

Dez Anos De Práticas De Avaliação Das Aprendizagens No Ensino Superior: Uma Síntese Da Literatura (2000-2009)*

Domingos Fernandes, Nuno Fialho
Instituto de Educação da Universidade de Lisboa

Resumo

O problema desta pesquisa consistiu em sintetizar literatura publicada num período de 10 anos (2000-2009), no domínio da avaliação das aprendizagens dos alunos, para identificar práticas associadas à melhoria do que, e como, se aprende. As sínteses de literatura constituem investigações cuja relevância é amplamente reconhecida, porque promovem a integração de trabalhos de natureza teórica e/ou de natureza empírica realizados por outros, resultando num conjunto de asserções que, em princípio, vão mais além do que a mera soma das partes analisadas. Tal como é mencionado por vários autores (e.g. Cooper, 2003; Slavin, 2004; Suri, 1999, 2002), as sínteses de literatura constituem investigações relevantes para conhecer o estado da pesquisa num dado domínio e para o “fazer avançar”. No espaço Iberoamericano e muito particularmente em Portugal, a investigação neste domínio é escassa e, por isso mesmo, pensou-se que este ponto de situação pode contribuir para informar e ajudar a desenvolver pesquisas futuras. Assim, o principal propósito desta investigação consistiu em produzir uma síntese de literatura que integrasse o que de mais essencial se pudesse afirmar acerca das práticas de avaliação das aprendizagens que se desenvolvem em contextos do ensino superior. Numa primeira fase foram identificados 63 artigos mas apenas 30, de natureza exclusivamente empírica, foram selecionados para efeitos desta investigação. A análise e a síntese da literatura revista permitiu, entre outras, elaborar as seguintes conclusões: a) as práticas de avaliação formativa, utilizando tarefas tais como relatórios, apresentações, posters e outros trabalhos escritos, estão associadas à melhoria das aprendizagens dos alunos; b) as práticas de auto e heteroavaliação contribuem de forma importante para envolver os alunos no processo de avaliação e para melhorar as suas aprendizagens; c) a avaliação de natureza formativa e criterial ajuda os alunos a compreender melhor o que são supostos aprender, orientando e regulando os seus esforços nesse sentido; e d) as práticas de avaliação formativa contribuem para desenvolver as capacidades mais complexas de pensamento dos alunos.

Palavras Chave: Avaliação das Aprendizagens; Práticas de Avaliação; Avaliação Formativa; Avaliação Sumativa; Pedagogia no Ensino Superior

* Este trabalho foi financiado por Fundos Nacionais através da FCT (Fundação para a Ciência e a Tecnologia) no âmbito do projecto PTDC/CPE-CED/114318/2009.

1. Introdução, problema e questões da investigação

A relação entre a natureza das práticas de avaliação dos professores nas salas de aula e as aprendizagens desenvolvidas pelos seus alunos está bem estabelecida na literatura. Em particular, é geralmente reconhecido que as práticas de avaliação formativa melhoram muito significativamente as aprendizagens de todos os alunos e, em particular, daqueles que têm mais dificuldades. A avaliação formativa, avaliação para as aprendizagens, está hoje indissociavelmente associada a mais e melhores aprendizagens e a processos tais como o feedback, a auto-avaliação e a participação dos alunos no desenvolvimento das atividades letivas (e.g. Black & Wiliam, 1998, 2006; Gardner, 2006; Stiggins & Chappuis, 2005).

Apesar destes trabalhos não terem sido desenvolvidos em contextos do ensino superior, onde a investigação daquelas questões é menos abundante, vários autores têm chamado a atenção para a necessidade de, também neste nível de ensino, se desenvolver uma avaliação mais orientada para apoiar os processos de ensino e de aprendizagem (e.g. Leathwood, 2005; Wilson & Scalise, 2006; Gijbels & Struyf, 2008; Garcia, 2009). No entanto, ainda que seja reconhecida a necessidade de se utilizarem práticas de ensino e de avaliação mais sofisticadas, a verdade é que a investigação vem sugerindo que tais práticas estão aquém do que a literatura recomenda, baseando-se quase exclusivamente na transmissão de conhecimentos que são avaliados através de testes no final de um maior ou menor período de tempo (Bryan & Clegg, 2006).

Yorke (2003) e Biggs (2006) referiram algumas das razões que, na sua opinião, ajudam a compreender a ausência quase generalizada da avaliação formativa nas salas de aula do ensino superior tais como: a) a pressão para a obtenção de determinados resultados através da definição de standards que acabam por determinar a utilização de avaliações de natureza sumativa; b) a organização curricular através de conteúdos modulares, reduz a oportunidade para distribuir feedback pelos alunos; c) o facto de, na prática, a produção de investigação ser considerada prioritária em relação ao ensino para efeitos da progressão na carreira dos professores universitários; d) a redução do investimento público no ensino superior tem-se traduzido num maior número de alunos por turma e na conseqüente redução do número de tarefas de ensino e avaliação realizadas pelos alunos, do feedback proporcionado aos alunos e numa maior utilização da avaliação de natureza sumativa.

