

VINTE E CINCO ANOS DE AVALIAÇÃO DAS APRENDIZAGENS: UMA SÍNTESE INTERPRETATIVA DE LIVROS PUBLICADOS EM PORTUGAL

Domingos FERNANDES¹

Introdução²

A investigação que se apresenta e discute neste artigo faz parte de um projecto mais amplo que tinha como finalidade descrever, analisar e interpretar as produções teóricas e as investigações desenvolvidas em Portugal no domínio da avaliação das aprendizagens. Trata-se de uma síntese elaborada a partir da revisão de literatura de natureza empírica e de literatura de natureza mais teórica e reflexiva, constante em teses de doutoramento, em dissertações de mestrado, em livros e em artigos publicados em revistas científicas. Desde o início do projecto, em Maio

1 Faculdade de Psicologia e de Ciências da Educação da Universidade de Lisboa.

2 O projecto em que se enquadra a investigação que aqui se apresenta e discute tem sido financiado pela Unidade de Investigação & Desenvolvimento de Ciências da Educação da Faculdade de Psicologia e de Ciências da Educação da Universidade de Lisboa. Agradeço na pessoa do seu Coordenador, Professor Doutor Rui Canário, todo o apoio concedido.

de 2003, até ao presente, foram consultadas cerca de 350 publicações e directamente analisadas mais de 200.

O projecto desenvolveu-se em dois momentos. No primeiro elaboraram-se quatro sínteses relativas a cada um dos seguintes conjuntos de literatura: a) teses de doutoramento realizadas nas universidades com assento no Conselho de Reitores das Universidades Portuguesas (CRUP); b) dissertações concluídas entre 1994 e 2005 (inclusive) no âmbito de programas de mestrado realizados em universidades com assento no CRUP; c) livros de autores portugueses publicados entre 1981 e 2005 (inclusive); e d) artigos publicados a partir de 1985 em revistas portuguesas de índole científica. No segundo momento é feita uma meta-síntese e uma meta-interpretação (Weed, 2005) das sínteses elaboradas na primeira fase.

O presente texto incide sobretudo na avaliação formativa e decorre de uma síntese baseada em livros que foi elaborada na primeira fase do projecto.

Para além desta introdução este artigo está organizado nas seguintes partes: a) Problema e questões da investigação; b) Enquadramento teórico e conceptual; c) Método; d) Uma caracterização geral dos livros analisados; e) Uma síntese interpretativa no domínio da avaliação formativa e f) Considerações e reflexões Finais.

Problema e questões da investigação

Nos últimos 25/30 anos foram publicados, por autores portugueses, algumas centenas de trabalhos de natureza diversa no domínio da avaliação das aprendizagens dos alunos. Este facto, em si mesmo, torna clara a necessidade de se desenvolver investigação que sintetize o que é que, em Portugal, tem vindo a ser produzido no domínio da avaliação das aprendizagens e que seja capaz de: a) caracterizar globalmente o que tem sido produzido pelos investigadores e estudiosos portugueses; b) identificar e discutir os problemas e os temas que mais têm merecido a sua atenção; c) analisar como é que esses problemas têm sido abordados;

e d) contribuir para o desenvolvimento teórico e prático de um domínio científico e pedagógico de reconhecida importância (Black e Wiliam, 1998; Stake, 2006).

A investigação foi desenvolvida tendo em conta as seguintes questões orientadoras:

1. Como é que se poderão caracterizar globalmente os livros analisados?
2. Como é que se poderão caracterizar as abordagens utilizadas pelos autores em áreas tão relevantes como a avaliação formativa ou a avaliação sumativa?
3. Que sínteses e reflexões se poderão elaborar relativamente à literatura revista?

Enquadramento teórico e conceptual

Esta parte do texto está organizada em duas secções. Na primeira apresentam-se e discutem-se algumas questões relativas à conceptualização da avaliação formativa e na segunda discutem-se brevemente alguns problemas inerentes à concepção e utilização de sínteses de literatura.

ACERCA DA AVALIAÇÃO FORMATIVA E DE OUTRAS QUESTÕES RELACIONADAS

Um dos momentos marcantes da história da avaliação educacional foi a elaboração dos conceitos de avaliação formativa e de avaliação sumativa no contexto da avaliação de programas (Scriven, 1967) que, em poucos anos, passaram a ser utilizados no âmbito da avaliação das aprendizagens dos alunos (Bloom, Hastings e Madaus, 1971). Estes investigadores tiveram um papel decisivo na introdução de práticas pedagógicas que, então, foram consideradas inovadoras tais como: a) as práticas de avaliação formativa nas salas de aula; b) as actividades de remediação; c) a auto-avaliação de alunos e professores; d) a diversificação de

métodos de recolha de informação avaliativa; e ainda e) a diversificação de objectos de avaliação, considerando não só elementos relativos aos conhecimentos académicos mas também os que se referiam aos aspectos emocionais e afectivos e psico-motores. Além disso, Bloom e os seus colaboradores consideravam fundamental que o ensino e as aprendizagens não se centrassem apenas nos níveis mais básicos propostos na sua taxonomia de objectivos educacionais (Bloom, 1956), mas que pudessem alargar-se a processos mais complexos de pensamento tais como os que se relacionam com a análise, a síntese e a avaliação. Consequentemente, as tarefas de resolução de problemas e as actividades experimentais eram recomendadas e incentivadas. Estes são alguns dos pilares da chamada *pedagogia para a mestria*, muito divulgada e posta em prática entre nós nos anos 70 e 80 e que veio dar uma relevância, até então nunca vista, à avaliação pedagógica e aos processos que lhe são inerentes.

Com a emergência de novas racionalidades e paradigmas de investigação surgem novas conceptualizações da avaliação das aprendizagens e, muito particularmente, da avaliação formativa. Em contraste com a avaliação formativa dos anos 70, muito orientada para os resultados e realizada *após* uma dada unidade curricular (Bloom, Hastings e Madaus, 1971), emerge, no início dos anos 80, uma avaliação formativa que ocorre *durante* o processo de ensino e aprendizagem, mais atenta aos processos sem descurar a importância dos resultados (e.g., Allal, 1979, 1986; Earl, 2003; Harlen e James, 1997; Torrance e Prior, 2001). Trata-se de uma avaliação interactiva e contínua que pressupõe a participação activa dos alunos, nomeadamente através dos processos de auto-avaliação, de auto-regulação e de auto-controlo.

Os procedimentos de recolha de evidências de aprendizagem têm como principal finalidade apoiar e orientar os alunos na melhoria das suas aprendizagens. Consequentemente, o *feedback* é essencial no processo de avaliação formativa e deve ser proporcionado de forma inteligente tendo em conta aspectos tais como a distribuição e a frequência e a sua natureza mais descritiva ou mais avaliativa (Fernandes, 2005a).

