamento da qualidade do Ensino Superior e a instauração progressiva de mecanismos de avaliação. Estes mecanismos, internos e externos à instituição, constituem hoje referenciais de qualidade do sistema do seu financiamento público. De entre toda a diversidade de possíveis definições de qualidade parece universalmente aceite que a investigação (da instituição como um todo) e o ensino (tal como experienciado pelos alunos) constituem dois ingredientes fundamentais dessa mesma definição (Perry, 1991).

De acordo com Perry (1995 in Ferreira, 2003) a qualidade é o maior e o mais interessante desafio do Ensino Superior na próxima década. Com efeito, as questões que dominam actualmente a discussão em torno da avaliação a qualidade referem-se sobretudo à prestação de contas e à melhoria das Instituições. Por sua vez, já em 1993, Veiga Simão afirmava que um resultado inevitável da avaliação seria exigir que a Universidade se abrisse mais ao exterior, à sociedade civil, facultando dados qualitativos e quantitativos respeitantes ao cumprimento da sua missão.

Assim, o mundo universitário tornou-se cada vez mais competitivo, com estudantes mais conscientes dos seus direitos e necessidades, e com maior poder reivindicativo (McKeachie, 1997). Por sua vez, na sociedade dita de conhecimento e de aprendizagem, em que a qualidade dos serviços oferecidos passa a ser um critério de escolha para os estudantes e suas famílias, as instituições, numa lógica de competitividade interinstitucional, são obrigadas a estabelecer planos de avaliação de qualidade e a alterar as suas posturas tradicionais se quiserem sobreviver.

Ao nível Europeu já se discute o papel central das Universidades na reforma já iniciada com a Declaração de Bolonha, onde a qualidade se impõe como um aspecto nuclear. Neste contexto, ganha particular relevância a função docente, admitindo-se o protagonismo do seu papel neste processo de reforma pela convicção de que as reformas *top-down* em curso devem ser protagonizadas por quem

está dentro do sistema, e neste caso, os docentes são agentes privilegiados para o fazer. Face às considerações feitas, emerge a interrogação: Será então o docente a interface onde serão acoplados todos os interesses, expectativas e objectivos que recaem no ensino superior? E, de entre todos, quais devem ser privilegiados na sua actividade? O ensino, a investigação ou a prestação de serviço à comunidade, ou ainda, as diferentes ponderações em cada uma das áreas. E, de todos os agentes ou *stakeholders*, quem será responsável por determinar os contornos dessa actividade?

No sentido de responder às preocupações com a promoção de uma cultura progressiva de endogeneização da qualidade, a Universidade do Minho tem procurado desenvolver, de há 14 anos a esta parte, dispositivos de avaliação inspirados no “modelo holandês” (modelo procura estabelecer um equilíbrio dinâmico entre a avaliação interna e avaliação externa) e a organização de acções de formação pedagógica de docentes⁴⁶. A avaliação das percepções do ensino ministrado pelos estudantes tem sido uma das peças do sistema instalado.

Todavia, a qualidade do ensino é um conceito multifacetado e, quando muito, os questionários apreendem uma parte do todo que o define. Assim, qualquer dispositivo de avaliação deve considerar fontes e instrumentos de avaliação diversificados que permitam triangular e confrontar as diversas informações (aspectos relativos à disciplina, ao curso e à instituição). Por sua vez, formas alternativas de avaliação do ensino (auto-avaliação docente, avaliação pelos pares, administradores e observadores experientes/externos) devem ser consideradas na agenda de preocupações dos diferentes elementos com responsabilidades nesta área. Em particular, no que concerne à avaliação da competência pedagógica, observa-se a emergência de outros processos de avaliação (e.g. *portfólios* sobre as actividades de ensino; elaboração de *software* didáctico; publicação de artigos de inovação sobre o ensino; participação em colóquios com apresentação de trabalhos; actividades de internacionalização dos cursos; desempenho de funções de tutoria, entre outras). Contudo, as questões associadas ao desempenho pedagógico do docente no Ensino Superior em Portugal, não têm sido utilizadas como critério fundamental para a sua promoção profissional e, por isso, emerge a necessidade de equacionar uma formação docente em novos moldes e em resposta a um contexto académico em que o peso da preparação pedagógica seja semelhante ao peso atribuído à preparação científica e de investigação.

Assim, e sem perder de vista a necessidade de uma avaliação multidimensional, uma das vias de aferição da qualidade do ensino decorre da análise das percepções dos alunos veiculadas através

⁴⁶ (cf. Despacho RT-04/2004: promove a formação de Docentes, seja na área do planeamento curricular e na planificação do ensino/aprendizagem, seja nos métodos de avaliação, na utilização das tecnologias ou na comunicação, integra-se num conjunto de acções orientadas genericamente ao vector da Qualidade).

da aplicação de questionários. Aliás, esta forma de avaliação tem gozado de maior suporte empírico, tem sido assumida pelos governos como indicador válido de desempenho e como uma medida directa de satisfação face à qualidade do ensino ministrado (Ramsden, 1991; Simão, 2003ab).

As vantagens desta avaliação foram documentadas em vários momentos desta dissertação, contudo, emergem dois a três aspectos fundamentais que justificam a continuidade da sua aplicação e monitorização. No essencial, sendo uma avaliação de papel-lápis é mais económica do que o recurso a observações de sala de aula que envolvem o recurso a observadores especializados e experientes. Sem prejuízo de uma avaliação transversal, fiável e rigorosa, este critério é fundamental numa instituição de Ensino Superior, como é o caso da Universidade do Minho, que apresenta um volume significativo de alunos que frequenta regularmente as aulas (uma média de 43 000 questionários a recolher por semestre). Adicionalmente, é defendido que a avaliação da qualidade docente inferida com base nas percepções dos alunos, quando positivamente assumida, pode cumprir vários objectivos de melhoria da qualidade do ensino, fornecendo aos docentes informações úteis para a melhoria das suas práticas e à instituição informações para a gestão mais adequada dos seus recursos (e.g. órgãos de gestão pedagógica).

No sentido de contribuir para a melhoria dos dispositivos de avaliação do ensino na Universidade do Minho, trabalhámos a adaptação e validação de um modelo de questionário que pudesse responder às preocupações/desafios actuais do Ensino Superior. Por outro lado, face à ideia generalizada que nem sempre os estudantes entendem o sentido dos itens colocados, ou não têm a necessária preparação técnica para o efeito, houve a preocupação de reunir estudantes e auscultá-los a propósito da coerência e compreensão do conteúdo dos itens. Neste sentido, procedemos à adaptação do questionário *Percepções do Ensino pelos Alunos* (PEA), inspirado no questionário *Student's Evaluation of Educational Quality* (SEEQ; Marsh, 1987, 1991, 2001), aliás, referência internacional no domínio da avaliação da qualidade educativa. O estudo de tradução, validação e sugestão de mudanças ao questionário consubstanciou a parte empírica desta dissertação com base numa amostra de 41 962 questionários.