Para Boud (2000) os esforços de ensino e de aprendizagem de professores e alunos têm que ser sistematicamente apoiados por práticas de avaliação que não sejam exclusivas dos professores, permitindo a intervenção ativa dos alunos. Na mesma linha, Taras (2002) referiu que a participação dos alunos nos processos de avaliação era essencial para o desenvolvimento da sua autonomia, sendo necessário introduzir mudanças nas práticas de avaliação que vão para além do feedback providenciando pelos professores. Para esta autora era igualmente necessário envolver os alunos no processo de avaliação, introduzir novas e inovadoras formas de auto e de heteroavaliação e

desenvolver tarefas de avaliação formativa e sumativa mais articuladas entre si e com o desenvolvimento do currículo. Também Leathwood (2005) sublinhou a necessidade de se explorarem novas formas e estilos de avaliação, tendo em conta estratégias que considerem o aluno no centro do processo de aprendizagem, recorrendo à auto avaliação, à avaliação por pares e à construção de portefólios e a uma maior e ampla discussão dos resultados de aprendizagem conseguidos pelos alunos.

Segundo Leach, Neutze & Zepke (2001) o desafio ao nível das práticas de avaliação no ensino superior reside na criação de condições para que todos os alunos se sintam como parte do processo e, como tal, mais comprometidos e motivados na e para a sua aprendizagem.

Nestas condições, o principal propósito desta investigação era produzir uma síntese de literatura que integrasse o que de mais essencial se pudesse afirmar acerca das práticas de avaliação descritas na literatura selecionada. Mereceram particular atenção as tarefas utilizadas, a participação dos alunos, a natureza e frequência do feedback, a natureza e as utilizações da avaliação e questões associadas à introdução de inovações nas práticas de avaliação. Assim, a investigação foi orientada pelas questões que a seguir se formulam.

1. Como se poderão caracterizar as tarefas de avaliação utilizadas e a participação dos alunos no processo do seu desenvolvimento?
2. O que se poderá dizer acerca da natureza do feedback utilizado e da frequência da sua distribuição?
3. Como se poderão descrever a natureza da avaliação e as utilizações que dela faziam os participantes nas investigações revistas?
4. Que reflexões se poderão inferir acerca das questões associadas à introdução de novas e inovadoras formas de avaliar no ensino superior?

2. Método e procedimentos

A elaboração da presente síntese da literatura, que é uma sistematização de um conjunto selecionado de investigações produzidas por outros ao longo de 10 anos, teve em conta contribuições teóricas de uma diversidade de autores (e.g. Major & Savin-Baden, 2010; Suri, 2002, Suri & Clarke, 1999, Weed, 2008). As sínteses de literatura constituem um fim em si mesmo, isto é, ajudam-nos a compreender um problema, neste caso as práticas de avaliação em contextos de ensino superior. Por isso, interessa distingui-las das clássicas revisões de literatura cujo principal propósito é, em regra, enquadrar conceptualmente uma investigação primária. Assim, esta investigação desenvolveu-se de acordo com as seguintes fases.

1. Identificação da área de investigação e definição do propósito da síntese de literatura.
2. Seleção de fontes e critérios para a identificação e recolha de literatura.

3. Identificação, recolha e análise temática e contextual dos artigos.
4. Definição do assunto chave, objecto da síntese de literatura.
5. Definição de questões orientadoras da investigação e primeiro processo iterativo.
6. Categorização dos relatórios de análise da literatura selecionada.
7. Processo de síntese (resultante da análise por categoria e por documento dentro de cada categoria) e segundo processo iterativo.

Seguindo as recomendações de Suri (1999, 2002) definiu-se um conjunto de literatura (e.g. revistas, livros) que garantisse a validade, diversidade e ecletismo dos dados a utilizar. Para tal foram considerados os seguintes critérios fundamentais: a) as revistas a seleccionar deveriam ser indexadas, preferencialmente orientadas para problemáticas do ensino superior e com a prática de revisão anónima por pares; b) a literatura tinha que ter como objeto central de investigação e/ou de reflexão a avaliação das aprendizagens dos alunos no ensino superior e abranger um período de 10 anos (2000-2009); e c) a literatura a seleccionar deveria ter em conta diferentes tradições científicas e culturais, nomeadamente a anglo saxónica, a francófona e a iberoamericana. Assim, foram selecionadas 29 revistas e periódicos, maioritariamente anglo saxónicos, incluindo uma revista portuguesa e cinco brasileiras. Foram ainda selecionadas uma revista de uma organização iberoamericana, as publicações de uma associação francófona de investigadores no domínio da avaliação e dois livros que abordavam questões relacionadas com a avaliação, o ensino e a aprendizagem no ensino superior (Bryan & Clegg, 2006; Falchikov, 2005). A maioria dos materiais foram identificados através da base de dados EBSCO.

A pesquisa inicial resultou na seleção de 63 artigos de natureza empírica e de natureza teórica e reflexiva que foram analisados por dois investigadores que, após uma primeira fase de reuniões em que acertaram os elementos a registar para cada artigo (e.g. referência, título, problema, metodologia), produziram as respetivas fichas de leitura. Numa segunda fase de reuniões analisaram-se em conjunto as fichas produzidas procedendo-se aos ajustamentos considerados necessários.