Ainda persistem alguns mal entendidos na literatura relativamente ao conceito de avaliação formativa. Por exemplo, nem sempre se faz a distinção entre a avaliação formativa de natureza behaviourista e a avaliação formativa mais próxima das perspectivas construtivistas, cognitivistas e socioculturais. Além disso, a avaliação formativa é por vezes identificada com um conjunto de práticas mais ou menos indiferenciadas que pouco terão a ver com a sua real natureza.

É importante que se saiba do que estamos verdadeiramente a falar quando se utiliza a expressão *avaliação formativa*. Por esta razão, tem-se utilizado a expressão *avaliação formativa alternativa* para designar uma avaliação formativa emergente, de raiz cognitivista e construtivista, e não a confundir com a avaliação de natureza mais técnica ou com as tais práticas indiferenciadas e pouco claras de avaliação que se designam como formativas sem que realmente o sejam (Fernandes, 2005a, 2005b, 2006).

Vários autores têm vindo a estabelecer uma distinção entre avaliação *das* aprendizagens e avaliação *para* as aprendizagens (e.g., Black e Wiliam, 2006a; Gardner, 2006a, 2006b; Harlen, 2006; Sebba, 2006) precisamente para sublinharem a diferença entre a avaliação sumativa e a avaliação formativa. A avaliação formativa está associada a todo o tipo de tomadas de decisão e de formas de regulação e de auto-regulação que influenciam de forma imediata os processos de ensino e aprendizagem, enquanto a avaliação sumativa proporciona informação sintetizada que, no fundo, se destina a registar e a tornar público o que parece ter sido aprendido pelos alunos.

Earl (2003) faz referência à avaliação *como* aprendizagem querendo com isto destacar o papel que os alunos devem ter no processo de avaliação, nomeadamente através da auto-regulação e do auto-controlo. Trata-se de uma concepção análoga à que foi desenvolvida por Nunziati (1990) com a avaliação formadora e que, na verdade, dificilmente se diferencia de uma avaliação formativa mais global e abrangente, tal como hoje é concebida pela generalidade da comunidade científica.

Harlen e James (1997) e Harlen (2005, 2006) são duas investigadoras que têm produzido trabalho de reflexão teórica acerca das relações entre

a avaliação formativa e a avaliação sumativa ou, se quisermos, entre a avaliação *para* as aprendizagens e a avaliação *das* aprendizagens. Para estas autoras a avaliação formativa tem uma dupla natureza. É *criterial*, porque, no decorrer dos processos de ensino-aprendizagem-avaliação, as aprendizagens dos alunos não são comparadas com algum padrão ou norma, mas analisadas em termos de critérios que são definidos previamente. É *ipsativa* porque está centrada no aluno, isto é, compara o aluno consigo mesmo, tendo em conta aspectos tais como o esforço, o contexto em que o trabalho se desenvolve e os seus progressos.

A avaliação sumativa tem também uma natureza *criterial* e é *normativa*, isto é, compara as aprendizagens dos alunos com uma norma (uma média, por exemplo) ou com as aprendizagens de um dado grupo. A partir destes pressupostos considera-se que, forçosamente, há uma forte articulação entre as duas modalidades de avaliação porque ambas partilham a sua natureza criterial.

Há, no entanto, relativamente a esta ideia, algumas questões sobre as quais convirá reflectir. Por exemplo: Será que a avaliação formativa tem necessariamente uma natureza criterial? Ou poderá orientar-se antes por referentes mais abrangentes e complexos que orientem o desenvolvimento das aprendizagens? Quando se considera que a avaliação formativa é de natureza criterial de que avaliação formativa estaremos realmente a falar? Quais as suas características? Partindo do princípio de que a avaliação formativa é criterial que diferenças haverá na utilização de critérios entre esta e a avaliação sumativa? Qual será o real sentido deste tipo de concepção? Não estaremos a diluir as significativas diferenças existentes entre as duas modalidades de avaliação?

No seu trabalho mais recente Harlen (2005, 2006) discute se os dados da avaliação sumativa podem ser usados na avaliação formativa e vice-versa e se a avaliação formativa e a avaliação sumativa são duas dimensões diferentes de um mesmo constructo ou se, pelo contrário, são dois conceitos distintos (visão dicotómica). Apesar de reconhecer que a avaliação formativa e a avaliação sumativa são conceptualmente diferentes, considera que, na prática, tais diferenças não serão tão evidentes. No entanto, aquela autora acaba por concluir que é complicado

utilizar os dados produzidos no âmbito da avaliação sumativa para efeitos formativos, uma vez que as avaliações sumativas: a) se destinam a sintetizar as aprendizagens dos alunos com base em critérios gerais; b) não têm normalmente o propósito de identificar e de orientar os passos a dar a seguir; e c) ocorrem com pouca frequência. Nestas condições é evidente que as avaliações sumativas não preenchem os exigentes requisitos das avaliações formativas.

Quanto à utilização das avaliações formativas para fins sumativos, Harlen vê essa possibilidade embora sem grande entusiasmo pois, no fundo, argumenta que as avaliações formativas têm problemas vários de validade e de fiabilidade por ocorrerem em contextos muito específicos, faltando-lhes, por isso, consistência para que possam servir propósitos sumativos. Porém, a possibilidade não é completamente posta de lado desde que, de acordo com aquela investigadora, se faça a distinção entre as evidências de aprendizagem e as interpretações que se fazem a partir delas. Para efeitos de avaliação formativa as evidências são interpretadas *localmente*, no contexto da sala de aula, e destinam-se a decidir o que fazer a seguir naquele mesmo contexto. Para efeitos de avaliação sumativa as evidências têm que ser interpretadas tendo em conta critérios mais gerais uma vez que vai ser atribuída uma classificação final e, neste caso, a autora questiona a validade e a consistência da avaliação formativa.

No que se refere à natureza das relações entre a avaliação formativa e a avaliação sumativa, é rejeitada a possibilidade de as considerar como conceitos dicotómicos preferindo-se enquadrá-las conceptualmente como dimensões de um mesmo constructo. Na verdade, Harlen (2006) vê a avaliação formativa informal (uma espécie de avaliação formativa *pura* ou *ingénua*) e a avaliação sumativa formal como dois pólos extremos de uma possível dimensão de propósitos e práticas de avaliação. Entre estes dois pólos estão a avaliação formativa formal e a avaliação sumativa informal que, curiosamente, apresentam muitas semelhanças.