Os resultados da análise da distribuição dos itens mostraram uma menor variabilidade de alguns itens do PEA, situação que pode sugerir a redução do número de pontos da escala original (1 a 9) sugerindo-se ainda, a possibilidade do aluno não emitir opinião, para distinguir esta tomada de decisão dum acto menos ou não consciente de resposta na situação efectiva de falta de informação ou dificuldade efectiva em se posicionar em determinados itens.

No que concerne à análise de consistência interna do PEA, os resultados mostraram níveis elevados de consistência (.97), situação que poderá decorrer do facto de estarmos perante um questionário com um número bastante elevado de itens, devendo, por isso, ser considerada a diminuição do tamanho do questionário, ou em alternativa tornar os itens mais heterogêneos na sua formulação e aspectos comportamentais avaliados. Aliás, este procedimento é tão mais importante quando se pretende validar um modelo de questionário tipo “*screening*” de formato mais curto como resposta à preocupação de avaliação institucional das instituições de Ensino Superior.

Da mesma forma, os índices de validade interna dos itens, calculados através de um coeficiente de correlação corrigido dos resultados item x total da dimensão/faceta, suplantaram claramente os índices mínimos exigidos, tendo-se observado uma amplitude de valores entre .23 e .83. Por sua vez, uma análise de natureza correlacional, cruzando as dimensões do PEA com o ano da disciplina e a tipologia de aula, permitiu a identificação de itens com coeficientes de correlação significativos que se encontrariam mais sobreponíveis em termos de conteúdo. Sem perder de vista a recolha de informação relevante, a preocupação foi avançar com sugestões no sentido da diminuição do número de itens por dimensão do PEA.

Os resultados da análise factorial dos 33 itens originais permitiram a identificação de cinco factores ou dimensões, reunindo o primeiro um conjunto de itens relacionados com a qualidade da aprendizagem; o segundo com o valor sócio-relacional do processo ensino/aprendizagem; o terceiro com valor do ensino e dos aspectos curriculares (transversalidade e profundidade dos assuntos abordados); o quarto com a avaliação e trabalhos/leituras e o quinto com o confronto pessoal do aluno com as características (exigências) da disciplina.

Historicamente, os órgãos de gestão pedagógica e docentes têm recebido das entidades responsáveis pelos processos avaliativos, relatórios circunscritos às médias e percentagens de respostas tomando cada item individualmente. Esta situação favorece a parcialidade e a fragmentação na comunicação dos resultados, não favorecendo a emergência de análises globais e de observação de tendências com base numa conceptualização multidimensional da qualidade do ensino. Por exemplo, num determinado nível de desempenho global o docente avaliado pode apresentar diferentes níveis de desempenho conforme a dimensão/faceta considerada. Por sua vez, a discriminação dos valores obtidos por dimensões (como ocorre agora com o presente questionário validado), a par de uma avaliação mais global, permitem acompanhar o desempenho docente numa lógica mais formativa favorável à implementação de uma política de acompanhamento do desempenho, menorizando julgamentos que favoreçam procedimentos de *ranking* e controlo institucional. Por sua vez,

entendemos que a realização deste trabalho, ao permitir definir um questionário suficientemente validado, viabiliza alguma investigação mais sistemática nesta área da avaliação da qualidade do ensino, tal como percepcionado pelos alunos, como também na fundamentação de medidas de intervenção assumidas como pertinentes.

Um dos aspectos mais consistentes e relevantes do conjunto de resultados obtidos é a avaliação positiva na generalidade dos itens, com especial destaque para a dimensão Relação Docente/Aluno. Este dado por si, constitui um indicador de satisfação com o ensino ministrado, e em particular, com a qualidade das relações interpessoais estabelecidas entre alunos e docentes na Universidade do Minho. Mesmo assim, as análises de variância dos resultados realizadas permitiram-nos encontrar oscilações significativas nos valores consoante o ano curricular em que a disciplina avaliada se integra, a natureza teórica ou prática da disciplina e o agrupamento de cursos a que respeita.

Os resultados da análise de variância sugerem que os alunos de anos curriculares mais avançados e no contexto das aulas práticas percepcionam mais positivamente a Organização/Clareza do docente, o mesmo ocorrendo na dimensão Profundidade na Abordagem dos Assuntos (só que aqui em relação às aulas teóricas) e na dimensão Trabalhos/Leituras. Nas demais seis dimensões do PEA (Relevância da Aprendizagem; Empenhamento Docente; Interacção da Turma; Relação Docente/Aluno; Avaliação/Classificações e Carga Trabalho/Dificuldade) observa-se um efeito simultâneo do ano e tipo de aula, apontando-se na generalidade dos casos para percepções da qualidade do ensino mais positivas por parte dos alunos no quadro das disciplinas dos anos curriculares intermédios e finais dos cursos e no contexto das aulas práticas. Por sua vez, recorrendo ao teste de contrastes (*scheffee*), as diferenças encontradas apontam para a manifestação de percepções do ensino mais positivas junto dos alunos das áreas das Ciências da Educação comparativamente com cada um dos restantes grupos considerados (Ciências e Engenharia; Ciências Económicas e Ciências Sociais e Humanas) na generalidade das nove dimensões do PEA.

No que se refere aos desenvolvimentos futuros na área, importa reafirmar que o modelo proposto constitui apenas uma das possíveis formas de avaliação do desempenho docente. Tal pressuposto reforça a utilidade desta medida enquanto importante instrumento de “*screening*” de eventuais problemas ou pontos fortes do ensino/aprendizagem experienciados pelos alunos no confronto com a vida universitária. Por sua vez, importa no futuro, e em jeito de sugestão, analisar conjuntamente os resultados obtidos neste questionário com outras formas de avaliação do desempenho docente e com o compromisso institucional dos docentes. Devemos, ainda, considerar

em futuros desenvolvimentos, e à semelhança do que se tem verificado com outros modelos de questionário na literatura, prever a possibilidade de desenvolver esforços de reorganização das dimensões/facetas, a partir da inclusão de novos constructos que possam revelar-se importantes na conceptualização e explicação do fenómeno em estudo no contexto cultural português, e desejavelmente, numa perspectiva longitudinal (e.g. variáveis mais directamente associadas à aprendizagem do aluno e às condições genéricas do ensino/aprendizagem, nomeadamente a abertura à inovação e a promoção da excelência nas actividades académicas, etc.). Efectivamente, do ensino centrado no docente não podemos passar para o extremo oposto para o ensino centrado no aluno (Bligh, 2003 in Ferreira, 2003), tal como os alunos, também o docente necessita de um ambiente de trabalho positivo em que seja possível progredir e usar totalmente as competências adquiridas. Um remate óbvio seria que, no futuro não muito distante pudéssemos usar tudo o que já sabemos sobre ensino/aprendizagem, apesar das contingências que envolvem o desenvolvimento institucional, tais como as prioridades consideradas no ambiente académico, os recursos existentes e, sobretudo, as questões pedagógicas na promoção na carreira académica.