No processo de identificação dos temas chave que estiveram na génese dos artigos identificados fez-se uma análise contextual e temática das fichas de leitura realizadas (Weed, 2008). Verificou-se que um conjunto significativo de artigos discutia investigações que tinham resultado em intervenções deliberadas na utilização de práticas de avaliação que, supostamente, melhoravam as aprendizagens dos alunos. Ou seja, práticas de avaliação formativa ou de avaliação para as aprendizagens.

Em seguida, foram definidas as questões de investigação acima formuladas e realizou-se o primeiro processo iterativo de revisão dos 63 artigos inicialmente selecionados. Nesta fase foi decidido excluir todos os artigos que não fossem de natureza empírica para que a investigação se pudesse centrar mais nas práticas reais de avaliação. Neste primeiro processo iterativo houve oportunidade para rever artigos que eventualmente tivessem ficado de parte e que, nesta fase da investigação, pudessem ser considerados relevantes.

Após os procedimentos referidos, foram selecionados para análise 30 dos 63 artigos inicialmente sinalizados tendo-se verificado que em 17 foi utilizada uma metodologia de natureza qualitativa, em sete uma metodologia mista e em seis uma metodologia quantitativa. Dezasseis artigos relatavam investigações em que os participantes eram alunos e professores enquanto nos restantes 14 apenas participaram alunos. A maioria das investigações (17) ocorreu em contextos de unidades curriculares da área das Ciências Sociais, seis na área das Ciências e das Tecnologias, duas na área das Ciências da Saúde e uma na área das Artes e Humanidades. Uma investigação envolveu as áreas das Ciências Sociais e das Artes e Humanidades e noutras três não era feita qualquer referência à área do conhecimento.

A análise contextual dos artigos selecionados permitiu considerar seis categorias de análise:

1. Tarefas de avaliação utilizadas.
2. Participação dos alunos.
3. Natureza e frequência do feedback.
4. Natureza da avaliação.
5. Utilização da avaliação e dos seus resultados.
6. Questões associadas à introdução de inovações nas práticas de avaliação.

Estas categorias permitiram realizar a análise vertical (por artigo ao longo das categorias) e a análise horizontal (por categoria ao longo dos artigos). Estas análises permitiram elaborar sínteses a partir das quais se produziram os resultados que se discutem na secção seguinte.

3. Apresentação e discussão dos resultados

Nesta secção apresenta-se uma breve súmula dos resultados obtidos através da análise dos artigos selecionados tendo em conta as categorias que foi decidido considerar na investigação.

Tarefas de Avaliação Utilizadas e Participação dos Alunos

Na maioria das investigações foram utilizadas uma variedade de tarefas de avaliação, em geral de forma sistemática, tais como portefólios (Albertino & De Sousa, 2004; Carless, 2002; Klenowski *et al.*, 2006; Moti & Barzilai, 2004); posters (Ballantyne *et al.*, 2002; Bloxham & West, 2004; Orsmond *et al.*, 2000; Orsmond *et al.*, 2002); relatórios ou trabalhos escritos (Berg *et al.*, 2006; Gibels *et al.*, 2008; Moti & Barzilai, 2004); questionários ou testes intermédios (Cook, 2001; Nicol, 2009; Tess, 2009; Trotter, 2006; Walker & Warhurst, 2000; Walker & Warhurst, 2000); e apresentações orais de trabalhos (Carless, 2007; Fallows & Chandramohan, 2001; Gibels *et al.*, 2008; Moti & Barzilai, 2004; Poon *et al.*, 2009). Em diversas situações foram utilizadas duas ou mais tarefas de avaliação preferenciais tais como apresentação de projeto e construção de poster (Ballantyne *et al.*, 2002),

relatórios, poster, apresentação, portfólio (Moti & Barzilai, 2004), apresentação e construção de modelos (Gibels *et al.*, 2008) e questionários e trabalho escrito de desenvolvimento de um tema (Nicol, 2009). Durante a realização das tarefas propostas os alunos eram, num número assinalável de casos, apoiados através de tutorias (Albertino & De Sousa, 2004; Bloxham & West, 2004; Giles *et al.*, 2004; Johnston & Miles, 2004; Klenowski *et al.*, 2006; Macmillan & Mclean 2005; Moti & Barzilai, 2004; Nicol, 2009; Trotter, 2006).

A participação dos alunos nos processos de avaliação desenvolveu-se, no essencial, através de três tipos de dinâmicas: a) organização dos alunos em grupos de trabalho para a realização e análise das tarefas propostas (Albertino, & De Sousa, 2004; Fallows & Chandramohan, 2001; Gale *et al.*, 2002; Gibels *et al.*, 2008; Johnston & Miles, 2004; Moti & Barzilai, 2004; Nicol, 2009; Orsmond *et al.*, 2000; Orsmond *et al.*, 2002; Poon *et al.*, 2009; Walker, & Warhurst, 2000); b) participação dos alunos no processo de definição de critérios de avaliação que eram tidos em conta na apreciação do trabalho realizado (Ballantyne *et al.*, 2002; Dancer & Kamvounias, 2005; Gale *et al.*, 2002; Moti & Barzilai, 2004; Orsmond *et al.*, 2002; Poon *et al.*, 2009); e c) utilização de tarefas de auto e heteroavaliação por parte dos alunos (Ballantyne *et al.*, 2002; Bloxham & West, 2004; Dancer & Kamvounias, 2005; Fallows & Chandramohan, 2001; Gale *et al.*, 2002; Johnston & Miles, 2004; Klenowski *et al.*, 2006; Miller *et al.*, 2009; Orsmond *et al.*, 2000; Orsmond *et al.*, 2002; Price & O'Donovan, 2006; Taras, 2002; Walker & Warhurst, 2000; Wanous *et al.*, 2009).