Da discussão anterior parece ter ficado claro que existe uma *zona* situada entre a avaliação formativa e a avaliação sumativa que se assemelha a uma espécie de *terra de ninguém*. Ou seja, onde se poderão

identificar e onde coexistem práticas mais ou menos indiferenciadas que ou se aproximam, ou se confundem, com um tipo de avaliação ou com o outro. Para apoiar e organizar uma discussão sobre este assunto formulou-se a seguinte questão: *Que articulações e relações é possível estabelecer entre a avaliação formativa e a avaliação sumativa?* (Fernandes, 2005a). Na resposta preconiza-se que a avaliação formativa alternativa deve ser claramente preponderante nas práticas de sala de aula e argumenta-se que os dados recolhidos através da avaliação formativa podem e devem ser utilizados pelos professores nas avaliações sumativas que são da sua responsabilidade. A avaliação sumativa pode assim tornar-se uma súmula das aprendizagens claramente mais rica, mais inteligente, mais contextualizada e mais útil para todos os intervenientes, muito particularmente para os alunos.

Considera-se assim que existe, quase naturalmente, uma forte articulação entre as duas modalidades de avaliação. De facto, defende-se que uma avaliação sumativa de qualidade nas salas de aula “deve estar subordinada aos princípios, aos métodos e aos conteúdos da avaliação formativa alternativa” (Fernandes, 2005a, p. 75). Este princípio tem um alcance significativo em termos pedagógicos e didáticos e, particularmente, em termos da integração da avaliação formativa com o ensino e com a aprendizagem, permitindo recolher informação de boa qualidade (veja-se, por exemplo, Black e Wiliam, 2006c; Stobart, 2006) e envolvendo activamente professores e alunos em todas as etapas fundamentais do desenvolvimento do currículo nas salas de aula. Verificando-se a prática de uma verdadeira avaliação formativa, a avaliação sumativa acaba por consistir num momento particularmente rico e devidamente ponderado de integração e de síntese da informação recolhida acerca do que os alunos sabem e são capazes de fazer numa variedade de situações. E isto significa que a informação obtida a partir dos processos decorrentes da avaliação formativa não pode deixar de ser devidamente considerada e integrada com outros que decorrem dos processos próprios da avaliação sumativa (por exemplo, um teste realizado com o propósito de se fazer um balanço intermédio ou final relativamente às aprendizagens desenvolvidas no âmbito de

uma dada unidade curricular e que pode permitir o estabelecimento de referentes comuns a todos os alunos).

O que toda esta argumentação significa é que a avaliação formativa que, por natureza, acompanha os processos de aprendizagem, não pode deixar de proporcionar informação de qualidade quanto ao desenvolvimento desses mesmos processos por parte dos alunos. Que não se fique com a absurda ideia, porém, de que a avaliação formativa se transforma num conjunto de pequenas avaliações sumativas!! Nada disso. Quando se verifica que um aluno não sabe o que é suposto saber, a avaliação formativa deve estar presente para o ajudar a melhorar, a vencer a dificuldade, com o seu esforço e o esforço do professor. Ora o que interessa ao professor em termos de avaliação sumativa, em termos de balanço, não é o facto de um aluno não saber algo num dado dia e, por isso, entrar com uma informação negativa para esse mesmo balanço. O que verdadeiramente lhe interessa é saber: a) se o aluno ficou a saber; b) como é que ultrapassou as dificuldades; c) as razões que poderão ter impedido que assim acontecesse; e d) o que foi efectivamente feito pelo aluno e pelo professor para dissipar as dificuldades.

A avaliação formativa e a avaliação sumativa não se podem confundir uma com a outra. Têm propósitos distintos, ocorrem em momentos distintos e têm uma *inserção* pedagógica distinta. Será que têm pressupostos epistemológicos também distintos? Se pensarmos no papel e na participação dos alunos e dos professores, no tipo de informação que se obtém num e noutro caso e no seu significado podemos considerar que há diferenças epistemológicas. Nestas condições não será fácil compreender a ideia de que ambas são dimensões distintas de um mesmo constructo. Talvez esta visão se aplique melhor à avaliação formativa e sumativa tal como foram conceptualizadas em 1967. Mas o conceito de avaliação formativa evoluiu muito em relação a essa altura. Por outro lado, também não me é fácil percepcioná-las numa visão dicotómica. Vejo-as mais numa relação complementar ou numa articulação motivada por razões que não têm a ver com a sua natureza intrínseca mas que, de certo modo, são impostas por imperativos relacionados com a referência social que a avaliação não pode contornar.

Temos aqui, seguramente, um interessante tema para uma discussão que merece ser aprofundada, nomeadamente tendo em conta os esforços de desenvolvimento teórico que têm vindo a ser concretizados no âmbito da avaliação formativa (e.g., Black e Wiliam, 2006b; Gipps, 1994).

ACERCA DAS SÍNTESES DE LITERATURA

As sínteses de literatura têm vindo a ganhar uma importância crescente como actividades primárias de investigação (Weed, 2005). É possível identificar uma variedade de abordagens à concepção e à elaboração de sínteses que revelam o reconhecimento crescente desta forma de investigação (ver, por exemplo, Cooper, 1988, 1998, 2003; Glass, 1976; Noblit e Hare, 1988; Slavin, 1984, 1986, 2004; Suri, 1999a, 1999b, 2002; Weed, 2005). As sínteses constituem reflexões sistematizadas que integram o trabalho teórico ou o trabalho de investigação empírica realizado por outros e que, no final, deverão traduzir mais do que *a soma das partes* analisadas.

Uma das vantagens dos trabalhos de síntese da literatura tem a ver com o facto de permitirem que os leitores e investigadores possam ter uma visão abrangente e profunda de desenvolvimentos teóricos e práticos numa dada área do conhecimento ao longo de um dado período de tempo. Interessa, porém, diferenciar entre sínteses ou revisões de literatura que são desenvolvidas pelos investigadores com o propósito de enquadrarem conceptualmente uma dada investigação primária e as sínteses da literatura cujo principal objectivo é o de contribuir para a compreensão de um dado problema ou domínio de investigação a partir da análise dos resultados e conclusões de um conjunto mais ou menos alargado de investigações primárias. A este propósito Suri (1999a, 1999b) considera que enquanto no primeiro caso a síntese é um meio para atingir um fim – contextualizar teórica e conceptualmente uma investigação – no segundo caso ela é um fim em si mesmo – produzir compreensão sobre um dado fenómeno. (Na presente investigação a síntese é encarada nesta última perspectiva.)

Vários autores (e.g., Cooper e Lindsay, 1998; Suri, 1999b, 2002; Suri e Clarke, 1999) referem que as sínteses são um desafio à imaginação para que se encontrem formas de tornar visível, de forma integrada e simples, a complexidade, a diversidade, os propósitos, métodos, perspectivas e conclusões constantes num número mais ou menos elevado de investigações ou outros trabalhos. Neste sentido, contribuem para identificar e delinear novos projectos de investigação.