Por último, procuramos destacar as potencialidades e as limitações deste estudo. Com efeito, a recolha de uma amostra significativa de questionários representa uma mais valia para a generalização dos resultados obtidos. Com efeito, os dados procuram reflectir a heterogeneidade e transversalidade científico-pedagógica dos cursos da Universidade do Minho. No entanto, o estudo padece de algumas limitações, nomeadamente o facto dos questionários terem sido recolhidos apenas numa instituição universitária e não de outras instituições do ensino superior português (e.g. Institutos Politécnicos, públicos e privados). Por outro lado, não se procedeu à confrontação necessária das percepções dos alunos e dos docentes no sentido de garantir uma avaliação mais completa do processo ensino/aprendizagem. Assim, mais do que um ponto de partida, importa, de acordo com um pensamento de Little Guidding, T. S. Elliot, *não cessar a exploração, sendo que o fim de toda a nossa pesquisa, consistirá em chegar ao ponto de partida, e conhecer o lugar pela primeira vez!*

Bibliografia

- Abrami, P. C. & d'Apollonia, S. (1990). The dimensionality of ratings and their use in personnel decisions. In M. Theall & J. Franklin (Eds.), *Student ratings of instruction: Issues for improve practice. New Directions for Teaching and Learning*, pp. 97-111. San Francisco: Jossey Bass.
- Abrami, P. C. & d'Apollonia, S. (1991). Multidimensional student's evaluations of teaching effectiveness: generalizability of n=1 research. Comment on Marsh (1991). *Journal of Educational Psychology*, 30, 221-227.
- Abrami, P. C. d'Apollonia, S. & Cohen, P. A. (1990). Validity of student ratings of instruction: What we know and what we do not. *Journal of Educational Psychology*, 82 (2), 219-231.
- Abrami, P. C. Dickens, W. J., Perry, R. P. & Leventhal, L. (1980). Do teacher standards for assigning grades affect student evaluations of teaching? *Journal of Educational Psychology*, 72 (1), 107-118.
- Abrami, P. C. Perry, R. P. & Leventhal, L. (1982). The relationship between student personality characteristics, teacher ratings, and student achievement. *Journal of Educational Psychology*, 74, 111-125.
- Abrantes, P. & Valente, M. (2000). *Estudo sobre a avaliação dos docentes do ensino superior: Desenvolvimento de instrumentos e avaliação de desempenho*. Relatório Final. Lisboa: DGES.
- Alarcão, I. (1996). A construção do conhecimento profissional. In M. R. Delgado-Martins, M. I. Rocheta & D. R. Pereira (Orgs.), *Formar professores de português, hoje*. (pp.91-95). Lisboa: Edições Colibri.
- Aleamoni, L. M. & Hexner, P. Z. (1980). A review of the research on student evaluation and a report on the effect of different sets of instructions on student course and instructor evaluation. *Instructional Science*, 9, 67-84.
- Aleamoni, L. M. (1999). Student rating myths versus research facts from 1924 to 1998. *Journal of Personnel Evaluation in Education*, 13 (2), 153-166.
- Almeida, L. S. & Freire, T. (2001). *Escalas de avaliação: Construção e validação*. In E. M. Fernandes & L. S. Almeida (Eds), Braga: Universidade do Minho, Centro de Estudos em Educação e Psicologia.
- Almeida, L. S. (2002). *Formatar o ensino a pensar na aprendizagem*. In A. S. Pouzada, L. S. Almeida & Vasconcelos R. M. (Eds). *Contextos e dinâmicas da vida académica*. Guimarães: Universidade do Minho.

- Almeida, L. S., Vasconcelos, R. Machado, C., Soares, A. P. & Morais, N. (2002). Perfil escolar e sócio-demográfico dos candidatos ao Ensino Superior: O caso dos estudantes da Universidade do Minho. In A. S. Pouzada, L. S. Almeida & Vasconcelos R. M. (Eds). *Contextos e dinâmicas da vida académica*. Guimarães: Universidade do Minho.
- Amaral, A. & Rosa, M. (2004). A alteração do relacionamento entre as instituições do ensino superior e o Estado e as suas consequências em torno da qualidade. In Seminário *Consequências institucionais da avaliação das universidades públicas*. Coimbra: Fundação das Universidades Portuguesas, 22 de Maio.
- Amaral, A. (1998). *Sistemas Europeus de Avaliação da Qualidade*. Editora Fundação das Universidades Portuguesas, CIPES.
- Andersen, K. & Miller, E. (1997). Gender and student evaluations of teaching. *Political Science & Politics*, 30 (2).
- Angulo, F., Fernández, J. & Martínez, M. R. (Eds). (1987). *La evaluación de la enseñanza universitaria*. Madrid: Instituto de Ciencias de la Educación.
- Arroteia, J. C. (1996). *O ensino superior em Portugal*. Aveiro: Universidade de Aveiro
- Atkins, M. (1987). *Effective teaching in higher education*. New York: Methuen.
- Aylett, R. & Gregory, K. (Eds.), (1996). *Evaluating teacher quality in higher education*. London: Falmer Press.
- Ballantyne, R., Borthwick, J. & Packer, J. (2000). Beyond student evaluation of teaching identifying and addressing academic staff development needs. *Assessment & Evaluation in Higher Education*, 25 (3), 221-236.
- Balsa, C., Simões, J. A. Nunes, R. C. & Campos, R. (2001). *Perfil dos estudantes do Ensino Superior: Desigualdades e diferenciação*. Lisboa: CEOS, Edições Colibri.
- Basow, S. & Silberg, N. T. (1987). Student evaluations of college professors: Are male and female professor's rates differently? *Journal of Educational Psychology*, 79 (3), 308-314.
- Bedggood, R. E. & Pollard, R. J. (1999). Uses and misuses of student opinion surveys in eight australian universities. *Australian Journal of Education*, 43 (2), pp.129-141.
- Bedin, V. (1995). A avaliação dos primeiros ciclos universitários: Uma prática social. In A. Estrela & P. Rodrigues (Coords.), *Para uma fundamentação da Avaliação em Educação*. Lisboa: Edições Colibri (pp. 61-74).
- Bennett, W. E. (1982). Student perceptions of and expectations for male and female instructors: Evidence relating to the question of gender bias in teaching evaluation. *Journal of Educational Psychology*, 74 (2), 170-179.