Em geral, os alunos eram orientados e preparados para desenvolverem tarefas de auto e de heteroavaliação, nomeadamente através da análise e discussão de conceitos avaliativos e da distribuição de feedback e da utilização dos critérios de avaliação (Ballantyne, *et al.*, 2002; Bloxham & West, 2004; Fallows & Balasubramanyan, 2001; Orsmond *et al.*, 2000; Orsmond *et al.*, 2002; Sluijsmans *et al.*, 2002; Taras, 2002;). Em várias situações os alunos praticavam utilizando trabalhos anteriormente realizados (Ballantyne *et al.*, 2002; Fallows & Balasubramanyan, 2001; Orsmond *et al.*, 2002).

Natureza e Frequência do Feedback

Para analisar o feedback descrito na literatura revista utilizou-se a caracterização desenvolvida por Vieira (2007).

Na grande maioria das investigações analisadas os processos de avaliação eram de referência criterial e, por isso, o feedback era normalmente orientado pelos respetivos critérios. Em cerca de metade dos estudos, sobretudo naqueles em que os alunos resolviam individualmente as tarefas propostas (e.g. relatórios questionários), o feedback foi distribuído apenas pelos docentes (e.g. Andrade & Du, 2007; Carless, 2002; Cook, 2001; Gilles *et al.*, 2004; Greer, 2001; Macmillan & Mclean, 2005; Miller, 2009; Mok *et al.*, 2006; Price & O'Donovan, 2006). Na outra metade, nos casos em que as tarefas estavam

associadas à utilização de processos de heteroavaliação, os alunos também distribuía feedback (e.g. Berg *et al.*, 2006; Gijbels *et al.*, 2008; Macmillan & Mclean, 2005; Walker & Warhurst, 2000).

Na maioria dos artigos revistos o processo de avaliação está organizado de forma a proporcionar feedback de imediato ou pouco tempo após a concretização da tarefa solicitada. A frequência do feedback parece estar associada à natureza das tarefas que são solicitadas aos alunos, sendo mais frequente ao longo da construção de portefólios ou do desenvolvimento de projetos, uma vez que alunos e professores trabalham mais vezes em conjunto.

Na maioria dos casos os professores preferiram distribuir feedback sob a forma escrita ou verbal que, em 18 dos 30 artigos analisados, incidiu sobre realizações concretas dos alunos na sequência do que lhes era proposto pelos respetivos professores (e.g. Ballantyne *et al.*, 2002; Johnston & Miles, 2004; Poon *et al.*, 2009). Nos restantes casos, o feedback incidiu igualmente sobre aspetos atitudinais e comportamentais e processos utilizados pelos alunos. O feedback providenciado pelos professores estava essencialmente associado às funções de orientar, regular e classificar o desempenho dos alunos. Frequentemente, as funções de orientar e regular surgiram juntamente com as funções de corrigir, informar e motivar (e.g. Carless, 2002; Macmillan & Mclean, 2005; Trotter, 2006). Nestes casos, o feedback tinha a configuração de um comentário e envolvia o desenvolvimento de processos que permitiam que os alunos tivessem acesso a informação detalhada sobre o seu desempenho na tarefa solicitada (e.g. apresentações, debates, relatórios, questionários). O feedback mais orientado para as classificações surgiu frequentemente associado a outras funções tais como as de corrigir, orientar, regular e informar (e.g. Greer, 2001; Nicol, 2009; Poon *et al.*, 2009). Nestes casos, o feedback podia surgir sob a forma de uma combinação entre a atribuição de uma classificação e a elaboração de um comentário, permitindo deste modo que os alunos tivessem acesso a informações mais detalhadas acerca do seu desempenho, como acontecia nos casos anteriormente referidos.

O feedback distribuído pelos alunos aos seus colegas foi geralmente estruturado e escrito, uma vez que era apoiado em instrumentos tais como tabelas, *checklists* e formulários. Na maior parte dos casos analisados o feedback incidia sobre os produtos realizados. Porém, sempre que se pedia aos alunos que se pronunciassem sobre o contributo dos seus pares para a realização dos trabalhos de grupo, o feedback incidia igualmente sobre aspetos atitudinais, comportamentais e processuais (e.g. Orsmond *et al.*, 2002; Walker & Warhurst, 2000; Wanous *et al.*, 2009).

O feedback entre pares tinha, por norma, as funções de corrigir, regular e orientar o desenvolvimento das aprendizagens. No entanto, nalguns casos, tinha igualmente a função de contribuir para a atribuição das classificações finais (Bloxham & West, 2004; Price & O'Donovan, 2006; Wanous *et al.*, 2009).