Há várias sínteses no domínio da avaliação das aprendizagens que se tornaram referências incontornáveis nesta área do conhecimento. Algumas das mais recentes foram elaboradas no âmbito da avaliação formativa por Allal e Lopez (2005), que fazem uma revisão da literatura teórica e empírica de trabalhos publicados em língua francesa, e por Koller (2005) que faz uma revisão da literatura empírica alemã. Bem conhecida é a revisão da literatura de Black e Wiliam (1998), também no domínio da avaliação formativa. São também incontornáveis os trabalhos de síntese já clássicos de Crooks (1988) e de Natriello (1987), ambos referentes a impactos dos processos de avaliação nos estudantes. Sousa (1995) fez uma síntese de literatura de investigação na área da avaliação das aprendizagens publicada no Brasil entre 1930 e 1980.

Em Portugal foram recentemente elaboradas duas sínteses centradas quase exclusivamente em trabalhos de investigação empírica (Barreira e Pinto, 2006; Neves, Jordão e Santos, 2004).

Suri e Clarke (1999a) identificam os seguintes tipos de sínteses de investigação existentes na literatura: a) as chamadas revisões narrativas de investigação; b) as meta-análises; c) as sínteses realizadas a partir das “melhores evidências” (*best-evidence synthesis*); e d) as sínteses de investigações de natureza qualitativa. É uma categorização que parece ter sido determinada pelos métodos utilizados em cada um dos tipos de síntese. Porém, em trabalhos mais recentes, Suri (1999b, 2002) faz referência a sínteses cuja categorização parece ter sido mais orientada pela natureza dos objectos de análise e/ou do *estado* em que se encontra a investigação no domínio de interesse. Assim são consideradas: a) *sínteses agregadoras*, quando estamos perante um conjunto mais ou menos homogéneo de evidências; b) *sínteses interpretativas*, quando

estamos a lidar com um conjunto heterógeneo de resultados decorrentes de diferentes metodologias e a examinar diferentes conceitos; c) *sínteses exploratórias*, utilizadas quando se pretende identificar linhas de investigação que de algum modo são promissoras ou quando se está perante um domínio pouco investigado; e d) *sínteses confirmatórias*, que serão preferíveis quando estamos a analisar estudos de uma área já bastante investigada e consolidada.

Naturalmente que, tal como acontece em estudos de avaliação e em investigações, as abordagens a utilizar nas sínteses de literatura devem estar mais dependentes dos seus contextos e propósitos do que de qualquer ortodoxia teórica ou metodológica.

Método

O processo de identificação, selecção e recolha dos livros incluídos nesta investigação foi desenvolvido a partir da Biblioteca Nacional, das Bibliotecas de Universidades, de Faculdades e/ou dos respectivos Departamentos, quase todas disponíveis *on-line*, do Centro de Documentação e Informação da Direcção Geral de Inovação e de Desenvolvimento Curricular (DGIDC), que inclui o acervo da Biblioteca do extinto Instituto de Inovação Educacional, e ainda a partir dos motores de pesquisa dos sítios de Editoras disponíveis na *world wide web*. Foram ainda pesquisadas as bases de dados ColCat, PORBASE e SIRIUS.

Após este processo de pesquisa organizou-se uma base de dados integrando as referências essenciais de cada um dos livros que se considerava ter condições para integrar a investigação. Para que um livro fosse seleccionado tinha que obedecer simultaneamente aos seguintes critérios:

1. A avaliação das aprendizagens dos alunos era um dos principais temas organizadores do livro ou de alguns dos seus capítulos.
2. O livro tinha de ter sido publicado em Portugal entre 1981 e 2005 (inclusive).

3. O primeiro autor do livro tinha de ser de nacionalidade portuguesa.
4. O livro tinha de ter uma tiragem igual ou superior a 500 exemplares.

Não foram, nesta fase da investigação, seleccionados para análise um conjunto ainda significativo de publicações tais como: a) monografias universitárias, trabalhos de fim de curso, “sebentas” de apoio aos alunos, livros ou brochuras publicadas no âmbito de faculdades, departamentos ou serviços editoriais universitários com tiragens limitadas e/ou circulação relativamente restrita (e.g., Costa, 1989; Ferreira, 2004; Silva, 1989; Sobral e Barreiros, 1980); b) livros ou outras publicações brochadas editadas por departamentos ou serviços do Ministério da Educação ou outras instituições independentes que funcionem no seu âmbito (e.g., Abrantes e Araújo (orgs.), 2002; Conselho Nacional de Educação, 1995; Fernandes, Ramalho e Lemos, 1991; Fernandes, s/d; Oliveira, Pereira e Fernandes, 1993); c) actas de encontros de natureza científico-pedagógica ou livros que eventualmente integrem artigos ou capítulos sobre avaliação das aprendizagens (e.g., Estrela e Nóvoa, 1992; Guimarães, Leal e Abrantes (orgs.), 1991; Pacheco e Zabalza, 1995); e d) livros que constituem sobretudo uma espécie de “guias de leitura” ou de interpretação dos normativos legais (e.g., Valada, 1993; Cunha, 2004). Uma lista exemplificativa das publicações não analisadas encontra-se no Anexo A.

Tendo em conta os critérios acima identificados foram seleccionados para análise 34 livros cuja lista organizada, de acordo com a data de publicação das primeiras edições, se apresenta no Anexo B.

Para efeitos de organizar a análise e a síntese dos livros seleccionados foram tidas em conta categorias tais como *Avaliação Formativa*, *Funções da Avaliação* e *Concepções e Práticas de Avaliação*. Em seguida foram utilizados procedimentos de recolha de informação que passaram pela leitura holística e focada dos livros ou das suas secções de interesse para os propósitos desta investigação. Procedeu-se então ao registo escrito, através de comentários breves (CB) de natureza descritiva, da informação referente a cada categoria. Ou seja, tal como se ilustra na Figura 1, cada categoria considerada deu origem a uma análise

horizontal *através* de todos os livros seleccionados. Além disso, relativamente a cada livro, procedeu-se a uma análise vertical *através* de todas as categorias consideradas. Assim, elaboraram-se sínteses para cada livro (verticais) e sínteses para cada categoria (horizontais).