- Bernard, H. (1992). *Processus d'évaluation de l'enseignement supérieur. Théorie et pratique*. Laval (Quebec): Éditions Études Vivantes.
- Bernard, M. E. Keefauver, L. W., Elsworth, G. & Maylor, F. D. (1981). Sex role behavior and gender in teachers-student evaluations. *Journal of Educational Psychology*, 73, 681-696.
- Bernard, R. & Normand, S. (1998). Une approche distincte de présentation et d'interprétation des résultats de l'appréciation de l'enseignement par les étudiants. *Mesure et Évaluation en Éducation*, 21(2), 85-117.
- Bernard, R. (1996). La nouvelle culture de l'évaluation: Passage difficile de la modernité à la postmodernité, 19 (2), 99-109.
- Biggs, J. B. (1990). Teaching for desired learning outcomes. In Entwistle (Ed.), *Handbook of educational ideas and practices*. London: Routledge.
- Biggs, J. B. (1999). *Teaching for quality learning at University. What the student does*. London. Society for Research into Higher Education & Open University Press.
- Bireaud, A. (1995). *Os métodos pedagógicos no ensino superior*. Porto: Porto Editora.
- Bonniol, J. J. & Vial, M. (2001). *Modelos de avaliação. Textos fundamentais*. Porto Alegre: Artmed Editora.
- Braga da Cruz, M., Cruzeiro, M. E., Ramos, A., Leandro, E., Nunes, J. S., Matias, N., Pedroso, P., Robinson, M. G. & Cavaco, V. (1995). *O desenvolvimento do ensino superior em Portugal: Situação e problemas de acesso*. Lisboa: Ministério da Educação, Departamento de Programação e Gestão Financeira.
- Braskamp, L. A. & Ory, J. C. (1994). *Assessing Faculty Work: Enhancing individual and institutional performance*. San Francisco: Jossey-Bass.
- Braskamp, L. A., Branderburg, D. C. & Ory, J. C. (1985). *Evaluating teaching effectiveness: A practical guide*. Beverly Hills, CA: Sage.
- Brennan, J. & Shah, T. (2000). *Managing quality in higher education. An international perspective on institutional assessment and change*. Buckingham: SRHE & Open University Press.
- Cachapuz, A. (2001). Em defesa do aperfeiçoamento pedagógico dos docentes do ensino superior. In C. Reimão (org.) (2001). *A Formação Pedagógica dos Professores do ensino superior*. Lisboa: Edições Colibri.
- Caetano, A. & Vala, J. (2002). *Gestão dos recursos humanos: contextos, processos e técnicas*. Lisboa: Editora RH.

- Canário, R. (2001). *A prática profissional na formação dos professores*. In B. P. Campos (Org.). *Formação profissional de professores no ensino superior*. Cadernos de formação de professores. INAFOP. Porto: Porto Editora, pp. 31-45.
- Cantero, J. M., Deus, R. M. & Abalde, E. (2002). Evaluación docente vs. evaluación de la calidad. http://www.uv.es/RELIEVE/v8n2/RELIEVEv8n2_4.htm
- Carson, B. H. (1996). Thirty years of stories: The professor's place in student memories. *Change*, 28 (6), 10-17.
- Cashin, W. E. & Downey, R. G. (1992). Using global student rating items for summative evaluation. *Journal of Educational Psychology*, 84 (4), 563-572.
- Cashin, W. E. (1979). *Motivating students*. IDEA Paper (1), Manhattan: Kansas State University, Center for Faculty Evaluation and Development.
- Cashin, W. E. (1989). Defining and evaluating college teaching. IDEA Paper (21), Manhattan: KS: Kansas State University, Center for Faculty Evaluation and Development.
- Cashin, W. E. (1990). Students do rate different academic fields differently. In M. Theall & J. Franklin (Eds). *Students ratings of instruction: Issues for improving practice. New Directions for Teaching and Learning*, (43). San Francisco: Jossey-Bass.
- Cashin, W. E. (1992). Student ratings: The need for comparative data. *Instructional Evaluation and Faculty Development*, 12 (2), 1-6.
- Cashin, W. E. (1995). *Student ratings of teaching: The research revisited (IDEA Paper No 32)*. Manhattan: Kansas State University, Center for Faculty Evaluation and Development.
- Castro, P. T. (1993). Exames das universidades e dos seus programas. *Boletim da Universidade do Porto*, 3 (17), 24-27.
- Centra, J. A. (1979). *Determining faculty effectiveness*. San Francisco, C. A: Jossey Bass.
- Centra, J. A. (1993). *Reflective faculty evaluation: Enhancing teaching and determining faculty effectiveness*. San Francisco, C. A: Jossey Bass.
- Centra, J. A. (2003). Will teachers receive higher student evaluations by giving higher grades and less course work? *Research in Higher Education*, 44 (5).
- Chonko, L. B., Tanner, J. F. & Davis, R. (2002). What are they thinking? Student's expectations and self-assessments. *Journal of Education for Business*, 77 (5).
- CIPES. (2003). *Avaliação, revisão e consolidação da legislação do ensino superior*. Porto: CIPES
- Ckickering, A. W. & Gamson, Z. F. (1987). Seven principles for good practice. *AAHE Bulletin*, 39 (7), 3-7.
- Coaldrake, P. & Stedman, L. (1998). *On the Brink: Australia's universities confronting their future*. St. Lucia, Qld. University of Queensland Press.

- Coffey, M. & Gibbs, G. (2001). The evaluation of student evaluation of educational quality questionnaire (SEEQ) in UK higher education. *Assessment & Evaluation in Higher Education*, 26 (1), 89-93.
- Cohen, P. A. (1980). Effectiveness of student rating feedback for improving college instruction. A meta-analysis of multisection validity studies. *Research in Higher Education*, 13 (4), 321-341.
- Coles, C. (2002). Variability of student ratings of accounting teaching: Evidence from a scottish business school. *International Journal of Management Education*, 2 (2).
- Conceição, P., Durão, D. F. G., Heitor, M. V. & Santos, F. (1998). *Novas ideias para a Universidade*. Lisboa: Instituto Superior Técnico.
- Costin, F., Greenough, W. T., Menges, R. J. (1971). Student ratings of college teaching: Reliability, validity, and usefulness. *Review of Educational Research*, 41 (5), 511-536.
- Cowan, J. (2004). Avaliação psicológica de docentes. *Encontro sobre a Avaliação Pedagógica no Ensino Superior*. Lisboa: Sindicato Nacional do Ensino Superior, 30-31 de Janeiro.
- Cramer, K. M. & Alexitch, L. R. (2000). Student evaluations of college professors: Identifying sources of bias. *Canadian Journal of Higher Education*, 30 (2).
- Cranton, P. A. & Hillgarten, W. (1981). The relationships between student rating and instruction behaviour: Implications for improving teaching. *Canadian Journal of Higher Education*, 11, 73-81.
- Crespo, V. (2003). *Ganhar Bolonha, ganhar o futuro. O ensino superior no espaço europeu*. Lisboa: Gradiva.
- Crumbley, L., Henry, B. K. & Stanley, H. K. (2001). Student's perceptions of the evaluation of college teaching. *Quality Assurance in Education*, 9 (4), 197.
- Curado, A. P. (2000). *Profissionalidade dos docentes: Que avaliar? Resultados de um estudo interativo de Delphi*. Lisboa: Ministério da Educação, Instituto de Inovação Educacional.
- D'Apollonia, S. & Abrami, P. (1997). Navigating student ratings of instruction. *American Psychologist*, 52 (11), 1198-1208.
- Das, M. & Das, H. (2001). Business student's perceptions of best university professors: Does gender role matter? *Sex roles*, 45 (9/10), 665-676.
- Davis, B. G. (1993). *Tools for teaching*. San Francisco: Jossey-Bass.
- De Ketéle, J. M. (2004). *Uma pedagogia da integração: Competências e aquisições no ensino*. (Trad. por Carolina Huang). Porto Alegre: Artmed Editora.
- Deberg, C. L. & Wilson, J. R. (1990). An empirical investigation of the potential confounding variables in student evaluation of teaching. *Journal of Accounting Education*, 6, 37-62.
- Decreto-Lei nº205/1998, de 11 de Julho. Cria o *Conselho Nacional de Avaliação do Ensino Superior* (CNAVES) Lisboa: Diário da República.