Natureza e Utilização da Avaliação

A literatura revista analisou experiências inovadoras no domínio da avaliação formativa, destinadas a melhorar as aprendizagens dos alunos através da sua participação ativa nos processos de aprendizagem e de avaliação e de uma criteriosa seleção das tarefas utilizadas. No entanto, em vários casos, foi possível constatar a utilização de procedimentos de avaliação sumativa complementares aos que são utilizados na avaliação formativa.

De modo geral, a avaliação utilizada nas investigações analisadas era de natureza criterial e tinha em vista a regulação e a auto-regulação das aprendizagens dos alunos. A ideia mais presente na literatura analisada era a de que a avaliação deveria contribuir para que os alunos compreendessem melhor o que eram supostos aprender e como o deveriam fazer, para que pudessem orientar o seu trabalho. Assim, eram selecionadas tarefas tais como a elaboração de portefólios, o desenvolvimento de projetos, a conceção e construção de posters e uma diversidade de trabalhos em grupo, em que o feedback e acompanhamento por parte dos professores se revelavam elementos fundamentais para que os alunos pudessem orientar, tanto quanto possível de forma autónoma, os seus esforços de aprendizagem.

A auto-avaliação foi utilizada numa diversidade de investigações com o propósito de induzir nos alunos práticas de reflexão acerca do seu desempenho (e.g. Dancer & Kamvounias, 2005; Klenowski *et al.*, 2006; Mok *et al.*, 2006). Este processo decorria tipicamente no final de uma unidade de ensino com os alunos a analisar o seu trabalho a partir de critérios de avaliação ou de modelos existentes e tendo em conta o feedback fornecido pelos professores ou pelos seus pares (e.g. Andrade, & Du, 2007; Nicol, 2009; Poon *et al.*, 2009; Taras, 2002). Desta forma, os alunos eram sistematicamente envolvidos numa diversidade de ações que, supostamente, contribuía para o desenvolvimento das suas capacidades metacognitivas, levando-os a terem maior consciência acerca do modo como o processo de avaliação se realizava e acerca dos processos que utilizavam para aprender através das tarefas que lhes eram propostas.

A heteroavaliação permitia melhorar as capacidades dos alunos no domínio da formulação de juízos acerca do valor e do mérito de um dado objeto. O desenvolvimento da capacidade para avaliar o trabalho de outros permitia que os próprios alunos regulassem o seu próprio processo de aprendizagem. Além disso, a possibilidade de analisar, discutir e avaliar o desempenho dos seus pares contribuía igualmente para melhorar a motivação dos alunos, ajudando-os a concentrar-se nos processos de aprendizagem (e.g. Ballantyne *et al.*, 2002; Gijbels *et al.*, 2008; Macmillan & Mclean, 2005).

Numa parte considerável dos artigos analisados verificou-se que as tarefas de avaliação eram utilizadas para fins formativos e sumativos, contribuindo para determinar as classificações dos alunos. No entanto, a ponderação das tarefas do dia a dia era sempre inferior à das tarefas de avaliação final tais como o exame, o projeto e o ensaio (e.g. Cook, 2001; Gale *et al.*, 2002; Nicol, 2009; Taras, 2002).

Em geral, pode dizer-se que a avaliação era utilizada com dois propósitos fundamentais: a) regular e orientar as aprendizagens dos alunos no contexto das tarefas que lhes eram propostas; e b) melhorar as competências metacognitivas dos alunos, tornando-os mais conscientes acerca dos seus próprios processos de aprendizagem.

Introdução de inovações nas práticas de avaliação

De acordo com a grande maioria dos alunos e professores intervenientes nos estudos analisados, a introdução de novas e inovadoras práticas de avaliação e/ou a redefinição das já existentes, teve resultados positivos quer ao nível dos processos de ensino quer ao nível dos processos de aprendizagem.

Relativamente à utilização de tutorias, os professores referiram que esta opção os ajudava a identificar problemas de aproveitamento dos alunos e a ajustar o processo de ensino (Moti & Barzilai, 2004). Além disso, as tutorias ajudavam a desenvolver capacidades complexas de pensamento, mais consentâneas com as exigências do ensino superior (Albertino & De Sousa, 2004). A perspectiva dos alunos sobre as tutorias também foi muito positiva, considerando que o frequente feedback dos professores era fundamental para conseguirem compreender melhor o que era necessário aprender (e.g. Albertino & De Sousa, 2004; Carless, 2002; Klenowski *et al.*, 2006; Macmillan & Mclean, 2005). Outras opiniões, no entanto, referiam que as tutorias impunham uma sobrecarga de trabalho para os professores (Macmillan & Mclean, 2005; Trotter, 2006).

Os estudos analisados permitiram constatar que o conhecimento e a compreensão dos critérios de avaliação estava associado à melhoria das aprendizagens por parte dos alunos (e.g. Johnston & Miles, 2004). Os processos de análise e discussão de aplicação de critérios de avaliação revelaram-se fundamentais para que os alunos adquirissem conhecimentos necessários para a sua correta e eficaz utilização (e.g. Price & O'Donovan, 2006). Além disso, avaliar tarefas resolvidas anteriormente utilizando critérios ajudou os alunos a distribuir feedback de natureza mais construtiva (Orsmond *et al.*, 2002).