	LIVRO A	LIVRO B	LIVRO C	LIVRO D	
CATEGORIA 1	CB1A	CB1B	CB1C	CB1D	SÍNTESE CATEGORIA 1
CATEGORIA 2	CB2A	CB2B	CB2C	CB2D	SÍNTESE CATEGORIA 2
CATEGORIA 3	CB3A	CB3B	CB3C	CB3D	SÍNTESE CATEGORIA 3
CATEGORIA 4	CB4A	CB4B	CB4C	CB4D	SÍNTESE CATEGORIA 4
	SÍNTESE LIVRO A	SÍNTESE LIVRO B	SÍNTESE LIVRO C	SÍNTESE LIVRO D	SÍNTESE GLOBAL

Fig. 1 – Esquema Geral de Análise dos Dados

A análise cruzada daqueles dois conjuntos de sínteses revelou-se importante para que se produzisse uma *síntese global*, identificando o que pareceram ser os aspectos que deveriam merecer destaque. Além disso, permitiu fundamentar a interpretação, a discussão e as reflexões conclusivas acerca dos dados recolhidos a partir do material seleccionado para a investigação.

Não foi possível, no contexto deste artigo, apresentar as análises e sínteses referentes a todas as categorias analisadas.

Uma caracterização geral dos livros analisados

Como acima se referiu foram analisados 34 livros publicados na área da avaliação das aprendizagens entre 1981 e 2005 (inclusive) por dez editoras comerciais, por dois serviços editoriais de universidades (Universidade Aberta e Universidade do Minho) e por uma Fundação (Calouste Gulbenkian). A maioria, dezoito, foi publicada na década de 90 enquanto que na década de 80 foram publicados apenas oito. Entre 2000 e 2005, publicaram-se oito livros relativos à avaliação das

aprendizagens dos alunos o que pode indiciar uma estabilização ou mesmo uma redução do número de livros publicados nesta área tanto mais que, em 2006, ainda não se identificou nenhum título. Vinte e dois livros foram publicados por apenas três Editoras: a Porto Editora, a ASA e a Texto, com nove, oito e cinco livros, respectivamente. Das restantes sete editoras comerciais duas publicaram dois livros cada uma e cinco apenas um cada uma. As três editoras sedeadas na cidade do Porto publicaram dezanove livros, as cinco da zona de Lisboa dez e as duas de Coimbra dois.

Este panorama editorial é algo desolador mas eventualmente não será muito diferente do que se refere a outras áreas igualmente relevantes no domínio da educação.

Apesar de se terem identificado outros livros ou brochuras editados por serviços editoriais universitários, a verdade é que só os dois livros seleccionados (Ribeiro e Ribeiro, 1989; Vieira e Moreira, 1993) tiveram tiragens e uma divulgação que ultrapassou claramente os “muros” das respectivas universidades.

O primeiro autor, ou um dos autores, de cerca de 70% dos livros (vinte e cinco) é docente do ensino superior universitário ou politécnico e maioritariamente do género feminino; de facto, vinte e três dos trinta e quatro livros analisados têm uma mulher como primeira autora ou como uma das co-autoras. Verifica-se ainda que metade dos livros analisados foram escritos por dois ou mais autores cada um e que 13 livros foram escritos por sete autores. Luísa Cortesão e Maria Arminda Torres publicaram três livros em conjunto (Cortesão e Torres, 1981, 1983, 1996) e a primeira destas autoras publicou ainda um outro livro (Cortesão, 1993). Alcino Vilar, Carlinda Leite, José Augusto Pacheco, Lucie Ribeiro e Valter Lemos publicaram, sozinhos ou em co-autoria, dois livros cada um.

OS LIVROS DOS ANOS 80

Todos os livros publicados na década de 80 estão mais ou menos influenciados por uma certa racionalidade técnica muito associada às

perspectivas behaviouristas do ensino e da aprendizagem. Mesmo em livros cujos autores parecem não subscrever propriamente as perspectivas comportamentalistas e associacionistas de ensino e de aprendizagem, a influência de Bloom e dos seus colaboradores, no que se refere à avaliação das aprendizagens, coexiste com influências cognitivistas ao nível dos processos de ensino ou ao nível da organização pedagógica (e.g., Cortesão e Torres, 1981, 1983; Domingos, Neves e Galhardo, 1981). As perspectivas defendidas por Bloom relativamente à organização do ensino e da aprendizagem e à integração da avaliação formativa naqueles processos, estão presentes noutros livros de uma forma mais ou menos assumida (e.g., Lemos, 1986; Ribeiro, 1989; Ribeiro e Ribeiro, 1989). E parece natural que assim tenha sido pois, na altura, os recursos teóricos mais presentes (e mais credíveis) nos meios académicos ainda eram os que provinham da esfera de influência da psicologia comportamentalista (e.g., Bloom, 1956; Bloom, Hastings e Madaus, 1971; Bloom e Krathwohl, 1956; Gronlund, 1970; Harrow, 1972; Krathwohl, Bloom e Mesia, 1964) que, em menor ou maior grau, são referidos ou mesmo discutidos em todos os livros portugueses publicados nos anos 80.

O trabalho de Linda Allal, apresentado em 1978 num colóquio da Universidade de Genebra, em que é feita a distinção entre a avaliação formativa proposta pelos behaviouristas e a avaliação formativa tal como era entendida pelos cognitivistas, foi publicado na Suíça em 1979 e em Portugal em 1986 (Allal, 1979, 1986; Allal, Cardinet e Perrenoud, 1979, 1986). Cortesão e Torres (1981) são os únicos autores que, nos anos 80, referem o trabalho de Allal que, como se sabe, abriu caminho para uma conceptualização cognitivista e construtivista da avaliação formativa.

De modo geral, os livros publicados na década de 80 estão claramente orientados pela *pedagogia para a mestria*, impropriamente designada por muitos como *pedagogia por objectivos*, e, consequentemente, por uma visão algo técnica da avaliação. Daí a ênfase na definição de objectivos em termos comportamentais, na sua organização em taxonomias e na construção de instrumentos, principalmente testes sumativos e formativos, destinados a verificar o seu

grau de consecução por parte dos alunos. Além disso, é evidente a ênfase dada em praticamente todos os livros da época à planificação, bastante estruturada, dos processos de ensino e de aprendizagem. O lugar da avaliação das aprendizagens ainda é algo exterior a estes dois processos e os seus objectos privilegiados são os conhecimentos académicos sob a forma de resultados demonstrados pelos alunos em testes formativos ou sumativos.

Um dos livros foi escrito no contexto da disciplina de Educação Física (Bento, 1987) e outro centrou-se na análise e no desenvolvimento de processos de observação e de inquéritos por questionário (Damas e De Ketele, 1985).

Dir-se-ia ainda que, talvez com a excepção dos livros de Corteção e Torres (1981, 1983), todos os outros têm um formato de manual bastante estruturado, contendo um conjunto de definições, de prescrições e de *explicações* com graus diferenciados de discussão e cuja tónica talvez não convide muito à reflexão e à problematização. A *ciência* da avaliação das aprendizagens surge como algo que está essencialmente construído e a questão parece estar na boa utilização de um conjunto de técnicas que se prescrevem para a fazer funcionar em pleno. Esta concepção continuará presente em vários livros das décadas seguintes.