- Departamento de Avaliação Prospectiva e Planeamento, (2000).
- Dewey, J. (1956). *The child and the curriculum: The school and society*. Chicago: University of Chicago Press.
- Doyle, K. O. (1983). *Evaluating teaching*. San Francisco: New Lexington Press.
- ECDU, (1979). Estatuto da Carreira Docente Universitária.
- Ellington, H. (2000). How to become an excellent tertiary-level teacher. Seven golden rules for university and colleague lecturers. *Journal of Further and Higher Education*, 24 (3).
- Elton, L. R. & Partington, P. (1991). *Teaching standards and excellence. Higher Education: Developing a culture for quality*. Occasional green paper n°1. Sheffield: University Staff Development and Training Unit.
- Emery, C.R., Kramer, T. R. & Tian, R. G. (2003). Return to academic standards: A critique of student evaluations of teaching effectiveness. *Quality Assurance in Education*, 11 (1), 37-46.
- Estanque, E. & Nunes, J. A. (2001). A universidade perante a transformação social e as orientações dos estudantes: O caso da Universidade de Coimbra. In A. Gonçalves, L. S. Almeida, R. Vasconcelos & S. Caires (Eds.), *Da universidade para o mundo do trabalho: Desafios para um diálogo*. Braga: Universidade do Minho, Conselho Académico.
- Estrela, M. T. & Simão, A. M. V. (2003). Algumas reflexões sobre práticas de avaliação do ensino universitário e dos docentes a partir da informação recolhida no Projecto EVALUE. *Revista Portuguesa de Educação*, 16 (1), 101-120.
- Estrela, M. T. (1986). Algumas considerações sobre o conceito de profissionalismo docente. *Separata da Revista Portuguesa de Pedagogia*, 301-310.
- Eurydice (2000). *Key data on education in the European union*. Brussels: European Commission.
- Felder, R. (2004). How to evaluate teaching. *Chemical Engineering Education*, 38 (3), 202-204.
- Feldman, K. A. (1976). Grades and college student's evaluations of their courses and teachers. *Research in Higher Education*, 4, 69-111.
- Feldman, K. A. (1977). Consistency and variability among college students in rating their teachers and courses: A review and analysis. *Research in Higher Education*, 6, 223-274.
- Feldman, K. A. (1978). Course characteristics and college student's ratings of their teachers and courses: What we know and what we do not. *Research in Higher Education*, 9, 199-242.
- Feldman, K. A. (1983). Seniority and experience of college teachers as related to evaluations they receive from their students. *Research in Higher Education*, 18, 3-124.
- Feldman, K. A. (1984). Class size and college student's evaluations of teachers and courses: A closer look". *Research in Higher Education*, 21, 45-116.

- Feldman, K. A. (1989a). Instructional effectiveness of college teachers as judged by teachers themselves, current and former students, colleagues, administrators, and external (neutral) observers. *Research in Higher Education*, 30, 137-194.
- Feldman, K. A. (1989b). Association between student ratings of specific instructional dimensions and student achievement: Redefining and extending the synthesis of data from multisection validity studies. *Research in Higher Education*, 30, 583-645.
- Feldman, K. A. (1993). College student's views of male and male college teachers: Evidence from student's evaluations of their classroom teachers. *Research in Higher Education*, 34, 151-211.
- Fernandes, A. T. (Coord.) (2001). *Estudantes do Ensino Superior no Porto: Representações e práticas culturais*. Porto: Edições Afrontamento.
- Fernández, J. & Mateo, M. A. (1994). The circular evaluative process of higher education: A spanish sample. In OECD (Ed.), *Evaluation and decision making process in higher education: French, german and spanish experiences*. Paris: OECD.
- Fernández, J. & Mateo, M. A. (1997). Student and faculty gender in ratings of university teaching quality. *Sex roles*, 37, 11/12.
- Fernández, J. (1992). Evaluation and decision making in the Complutense University of Madrid. *Higher Education Management*, 4, 336-345.
- Fernández, J., Mateo, M. A. & Muñoz, J. (1995). Evaluation of the academic setting in Spain. *European Journal of Psychological Assessment*, 11, 134-138.
- Fernández, J., Mateo, M. A. & Muñoz, J. (1996). Valoración por parte del profesorado de la evaluación docente realizada por los alumnos. *Psicothema*, 8 (1), 167-172.
- Ferreira, A. (2003). *O Papel do Professor na Educação Médica: Contributos para um ensino de qualidade no ensino superior*. Dissertação de Mestrado. Braga: Universidade do Minho.
- Ferreira, A. F. (5 de Dezembro de 2004). *Carpaccio Bolonhês*, in Semanário Económico.
- Fiolhais, C. (2004). Condicionantes do sucesso académico. *Encontro sobre a Avaliação Pedagógica no Ensino Superior*. Lisboa: Sindicato Nacional do Ensino Superior, 30-31 de Janeiro.
- Fleming, N. D. (2003). Establishing rapport: Personal interaction and learning. Idea Paper (39). Idea Center. www.vark-learn.com.
- Fonte, R. & Vasconcelos, R. (2001). A formação pedagógica de docentes universitários. A experiência na Escola de Engenharia da Universidade do Minho. Braga: Universidade do Minho, Conselho Académico.
- Fraser, B. J. (1991). Two decades of classroom environment research. B. J. Fraser & H. J. Walberg (Eds.), *Educational environments: Evaluation, antecedents, and consequences*. Oxford: Pergamon.