As práticas de auto e heteroavaliação ajudaram os alunos a focarem-se no que era essencial aprender, a estarem mais atentos nas aulas e a motivarem-se para o trabalho escolar (e.g. Berg *et al.*, 2006; Bloxham & West, 2004; Orsmond *et al.*, 2000). Em particular, a reflexão dos alunos sobre o seu trabalho escolar através da auto-avaliação referida a critérios, parece ter contribuído para melhorar as suas capacidades metacognitivas e de auto-regulação (e.g. Klenowski *et al.*, 2006; Mok *et al.*, 2006). Os resultados de vários estudos mostraram ser necessário desenvolver práticas sistemáticas de auto e heteroavaliação para que os alunos se familiarizem com processos para os quais não estão preparados (e.g. Orsmond *et al.*, 2000; Price & O'Donovan, 2006). Interessa ainda referir que, em geral, os alunos

revelaram maior imparcialidade e objetividade nas práticas de heteroavaliação do que nas práticas de auto-avaliação (e.g. Dancer & Kamvounias, 2005; Walker & Warhurst, 2000).

De modo geral, os alunos também avaliaram muito positivamente as práticas de auto e/ou heteroavaliação considerando-as relevantes para as suas aprendizagens pois permitiam identificar o que se pretendia e analisar os trabalhos que produziam (e.g. Fallows & Balasubramanyan, 2001; Moti & Barzilai, 2004; Taras, 2002). No entanto, em algumas investigações os alunos referiram que não se sentiam confortáveis perante os restantes colegas com as práticas de heteroavaliação, chegando a considerar que não tinham competência para avaliar os seus pares nem a si mesmos (e.g. Andrade & Du, 2007; Ballantyne *et al.*, 2002; Bloxham & West, 2004; Taras, 2002). Na opinião desses alunos a responsabilidade de avaliar devia ser exclusiva dos professores.

A avaliação por pares revelou-se uma estratégia eficaz no desenvolvimento das aprendizagens do alunos, contribuindo igualmente para avaliar a sua participação individual nos trabalhos de grupo (e.g. Ballantyne *et al.*, 2002; Berg *et al.*, 2006; Bloxham & West, 2004; Johnston & Miles, 2004).

4. Síntese e reflexão final

Na maioria das investigações analisadas foram utilizadas uma diversidade de tarefas, muitas vezes mais do que uma ao longo de um dado período, que se revelaram adequadas para a aprendizagem dos alunos (e.g. posters, relatórios, trabalhos escritos, relatórios) e que mostraram que é possível ensinar, aprender e avaliar doutras formas, envolvendo ativamente os estudantes nestes processos. A participação dos alunos no processo de avaliação concretizou-se através da análise de tarefas em dinâmicas de grupo, da definição de critérios e das práticas de auto e heteroavaliação.

O feedback era baseado em critérios e distribuído oralmente ou por escrito, incidindo quer sobre resultados quer sobre processos utilizados. Assim, o feedback tanto era utilizado para orientar e regular as aprendizagens como para ajuizar acerca do valor e do mérito de um determinado produto ou resultado. Os alunos forneciam feedback sobretudo no contexto das atividades de heteroavaliação mas também quando as tarefas propostas implicavam mais trabalho conjunto com os professores.

Na grande maioria dos estudos analisados a avaliação era de natureza formativa e criterial e essencialmente orientada para melhorar, regular e orientar o desenvolvimento das aprendizagens. Porém, numa variedade de casos, a avaliação sumativa era utilizada como complementar, associada aos pontos de situação que se faziam acerca do que tinha sido aprendido no final de um certo período de tempo. A avaliação, tal como descrita na maioria dos estudos, parece ter tido um papel determinante na tomada de consciência por parte dos alunos acerca do que tinham realmente que aprender e dos esforços que necessitavam de fazer para o concretizar. Além disso, a avaliação era utilizada para que os alunos pudessem refletir sobre o que tinham aprendido e sobre as melhores

formas de se organizarem para aprender. Por isso, as práticas de auto e de heteroavaliação se revelaram tão relevantes no desenvolvimento de um importante conjunto de capacidades, nomeadamente as de natureza metacognitiva.

Esta investigação permitiu constatar que é possível melhorar as práticas de avaliação e de ensino no contexto do ensino superior e, conseqüentemente, melhorar substancialmente o que, e como, os alunos aprendem. Os resultados das investigações revistas foram geralmente muito positivos e encorajadores, sobretudo porque mostraram que a participação ativa dos alunos nos processos de avaliação e de aprendizagem conduz invariavelmente ao desenvolvimento das capacidades mais complexas de pensamento. As práticas sistemáticas de auto e de heteroavaliação revelaram-se fundamentais porque permitiram que, em cada momento, os alunos compreendessem em que situação se encontravam e o que, e como, tinham que aprender.