OS LIVROS DOS ANOS 90

Os anos 90 ficaram marcados por importantes alterações curriculares na sequência da publicação da Lei de Bases do Sistema Educativo em 1986 (Lei n.º 46/86) e do Decreto-Lei n.º 286/89 onde se definiu a nova organização curricular dos ensinos básico e secundário. O sistema de avaliação dos alunos destes níveis de ensino acabaria por ser regulado através de dois despachos: o Despacho Normativo n.º 98-A/92 e o Despacho n.º 338/93 para a avaliação das aprendizagens nos ensino básico e secundário, respectivamente. Entre outras medidas, ambos os despachos definem que a avaliação formativa deve ser a modalidade de avaliação predominante nas salas de aula, advogam a partilha

do processo de avaliação com os alunos, referem que a progressão dos alunos deve ser a regra e a retenção uma excepção e instituem mecanismos de avaliação externa (e.g. provas aferidas no ensino básico e exames nacionais no ensino secundário).

Contam-se por largas dezenas os artigos, as investigações de mestrado e de doutoramento, os livros e outro tipo de publicações que foi possível identificar no âmbito deste projecto de investigação, que *reagiram* ao chamado *novo sistema de avaliação dos alunos*. Na verdade, a sociedade portuguesa talvez nunca tenha discutido tanto a avaliação como nos anos 90 e, por isso mesmo, se considerou então que se vivia o *Tempo da Avaliação* (Fernandes, 1992).

Uma marca importante dos livros publicados na década de 90, reflectindo claramente o período sócio-político e educativo que então se viveu, é pois a reacção mais ou menos apaixonada, mais ou menos crítica, mais ou menos fundamentada teoricamente, aos chamados despachos da avaliação e, muito particularmente, ao Despacho Normativo 98-A/92. Na verdade, dos 18 livros publicados naquela década, 11 referem-nos explicitamente dedicando-lhes capítulos ou partes muito substanciais.

Nalguns casos os livros estão organizados para apoiar os professores no desenvolvimento prático fundamentado de uma avaliação mais consentânea com o que se preconizava nos referidos despachos (e.g., Lemos, Neves, Campos, Conceição e Alaíz, 1992; Vilar, 1993). Noutros casos os autores enquadram teórica e conceptualmente o conteúdo dos despachos ou promovem algum tipo de discussão, mais ou menos sustentada teoricamente, mais ou menos profunda, acerca de conceitos que neles se definem (e.g., Leite (org.), 1993; Pacheco, 1994; Vila Nova, 1997; Vilar, 1993). Ainda noutros casos o sistema de avaliação proposto oficialmente é o pretexto para que se elaborem considerações acerca das políticas educativas ou para que se expressem opiniões acerca das causas e consequências do insucesso escolar dos alunos (e.g., Alves e Formosinho, 1992; Machado, 1994; Pacheco, 1994). Um dos livros relata e divulga uma investigação empírica cujo principal objectivo foi o de estudar alguns impactos da aplicação do Despacho 98-A/92 (Lobo, 1998).

E de alguma forma surpreendente que a parte mais significativa do conteúdo dos livros publicados numa década seja, no essencial, inspirada em despachos do Ministério da Educação. Parece importante que a agenda da comunidade científica esteja menos dependente das iniciativas de política educativa e mais orientada por planos estratégicos próprios e autónomos que se concretizem através de programas de investigação empírica (a este propósito deve referir-se que apenas em quatro dos 34 livros analisados é feita referência a investigação empírica desenvolvida pelos próprios autores).

Mas os livros dos anos 90 não ficam só marcados pelas reacções aos despachos ministeriais da avaliação das aprendizagens dos alunos. Em muitos deles é clara a preocupação em conceptualizar a avaliação de forma mais integrada, considerando o seu lugar nas sociedades e nos sistemas educativos e o papel que os professores poderão ter no desenvolvimento de práticas que ajudem todos e cada um dos alunos a melhorar as suas aprendizagens. A elaboração teórica é mais densa e sofisticada. E há vários factos que ilustram estas características dos livros publicados na década de 90 entre os quais se assinalam os seguintes: a) a diversidade e a abrangência das referências bibliográficas utilizadas por vários autores (e.g., Lobo, 1998; Valadares e Graça, 1998; Vieira e Moreira, 1993); b) a preocupação em enquadrar teoricamente e em distinguir diferentes perspectivas de avaliação (e.g., Damião, 1996; Vilar, 1992; Vila Nova, 1997); c) o cuidado revelado por alguns autores com o rigor conceptual, nomeadamente no que se refere à definição de avaliação formativa (e.g., Pacheco, 1994; Vila Nova, 1997; Lobo, 1998); e d) a discussão empreendida por alguns autores relativamente ao compromisso ético e político que está inerente ao desenvolvimento de uma avaliação que sirva efectivamente para melhorar o ensino e as aprendizagens, ajudando a combater o insucesso escolar, e ao papel que os professores podem ter neste processo (e.g., Baptista, 1999; Cortesão, 1993; Cortesão e Torres, 1996; Vieira e Moreira, 1993).

Pode dizer-se que, em geral, muitos dos livros publicados na década de 90 se aproximam das perspectivas pós-behaviouristas de avaliação,

influenciados pelas teorias cognitivistas e construtivistas das aprendizagens. Apesar de as perspectivas próximas da avaliação como técnica ou como medida, muitas vezes mescladas com outras perspectivas, continuarem a ter alguma preponderância em vários livros (e.g., Valadares e Graça, 1998; Vila Nova, 1997), surgem outras visões que conceitualizam a avaliação como um processo eminentemente pedagógico e didáctico mas também político e ético (e.g., Cortesão, 1993; Cortesão e Torres, 1996; Vieira e Moreira, 1993). A ênfase será ainda, em vários casos, na formulação de objectivos e no estabelecimento de procedimentos para medir a sua consecução, mas começam a surgir perspectivas mais holísticas e integradas das aprendizagens e da sua avaliação. Por outro lado, começa a emergir a perspectiva de uma avaliação integrada nos processos de ensino e de aprendizagem, mais contextualizada e mais abrangente no seu âmbito, compreendendo um alargado leque de saberes sociais, académicos, práticos e afectivos que devem ser objecto de avaliação.

Como resultado dos desenvolvimentos teóricos internacionais mas também da entrada em vigor de um novo sistema de avaliação dos alunos em Portugal, alguns livros realçam o papel da avaliação formativa na melhoria das aprendizagens dos alunos (e.g., Lemos *et al.*, 1992; Cortesão, 1993; Vieira e Moreira, 1993; Pais e Monteiro, 1996).