- Freitas, M. (2004). *Estratégia Nacional de Educação para o Desenvolvimento Sustentável. Resumo de anteprojecto de estrutura geral e linhas orientadoras*. Braga: Universidade do Minho, Instituto de Educação e Psicologia.
- Gago, J. M., Amaral, J. F., Grácio, S., Rodrigues, M. J., Fernandes, L., Ruivo, B., Ambrósio, T., Silva, C. M., Duarte, T., Teixeira, M., Proença, L., Alves, M. G. & Lisboa, M. (1994). *Prospectiva do ensino superior em Portugal*. Lisboa: Ministério da Educação, Departamento de Programação e Gestão Financeira.
- Gama, A. (2004). Modelos de avaliação do ensino e da docência. *Encontro sobre a Avaliação Pedagógica no Ensino Superior*. Lisboa: Sindicato Nacional do Ensino Superior, 30-31 de Janeiro.
- García, J. M. & Congosto, E. (2000). Evaluación y calidad del profesorado. In T. González Ramírez (Coord.) *Evaluación y gestión de la calidad educativa. Un enfoque metodológico*. Ed. Aljibe. Málaga. 127-157.
- Gillmore, G. M., Kane, M. T. & Naccarato, R. W. (1978). The generalizability of student ratings of instruction: Estimates of teacher and course components. *Journal of Educational Measurement*, 15, 1-13.
- Goldberg, G. & Callahan, J. (1991). Objectivity of student evaluations of instructors. *Journal of Education for Business*, 66 (6), 377-379.
- Gomes, J. F. (2003). *Estatutos dos Docentes do Ensino Superior Público: Princípios orientadores* In Seminário. *Repensar Bolonha*, Porto: Universidade do Porto, 20 de Abril.
- Gonçalves, N. I. (2004). *Modelos de avaliação do ensino e da docência. Encontro sobre a Avaliação Pedagógica no Ensino Superior*. Lisboa: Sindicato Nacional do Ensino Superior, 30-31 de Janeiro.
- Gorham, J. (1988). The relationship between verbal teacher immediacy behaviors and student learning. *Communication Education*, 37, 40-53.
- Green, D. (1995). What is quality in higher education? Concepts, policy and practice. In D. Green (Ed.). *What is quality in Higher Education?* Buckingham: SRHE & Open University Press (3-20).
- Greenwald, A. G. (1997). Validity concerns and usefulness of student ratings of instruction. *American Psychologist*, 52 (11), 1182-1186.
- Greenwald, A., G. & Gillmore, G. (1997). Grading leniency is a removable contaminant of student ratings. *American Psychologist*, 52 (11), 1209-1217.
- Hadji, C. (1994). *A avaliação, regras do jogo*. Porto: Porto Editora.
- Hadji, C. (1995). A avaliação dos professores. Linhas directivas para uma metodologia pertinente. In A. Estrela & P. Rodrigues (Coord.). *Para uma fundamentação da avaliação em educação*. (pp. 27-36). Lisboa: Edições Colibri.

- Harrison, P. D., Deanna, K. D. & Burdsal, C. A. (2004). The relative merits of different types of overall evaluations of teaching effectiveness. *Research in Higher Education*, 45 (3).
- Harvey, I. (2003). Student feedback. *Quality in Higher Education*, 9 (1), 3-20.
- Howard, G. S. & Maxwell, S. E. (1980). Correlation between student satisfaction and grades: A case of mistaken causation? *Journal of Educational Psychology*, 72, 810-820.
- Howard, G. S., Conway, C. G. & Maxwell, S. E. (1985). Construct validity of measures of college teaching effectiveness. *Journal of Educational Psychology*, 77, 187-196.
- Kaufman, D. V. (2002). Student's evaluation of faculty teaching. SFUFA Forum.
- Knight, P. T. (2002). *Being a teacher in higher education*. Buckingham: Society for Research into Higher Education & Open University Press.
- Koh, H. C & Tan, T. M. (1997). Empirical investigation of the factors affecting SET. *International Journal of Educational Management*, 11 (4), 170-178.
- Kulik, J. A. & Mc keachie, W. J. (1975). The evaluation of teachers in higher education. In F. N. Kerlinger (Ed.) *Review of Research in Education*. Itasca, Ill.: Peacock.
- Kulik, J. A. (2001). Student ratings: Validity, utility and controversy. *New Directions for Institutional Research*, 109 (1), 9-25.
- Lachapelle, A. (1998). Un processus d'évaluation de l'enseignement conçu et administré par les pairs. *Mesure et Évaluation en Éducation*, 21(2), 5-35.
- Langbein, L. L. (1994). *The validity of student evaluations of teaching*. *Politics Science & Politics*, pp.545-553.
- Liaw, S. H & Goh, K. L. (2003). Evidence and control of biases in student evaluations of teaching. *International Journal of Educational Management*, 17 (1), 37-43.
- Linn, R. L., Centra, J. A. & Tucker, L. (1975). Between, with and total group factor analysis of student ratings of instruction. *Multivariate Behavioral Research*, 10, 277-288.
- Lopes, A. (2002). Construção de identidades docentes e selves profissionais: um estudo sobre a mudança pessoal nos professores. *Revista de Educação*, 11 (2), 35-52.
- Lourtie, P. (1998). Avaliação da carreira docente. *Documenta*, 1, 38-46.
- Macedo, S. G. (2001). *Desempenho docente pela avaliação discente: Uma proposta metodológica para subsidiar a gestão universitária*. Dissertação de Doutorado. Florianópolis: Universidade Federal de Santa Catarina.
- Magalhães, F. (2004). Avaliação Psicológica de Docentes. *Encontro sobre a Avaliação Pedagógica no Ensino Superior*. Lisboa: Sindicato Nacional do Ensino Superior, 30-31 de Janeiro.

- Marsh, H. W. & Bailey, M. (1993). Multidimensional of student's evaluations of teaching effectiveness. A profile analysis. *Journal of Higher Education*, 64, 1-18.
- Marsh, H. W. & Dunkin, M. J. (1992). Students' evaluation of university teaching: A multidimensional perspective. *Higher education: Handbook of theory and research*, 8, (143-234). New York: Agathon.
- Marsh, H. W. & Hocevar, D. (1991a). The multidimensionality of student's evaluation of teaching effectiveness: The generality of factor structures across academic discipline, instructor level, and course level. *Teaching and Teacher Education*, 7, 9-18.
- Marsh, H. W. & Hocevar, D. (1991b). Student's evaluation of teaching effectiveness. The stability of mean ratings of the same teachers over a 13-year period. *Teaching and Teacher Education*, 7, 303-314.
- Marsh, H. W. & Overall, J. U. (1981). The relative influence of course level, course type, and instructor on students' evaluations of college teaching. *American Educational Research Journal*, 18, 103-112.
- Marsh, H. W. & Roche, L. A. (1992). The use of student's evaluation of university teaching in different settings: The applicability paradigm. *Australian Journal of Education*, 36, 278-300.
- Marsh, H. W. & Roche, L. A. (1993). The use of student's evaluation and an individually structured intervention to enhance university teaching effectiveness. *American Educational Research Journal*, 30, 217-251.
- Marsh, H. W. & Roche, L. A. (1994). *The use of student's evaluation of university teaching to improve teaching effectiveness*. Canberra, ACT: Australian Department of Employment, Education, and Training.
- Marsh, H. W. & Roche, L. A. (1997). Making student's evaluations of teaching effectiveness: The critical issues of validity, bias and utility. *American Psychologist*, 52, 1187-1197.
- Marsh, H. W. & Ware, J. E. (1980). Effects of expressiveness, content coverage, and incentive on multidimensional student rating scales: New interpretations of Doctor Fox Effect. *Journal of Educational Psychology*, 74, 126-134.
- Marsh, H. W. (1980). The influence of student, course, and instructor characteristics on evaluation of university teaching. *American Educational Research Psychology*, 72, 810-820.
- Marsh, H. W. (1982). Validity of student's evaluations of college teaching: A multitrait-multimethod analysis. *Journal of Educational Psychology*, 74, 264-279.
- Marsh, H. W. (1984). Student's evaluations of university teaching: Dimensionality, reliability, validity, potential biases and utility. *Journal of Educational Psychology*, 76, 707-754.