No entanto, tal como afirmam alguns investigadores, os métodos e processos de investigação necessitam de ser mais elaborados para que os resultados que apontam para as melhorias das aprendizagens dos alunos possam ter outra credibilidade e consistência (Greer, 2001; Miller, 2009; Nicol, 2009). A questão da utilização do feedback e dos resultados das avaliações em geral deverá ser objeto de mais atenção para que os alunos possam usufruir dessa informação para melhorar as suas aprendizagens (Orsmond *et al.*, 2002).

Referências Bibliográficas

- Albertino, F. & De Sousa, N. (2004). Avaliação da aprendizagem: o portefólio como auxiliar na construção de um profissional reflexivo. *Estudos em Avaliação Educacional*, 29 (jan-jun), 169-189.
- Andrade, Heidi & Du, Ying (2007). Student responses to criteria-referenced self-assessment. *Assessment and Evaluation in Higher Education*. 31 (3), 341-356.
- Ballantyne, Roy, Hughes, Karen & Mylonas, Aliisa (2002). Developing procedures for implementing peer assessment in large classes using an action research process. *Assessment and Evaluation in Higher Education*, 27 (5), 427-441.
- Berg, Ink, Admiraal, Wilfried & Pilot, Albert (2006). Design principles and outcomes of peer assessment in higher education. *Studies in Higher Education*, 31 (3), 341-356.
- Biggs, Graham (2006). How assessment frames student learning. In C. Bryan & K. Clegg (Orgs.), *Innovative Assessment in Higher Education* (pp. 23-36). New York: Taylor and Francis Group.
- Black, Paul & Wiliam, Dylan (1998). Assessment and classroom learning. *Assessment in Education: Principles, Policy & Practice*, 5(1), 7-74.

- Black, Paul & Wiliam, Dylan (2006). Assessment for learning in the classroom. In John Gardner (Ed.), *Assessment and learning* (pp. 9-25). London: Sage.
- Bloxham, Sue & West, Amanda (2004). Understanding the rules of the game: marking peer assessment as a medium for developing students' conceptions of assessment. *Assessment and Evaluation in Higher Education*, 29 (6), 721-732.
- Bryan, Cordelia, & Clegg, Karen (2006). Introduction. In C. Bryan & K. Clegg (Orgs.), *Innovative Assessment in Higher Education* (pp. 3-10) New York: Taylor and Francis Group.
- Boud, David (2000). Sustainable assessment: Rethinking assessment for the learning society. *Studies in Continuing Education* 22(2), 151-167.
- Carless, David R. (2002). The 'Mini-Viva' as a tool to enhance assessment for learning. *Assessment and Evaluation in Higher Education*, 27 (4), 354-363.
- Cook, Averil (2001). Assessing the use of flexible assessment. *Assessment and Evaluation in Higher Education*, 26 (6), 539-549.
- Cooper, H. (2003). Editorial. *Psychological Bulletin*, 129(1), pp. 3-9.
- Dancer, Diane, & Kamvounias, Patty (2005). Student involvement in assessment: A project designed to assess class participation fairly and reliably. *Assessment and Evaluation in Higher Education*, 30 (4), 445-454.
- Fallows, Stephen & Balasubramanyan, Chandramohan (2001). Multiple approaches to assessment: Reflections on use of tutor, peer and self-assessment. *Teaching in Higher Education*, 6 (2), 229-246.
- Falchikov, Nancy (2005). *Improving assessment through student involvement: Practical solutions for aiding learning in higher and further education*. New York: Routledge Farmer.
- Gale, Ken, Martin, Ken & Mcqueen, Gillian (2002). Triadic assessment. *Assessment and Evaluation in Higher Education*, 27 (6), 557-567.
- Garcia, Joe (2009). Avaliação e aprendizagem na educação superior. *Estudos em Avaliação Educacional*. 20, (43) 201-213.
- Gardner, John (2006). Assessment and learning: An introduction. In John Gardner (Ed.), *Assessment and learning* (pp. 1-5). London: Sage.
- Gijbels, David, Segers, Mien & Struyful, Elke (2008). Constructivist learning environments and the (im)possibility to change students' perceptions of assessment demands and approaches to learning. *Instructional Science*, 36, 431-443.
- Gilles, Anna, Martin, Sylvia, Bryce, Deborah & Hendry, Graham (2004). Students as partners in evaluation: student and teacher perspectives. *Assessment and Evaluation in Higher Education*, 29 (6), 681-685.
- Johnston, Lucy & Miles, Lucy (2004). Assessing contributions to group assignments. *Assessment and Evaluation in Higher Education*, 29 (6), 751-768.