O livro de Vieira e Moreira (1993) está orientado para a didáctica da língua estrangeira, o de Vilhena (1999) centra-se na avaliação de um espectro de saberes que vai para além dos que decorrem directamente do conteúdos curriculares e o de Sim-Sim (1997) decorre de uma investigação empírica cujo produto principal é um instrumento de avaliação da linguagem oral de crianças de 4, 6 e 9 anos de idade. Todos os outros livros são de natureza genérica e transversal.

É difícil considerar que um dado livro pertence a uma dada categoria porque, na realidade, ele pode ter características que o incluem em mais do que uma. Apesar disso, e ressalvando a real possibilidade de um livro poder incluir-se em mais do que uma categoria, a análise realizada sugere que a maioria dos livros publicados nos anos 90 se possa distribuir por três categorias principais.

Há livros que seguem o formato próximo dos chamados manuais, dando mais relevância à apresentação de factos, definições e teorias, no estrito âmbito da avaliação das aprendizagens, do que à sua discussão ou problematização em contextos mais alargados (e.g., Damião, 1996; Lemos *et al.*, 1992; Valadares e Graça, 1998; Vila Nova, 1997). Por outro lado, há livros de âmbito mais alargado, que contextualizam a avaliação no conjunto do sistema educativo ou na própria sociedade, procurando estabelecer relações com outros elementos e promover uma reflexão teórica e/ou prática, muitas vezes inspirada em perspectivas pedagógicas, sociológicas, curriculares, psicológicas ou mesmo filosóficas (e.g., Baptista, 1999; Cortesão, 1993; Cortesão e Torres, 1996; Pacheco, 1994; Vilar, 1992). Uma terceira categoria inclui livros, com formato mais ou menos próximo do manual, mas que estão orientados para a didáctica de disciplinas curriculares ou em que a perspectiva didáctica está subjacente à organização dos temas que se apresentam e discutem. Não se limitando a apresentar a *matéria*, estes livros suscitam a reflexão e a leitura críticas (e.g., Vieira e Moreira, 1993).

OS LIVROS DA PRIMEIRA METADE DA PRESENTE DÉCADA

Foram analisados oito livros publicados entre 2000 e 2005 (inclusive). Por contraste com o que se passou nos anos 80, os livros desta década parecem inspirados em racionalidades mais críticas e sociocríticas. A avaliação das aprendizagens é entendida como um processo complexo que faz parte integrante das rotinas pedagógicas das salas de aula. É notória em alguns livros (e.g., Alves, 2004; Fernandes, 2005a; Leite e Fernandes, 2002) a presença das duas tradições de investigação e de produção teórica mais influentes no domínio da avaliação das aprendizagens: a) a tradição francófona, através dos trabalhos de Allal (1986, 1988), Allal, Cardinet e Perrenoud (1986), Bonniol (1986), Figari (1996, 2001), Hadji (1992) e tantos outros; e b) a tradição anglo-saxónica, através de trabalhos tais como os de Black e Wiliam (1998, 2006a), Berlak (1992a, 1992b), Gipps (1994), Resnick (1987), Stiggins e Conklin (1992) e Wiggins (1998).

Dos oito livros analisados deste período três elaboram acerca de estratégias de recolha de informação avaliativa tais como os portefolios (Coelho e Campos, 2003; Miranda e Marques, 2003) e os registos de observação (Veríssimo, 2000) enquanto que quatro se poderão considerar livros de reflexão teórica com ênfases diferentes ao nível das recomendações de natureza prática (Alves, 2004; Fernandes, 2005a; Leite e Fernandes, 2002; Rosado e Colaço, 2002). Por exemplo, o livro de Leite e Fernandes está mais orientado para apoiar directamente as práticas dos professores do que qualquer um dos outros. A partir das perspectivas do seu autor acerca do insucesso escolar dos alunos, um dos livros faz claras incursões no domínio das políticas educativas quer sob a forma de questões que são formuladas, quer sob a forma de reflexões que acabam por dar origem a recomendações concretas (Fernandes, 2005a). Outro reúne num reduzido número de páginas algumas perspectivas genéricas sobre avaliação (Cabral, 2003).

Um dos livros foi escrito tendo em vista a avaliação das aprendizagens no domínio das actividades físicas (Rosado e Colaço, 2002). Alves (2004) e Leite e Fernandes (2002) dão vários exemplos concretos de avaliação, nomeadamente no âmbito das áreas curriculares não disciplinares do ensino básico. Por seu lado, Coelho e Campos (2003) contextualizam a utilização do portefolio no âmbito da disciplina de Português do ensino secundário.

Tendo em conta as três categorias em que se procuraram organizar os livros publicados na década de 90, parece adequado dizer-se que os livros de Coelho e Campos (2003), Leite e Fernandes (2002), Miranda e Bernardes (2003) e Veríssimo (2000) se incluem na primeira daquelas categorias pois têm a natureza de manuais de apoio directo às práticas dos professores, com a inclusão, nalguns casos, de exemplos e de materiais passíveis de ser directamente aplicados. Os livros de Alves (2004) e Fernandes (2005a) parecem ter um pendor mais teórico e, por isso, seriam incluídos na segunda categoria. Finalmente, na terceira categoria, inclui-se o livro de Rosado e Colaço (2002) pela didáctica das actividades físicas que lhe está subjacente.

Uma síntese interpretativa no domínio da avaliação formativa

A avaliação formativa é um processo intrinsecamente ligado a todas as actividades de ensino e de aprendizagem desenvolvidas por alunos e/ou professores; por natureza, ela gera informação que pode ser utilizada como *feedback* para melhorar o ensino e a aprendizagem. Trata-se de um conceito que tem vindo a ser considerado cada vez mais relevante no processo de ajudar os alunos a aprender melhor, com compreensão e com significado.