- Marsh, H. W. (1987). Students' evaluation of university teaching: research findings, methodological issues and directions for future research. *International Journal of Educational Research*, 11 (3), 253-388.
- Marsh, H. W. (1991). A multidimensional perspective student's evaluations of teaching effectiveness. A test of alternative higher-order structures. *Journal of Educational Psychology*, 83, 285-296.
- Marsh, H. W. (2001). *Student's evaluations of university teaching*. In Seminário realizado em 13 de Junho na Universidade do Minho. Conselho Académico.
- Mateo, M. A. & Fernández, J. (1996). Evaluation of the setting in which university faculty carry out their teaching and research functions: The ASEQ. *Educational and Psychology Measurement*, 55 (2).
- Mateo, M. A. & Sanchez, J. F. (1993). Dimensiones de la calidad de la enseñanza universitaria. *Psicothema*, 5 (2), 265-275.
- McKeachie, W. J. (1979). Student ratings of faculty: A reprise. *Academe*, 384-397.
- McKeachie, W. J. (1997). The validity of use. *American Psychologist*, 52 (11), 1218-1225.
- McMillan, J., Wergin, J. E., Forsyth, D. R. & Brown, J. C. (1986). Student ratings of instruction: A summary of literature. *Instructional Evaluation*, 9 (1), 2-9.
- ME/SEES (1997). *Autonomia e qualidade: Documento orientador das políticas para o ensino superior*. Lisboa. ME/SEES.
- Melo, M. C., Silva, J. L., Gomes, A. & Vieira, F. (2000). Concepções de Pedagogia Universitária: Uma análise do questionário de avaliação do ensino ministrado na Universidade do Minho. *Revista Portuguesa de Educação*, 13 (2), 125-156.
- Millman, J. (Ed.), (1981). *Handbook of teacher evaluation*. Beverly Hills, CA: SAGE
- Mintzberg, H. (1995). *Estrutura e Dinâmica das Organizações*. Lisboa: D. Quixote.
- Montenegro, M. I. (2002). A avaliação de projectos de ensino: A experiência da Universidade do Minho. In A. S. Pouzada, L. S. Almeida & R. M. Vasconcelos (Eds.). *Contextos e dinâmicas da vida académica*. Guimarães: Conselho Académico, Universidade do Minho.
- Murray, H. G. (1980). *Evaluating university teaching: A review of research*. Toronto, Canada: Ontario Confederation of University Faculty Associations.
- Murray, H. G. (1983). Low inference classroom teaching behaviours and student ratings of college teaching effectiveness. *Journal of Educational Psychology*, 71, 856-865.
- Murray, H. G. (1984). The impact of formative and summative evaluation of teaching in North American universities. *Assessment and Evaluation in Higher Education*, 9 (2), 117-132.
- OECD (2001). *Education at a glance*. Paris: OECD Publications Service.

- Ory, J. C. (2001). Faculty thoughts and concerns about student ratings. *New Directions for Teaching and Learning*, 87. John Wiley & Sons.Inc.
- Overall, J. U. & Marsh, H. W. (1980). Student's evaluations of instruction: A longitudinal study of their stability. *Journal of Educational Research*, 72, 321-325.
- Perales, M. J., Sánchez, P. & Chiva, I. (2002). El "Curso de Iniciación a la Docencia Universitaria" como experiencia de formación de profesores universitarios noveles en la Universitat de València. Un sistema de evaluación. *Revista Electrónica de Investigación y Evaluación Educativa*, 8(1).
- Perry, B. P. (1991). Quality in Higher Education. In T. Schuller (Ed.), *The future of Higher Education*. Buckingham: SRHE & Open University Press (91-99).
- Perry, B. P., Abrami, P. C., Dickens, W. J., & Leventhal, L. (1979). Instruction Reputation: An expectancy relationship involving students ratings and achievement. *Journal of Educational Psychology*, 71, 776-787.
- Pimenta, S. L. & Anastasio, L. G. (2002). *Docência no Ensino Superior*. São Paulo: Coleção Docência em Formação.
- Poissant, H. (1995). *L'évaluation de l'enseignement universitaire*. Montreal: Les Éditions Logiques.
- Poissant, H. (1996). *L'évaluation de l'enseignement Universitaire*. Québec: Les Éditions Logiques.
- Ramsden, P. (1991). A performance indicator of teaching quality in higher education: The course experience questionnaire. *Studies in Higher Education*, 16 (2), 33-52.
- Ramsden, P. (1998). *Learning to lead in higher education*. London: Routledge.
- Ramsden, P. (2003). *Learning to teach in higher education* (2ªEd.). London: Routledge & Falmer.
- Rego, A. & Sousa, L. (2000). Impactos dos comportamentos de cidadania docente sobre os alunos universitários. A perspectiva dos estudantes e dos professores. *Linhas Críticas*, 6 (10), 9-29.
- Rego, A. (2000). Eficácia comunicacional dos docentes universitários: Operacionalização de um constructo. *Psicologia, Educação e Cultura*, 4 (1), 127-140.
- Rego, A. (2001). Eficácia comunicacional na docência universitária: A perspectiva de estudantes e professores. *Psicologia: Teoria e Pesquisa*, 17 (3).
- Rego, A. (2003). *Comportamentos de cidadania docente: Na senda da qualidade no ensino superior*. Coimbra: Editora Quarteto.
- Rindermann, H. & Schofield, N. (2001). Generalizability of multidimensional student ratings of university instruction across courses and teachers. *Research in Higher Education*, 42 (4).
- Rosário, P. (2003). *Motivação e aprendizagem: Uma rota de leitura*. In Taveira, M. C. (no prelo). *Temas de Psicologia Escolar: Uma proposta científico-pedagógica*. Coimbra: Quarteto Editora.