- Klenowski, Val, Askew, Sue & Carnell, Eileen (2006). Portfolios for learning, assessment and professional development in higher education. *Assessment and Evaluation in Higher Education*, 31 (3), 267-286.
- Leathwood, Carole (2005). Assessment policy and practice in higher education: Purpose, standards and equity. *Assessment and Evaluation in Higher Education*, 30 (3), 307-324.
- Leach, L., Neutze, G. & Zepke, N. (2000) Learners' perceptions of assessment: Tensions between philosophy and practice. *Studies in the Education of Adults*, 32(1), 107-119.
- Macmillan, John & Mclean, Monica (2005). Making first-year tutorials count: Operationalizing the assessment-learning connection. *Active Learning in Higher Education*. 6 (2), 94-105.
- Major, Clair & Savin-Baden, Maggi (2010). Exploring the relevance of qualitative research synthesis to higher education research and practice. *London Review of Education*, 8, (2), 127-140.
- Miller, Tess (2009). Formative computer-based assessment in higher education: The effectiveness of feedback in supporting student learning. *Assessment and Evaluation in Higher Education*, 34 (2), 181-192.
- Mok, Madalena, Lung, Ching, Cheng, Doris, Cheung, Rebecca & Ng, Lee (2006). Self-assessment in higher education: Experience in using a metacognitive approach in five case studies. *Assessment and Evaluation in Higher Education*, 31 (4), 415-433
- Moti, Frank & Barzilai, Abigail (2004). Integrating alternative assessment in a project-based learning course for pre-service science and technology teachers. *Assessment and Evaluation in Higher Education*, 29 (1), 41-61.
- Nicol, David (2009). Assessment for learner self evaluation: Enhancing achievement in first year using learning technologies. *Assessment and Evaluation in Higher Education*. 34 (3), 335-352.
- Orsmond, Paul, Merry, Stephen & Reiling, Kevin (2000). The use of student derived marking criteria in peer and self-assessment. *Assessment and Evaluation in Higher Education*, 1 (1), 23-39.
- Orsmond, Paul, Merry, Stephen & Reiling, Kevin (2002). The use of exemplars and formative feedback when using student derived marking criteria in peer and self-assessment. *Assessment and Evaluation in Higher Education*, 27 (4), 309-323.
- Poon, Wai-Yin, McNaughton, Carmel, Lamb, Paul & Kwan, H. S. (2009). Improving assessment methods in university science education with negotiated self- and peer-assessment. *Assessment and Evaluation in Higher Education*, 16 (3), 331-346.
- Price, Margaret & O'Donovan, Berry (2006). Improving performance through enhancing student understanding of criteria and feedback. In C. Bryan & K. Clegg (Eds.), *Innovative Assessment in Higher Education* (pp. 100-109). New York: Taylor and Francis Group.
- Slavin, R. (2004). Education research can and must address "what works" questions. *Educational Researcher*, 33(1), pp. 27-28.

- Sluijsmans, D. M. A., Brand-Gruwel, S. & Merrienboer, J. J. G., (2002). Developing procedures for implementing peer assessment in large classes using an action research process. *Assessment and Evaluation in Higher Education*, 27 (5), 442-454.
- Stiggins, Richard & Chappuis, Jan (2005). Using student-involved classroom assessment to close achievement gaps. *Theory into Practice*, 44(1), 11-18.
- Suri, Harsh (1999). *A methodologically inclusive model for research synthesis*. Retirado em Março 5, 2012 de <http://www.aare.edu.au/99pap/sur99673.htm>
- Suri, Harsh (2002). *Essential features of methodologically inclusive research synthesis*. Retirado em Março 5, 2012 de http://www.eric.ed.gov/ERICWebPortal/search/detailmini.jsp?_nfpb=true&_ERICExtSearch_SearchValue_0=ED464923&ERICExtSearch_SearchType_0=no&accno=ED464923
- Suri, Harsh & Clarke, David (1999). *Revisiting methods of literature synthesis*. Retirado em Março 5, 2012 de http://www.eric.ed.gov/ERICWebPortal/search/detailmini.jsp?_nfpb=true&_ERICExtSearch_SearchValue_0=ED464958&ERICExtSearch_SearchType_0=no&accno=ED464958
- Vieira, Isabel (2007). *O feedback nas práticas avaliativas de dois professores de Português do ensino secundário*. Dissertação de Mestrado em Ciências da Educação (Avaliação em Educação). Universidade de Lisboa: Faculdade de Psicologia e de Ciências da Educação, Lisboa, Portugal.
- Taras, Madalena (2002). Using assessment for learning and learning from assessment. *Assessment and Evaluation in Higher Education*, 27, (6), 501-510.
- Trotter, Eileen (2006). Student perceptions of continuous summative assessment. *Assessment and Evaluation in Higher Education*, 31 (5), 505-521.
- Yorke, Mantz (2006). Formative assessment in higher education: Moves towards theory and the enhancement of pedagogic practice. *Higher Education*, 47, 477-501.
- Walker, Melanie & Warhurst, Chris (2000). In most classes you sit around very quietly at a table and get lectured at...? debates, assessment and student learning. *Teaching in Higher Education*, 1 (1), 33-49.
- Wanous, Mohammed, Procter, Brian & Murshidb, Kinaz (2009). Assessment for learner self evaluation: Enhancing achievement in first year using learning technologies. *European Journal of Engeneiring Education*. 34 (1), 77-85.
- Weed, Mike (2008). A potential method for the interpretive synthesis of qualitative research: Issues in the development of 'meta-interpretation'. *International Journal of Social Research Methodology*, 11 (1) 13-28.
- Wilson, Mark & Scalise, Kathleen (2006). Assessment to improve learning in higher education: The BEAR Assessment System. *Higher Education*, 52, 635-663.