- Rowley, J. (2003). Designing student feedback questionnaires. *Quality Assurance in Education*, 11 (3), 142-149.
- Sakai, J. M., Duarte, W. F. & Ukita, G. (2002). Características do professor eficiente sob a ótica de uma amostra de docentes universitários. *Boletim de Psicologia*, 53, 203-220.
- Santiago, R. A. (1997). O conceito de qualidade no ensino superior: Mitos e realidades. In Luís, Ana et al., (Ed.), *A administração da Educação: Investigação, formação e práticas*. Lisboa: FPAE.
- Santiago, R. A., Tavares, J., Taveira, M. C., Lencastre, L. & Gonçalves, F. (2001). *Promover o sucesso académico através da avaliação e intervenção na universidade*. Braga: Universidade do Minho.
- Santos, L. (2001). *Adaptação académica e rendimento escolar. Estudo com alunos universitários do 1º ano*. Braga: Universidade do Minho, Conselho Académico.
- Scriven, M. (1983). Summative teacher evaluation. In Millman (Ed.) *Handbook of Teacher Evaluation*. Thousand Oaks, Calif.:Sage.
- Seixas, A. M. (2003). *Políticas educativas e ensino superior em Portugal*. Coimbra: Quarteto.
- Seldin, P. & Associates. *How administrators can improve teaching*. San Francisco: Jossey-Bass.
- Seldin, P. (1984). *Changing practices in faculty evaluation: A critical assessment and recommendations for improvement*. San Francisco: Jossey-Bass.
- Seldin, P. (1989). Using student feedback to improve teaching. *New Directions for Teaching and Learning*, 37, 89-97.
- Seldin, P. (1993). *The teaching portfolio*. Boston: Anker Publishing.
- Seldin 2000**
- Simão, J. V. (2003a). *Por um modelo de avaliação contratualizante e responsabilizante*. In *Colóquio Percurso da Avaliação do Ensino Superior*. Lisboa: Conselho Nacional de Avaliação do Ensino Superior, 20 de Maio.
- Simão, J. V. (2003b). *Modernização do ensino superior: Da ruptura à excelência*. Fundação das Universidades Portuguesas. Conselho de Reitores das Universidades Portuguesas.
- Simão, J. V., Santos, S. M. & Costa, A. A. (2003). *Ensino Superior: Uma visão para a próxima década*. Lisboa: Gradiva.
- Simões, G. A. (2000). *A avaliação do desempenho docente: Contributos para uma análise crítica*. Lisboa: Texto Editora.
- Slotte, V. Wang, M. & Lonka, K. (2001). Information technology in medical education: A nationwide project on the opportunities of the new technology. *Medical Education*, 35, pp.990-995.

- Soares, A. P. & Almeida, L. S. (2001a). Transição para a universidade: Apresentação e validação do Questionário de Expectativas Académicas (QEA). In B. D. Silva & Almeida, L. S. (Orgs.), *Actas do VI Congresso Galaico-Português de Psicopedagogia*. Braga: Universidade do Minho.
- Soares, A. P. C. & Almeida, L. S. (2001b). Expectativas académicas e adaptação a universidade: Um estudo com alunos do 1º ano da Universidade do Minho. *Actas do V Seminário Investigação e Intervenção no Ensino Superior*. Viana do Castelo: Instituto Superior Politécnico.
- Soares, A. P. C. & Almeida, L. S. (2002). Ambiente académico e adaptação à Universidade: Contributos para validação do Classroom Environment Scale. In A. S. Pouzada, L. S. Almeida & R. M. Vasconcelos (Eds.). *Contextos e dinâmicas da vida académica*. Guimarães: Conselho Académico, Universidade do Minho.
- Soares, A. P. C. (2004). *Desenvolvimento de um modelo multidimensional de ajustamento de jovens universitários*. Dissertação de Doutoramento. Braga: Instituto de Educação e Psicologia, Universidade do Minho.
- Soares, A. P. C., Vasconcelos, R. M. & Almeida, L. S. (2002). Adaptação e satisfação na Universidade: Apresentação e validação do Questionário de Satisfação Académica (QSA). In A. S. Pouzada, L. S. Almeida & R. M. Vasconcelos (Eds.). *Contextos e dinâmicas da vida académica*. Guimarães: Conselho Académico, Universidade do Minho.
- Sobrinho, J. D. (2000). *Avaliação da Educação Superior*. Petrópolis, RJ: Vozes.
- Sorenson, D. L. & Reiner, C. (2003). Charting the uncharted seas of online student ratings of instruction. *New Directions for Teaching and Learning*, 96. John Wiley & Sons, Inc.
- Spencer, J. K. & Schmelkin, L. P. (2002). Student perspectives on teaching and its evaluation. *Assessment & Evaluation in Higher Education*, 27 (5).
- Tavares, J. (2000). Pedagogia universitária e inovação. In B. D. Silva & Almeida, L. S. (Orgs.), *Actas do VI Congresso Galaico-Português de Psicopedagogia*. Braga: Universidade do Minho.
- Teixeira, N. (2004). Modelos de avaliação do ensino e da docência. *Encontro sobre a Avaliação Pedagógica no Ensino Superior*. Lisboa: Sindicato Nacional do Ensino Superior, 30-31 de Janeiro.
- Terenzini, P. T. & Pascarella, E. T. (1994). Living with myths: Undergraduate education in America. *Change*, 26 (1), 28-32.
- Theall, M. & Franklin, J. L. (1990). Student ratings in the context of complex evaluation systems. In M. Theall & J. Franklin (Eds.) *Student ratings of instruction: Issues for improve practice*. *New Directions for Teaching and Learning*, 43. San Francisco: Jossey Bass.

- Theall, M. & Franklin, J. L. (2001). Looking for bias in all the wrong places: A search for truth or a witch hunt in student ratings of instruction. *New Directions for Teaching and Learning*, 109. San Francisco: Jossey Bass.
- Tremblay, M. C. (1998). Élaboration d'une grille d'auto-évaluation: Instrument de réflexion sur son enseignement. *Mesure et Évaluation en Éducation*, 21(2), 37-63.
- UNESCO, (1998). *Summary of the world declaration on higher education for the twenty-first century: Vision and action*. Paris: UNESCO.
- Van Vught, F. & Westerheijden, D. F. (1993). *Quality management and quality assurance in european higher education: Methods and mechanisms*. Luxembourg: Office of the Official Publications of the European Commission.
- Vieira, F., Gomes, A., Silva, J. L., Moreira, M. A, Melo, M. C. & Albuquerque, P. B. (2002). *Concepções de Pedagogia Universitária: Um estudo na Universidade do Minho*. Braga: Universidade do Minho.
- Vroeijenstijn, A. I. (1995). *Improvement and accountability: Navigating between scylla and charybdis. Guide for external quality assessment in higher education*. London: Jessica Kingsley.
- Watkins, D., Marsh, H. W. & Young, D. (1987). Evaluating tertiary teaching: A New Zealand perspective. *Teaching & Teacher Education*, 3, 41-53.
- Wilkerson, J. Manatt, P., Rogers, M. A. & Maughan, R. (2000). Validation of student, principal, and self-ratings in 360° feedback for teacher evaluation. *Journal of Personnel Evaluation in Education*, 14 (2), 179-192.
- Wilson, K. L., Lizzio, A. & Ramsden, P. (1997). The development, validation and application of the course experience questionnaire. *Studies in Higher Education*, 22 (1), 33-52.
- Wilson, R. C. (1986). Improving faculty teaching: Effective use of student evaluations and consultants. *Journal of Higher Education*, 57, 196-211.
- Zabalza, M. (1999). Coordenadas básicas para una docencia de calidad en la Universidad: Qué caracteriza la calidad de los docentes y la docencia universitaria. Conferência apresentada no I *Simposium Iberoamericano sobre Didáctica*. Santiago de Compostela: Universidade de Santiago de Compostela, 2-4 Dezembro.

ANEXOS

Anexo A – QAEM

Anexo B – PEA

Anexo C – Carta aos Docentes