Há uma abundante literatura de natureza teórica e empírica em que se discute o conceito de avaliação formativa. Particularmente nas últimas duas décadas tem havido uma evolução muito significativa decorrente dos desenvolvimentos teóricos ocorridos em domínios científicos que lhe estão naturalmente associados (e.g., teorias do currículo, da aprendizagem, da comunicação, da cognição e da metacognição). Tal evolução tem sido sobretudo marcada pelos trabalhos de autores que, como acima já se referiu, representam as duas tradições teóricas mais influentes ao nível internacional: a francófona e a anglo-saxónica. A primeira está mais próxima das abordagens sociocognitivas, em que a regulação é um conceito chave (e.g., Abrecht, 1991; Allal, 1979; Bonniol, 1989; Cardinet, 1986, 1991; De Ketele, 1986, 1993, 2001; Figari, 1996, 2001; Grégoire, 1996; Hadji, 1992; Jorro, 2000; Nunziati, 1990; Perrenoud, 1998a, 1998b; Scallon, 2004; Weiss (Dir.), 1991) e a segunda está mais próxima das abordagens socioculturais, com o *feedback* a ocupar o lugar central (e.g., Berlak, 1992a, 1992b; Black e Wiliam, 1998; Gifford e O' Connor, 1992; Nuttal, 1986; Phye, 1997; Sadler 1998; Shepard, 2000, 2001; Stiggins e Conklin, 1992; Torrance e Prior, 2001; Tunstall e Gipps; 1996; Wiggins, 1998). A primeira dá maior destaque ao papel dos alunos na regulação das suas próprias aprendizagens e a segunda ao papel dos professores como *distribuidores* inteligentes de *feedback* (Fernandes, 2005a, 2006).

Da análise desta literatura internacional (Fernandes, 2005a, 2006) podem identificar-se áreas que, no domínio da avaliação formativa, parecem ser as que mais têm merecido a atenção dos investigadores tais como: a) relações entre a avaliação formativa e a avaliação sumativa; b) relações entre as práticas de avaliação formativa e as aprendizagens efectivamente realizadas pelos alunos; c) papéis de alunos e de professores no processo de avaliação formativa; d) utilizações da informação gerada pelas práticas de avaliação formativa; e) feedback, regulação e avaliação formativa; f) relações entre conhecimentos, concepções e práticas de avaliação formativa dos professores; g) desenvolvimento curricular, selecção de tarefas e avaliação formativa (integração da avaliação formativa nos processos de ensino e aprendizagem); h) relações entre a avaliação formativa e as teorias da aprendizagem; i) qualidade das avaliações formativas (validade e fiabilidade); e j) construção de uma teoria da avaliação formativa.

Esta lista de dez áreas problemáticas e de investigação no domínio da avaliação formativa contribuiu significativamente para enquadrar a literatura que aqui se apresenta e discute.

Em 25 dos 34 livros analisados os autores discutem explicitamente, ainda que com graus muito diferenciados de profundidade e de abrangência, o conceito de avaliação formativa. Alguns autores que não o fazem centraram-se num método específico de recolha de informação avaliativa (e.g., Bernardes e Miranda, 2003; Coelho e Campos, 2003; Damas e De Ketele, 1985; Sim-Sim, 1997), outros referem o conceito sem que verdadeiramente o discutam (e.g., Alves e Formosinho, 1993) e noutros casos há algum tipo de discussão que tem implícitas características importantes do conceito, apesar de não haver propriamente uma discussão a propósito do mesmo (e.g., Leite (org.), 1993).

De acordo com Cortesão (1993) e Lobo (1998) a avaliação formativa começa a ser assumida e recomendada em documentos produzidos pelo Ministério da Educação a partir de 1976, nomeadamente para dar orientações aos professores no contexto da avaliação dos alunos do ensino unificado. Este é um facto interessante porque o livro que *introduz* a avaliação formativa no contexto escolar foi publicado em

1971 (Bloom, Hastings e Madaus, 1971) após a criação do conceito, em 1967, no âmbito da avaliação de programas (Scriven, 1967).

De modo geral pode dizer-se que os livros analisados foram acompanhando os desenvolvimentos internacionais. Para além das referências feitas aos autores ditos clássicos, é interessante constatar a notória influência de Linda Allal (1986, 1979) no que se refere à distinção entre a avaliação formativa de natureza behaviourista e a avaliação formativa de natureza construtivista e cognitivista. Efectivamente aquela distinção é feita em alguns livros (Alves, 2004; Fernandes, 2005a; Lobo, 1998; Pacheco, 1994; Rosado e Colaço, 2002; Vila Nova, 1997). Curiosa é também a referência a Nunziati (1990) relativamente ao conceito de avaliação formadora que é apresentado e discutido por alguns autores (e.g., Alves, 2004; Leite e Fernandes, 2002; Pais e Monteiro, 1996; Vila Nova, 1997). É interessante esta escolha num contexto em que se poderiam considerar outras *avaliações formativas* em que os alunos, tal como acontece em qualquer avaliação formativa, também têm um papel destacado: a) a *avaliação autêntica* (Wiggins, 1998); b) a *avaliação contextualizada* (Berlak, 1992a); c) a *avaliação reguladora* (Allal, 1986; Perrenoud, 1998b); d) a *regulação controlada dos processos de aprendizagem* (Perrenoud, 1998a); ou ainda e) a *avaliação educativa* (Gipps, 1994).

É ainda importante destacar o facto de há 25 anos atrás Domingos *et al.* (1981) sublinharem de forma tão clara que a avaliação deve estar associada à melhoria do ensino e da aprendizagem e que a avaliação formativa está integrada no processo de ensino-aprendizagem. Aquelas autoras, que seguiam de perto a abordagem behaviourista às questões da avaliação das aprendizagens dos alunos que, então, era o recurso teórico mais influente, diziam o seguinte: “Se aceitarmos que a principal finalidade da avaliação é melhorar o ensino e a aprendizagem então este tipo de avaliação – avaliação formativa – desempenha papel preponderante. (...) É durante o processo de ensino-aprendizagem que se faz a avaliação formativa” (p. 212).

Vilar (1992) destaca funções da avaliação formativa (melhorar, desenvolver) e da avaliação sumativa (separar, seleccionar) considerando

que são contraditórias: “A ‘avaliação formativa’ põe em evidência a contradição entre a lógica do desenvolvimento pessoal dos intervenientes e a lógica social da triagem e selecção” (p. 30). Trata-se de uma questão teórica relevante, as relações entre a avaliação formativa e a avaliação sumativa, que tem sido pouco debatida entre nós. Um dos problemas relacionados com essa questão consiste em saber se as duas modalidades de avaliação são conceitos dicotómicos, contraditórios, complementares ou duas dimensões de um mesmo constructo (veja-se, por exemplo, Harlen, 2006).

Cortesão (1993) afirma que “(...) a avaliação formativa na sua forma ideal acontecerá ao longo do processo de ensino-aprendizagem e nunca poderá, formalmente, ser usada para classificar e muito menos para decidir da passagem ou da reprovação do aluno”(p. 12). Esta afirmação também suscita uma questão teórica importante: Podem as informações obtidas no âmbito da avaliação formativa ser utilizadas na avaliação sumativa? Dito de outra forma: Será que deveremos ignorar todas as evidências de aprendizagem obtidas através das avaliações formativas, quando temos que fazer um juízo global acerca do que os alunos sabem e são capazes de fazer? Não há uma resposta linear para estas questões com alguns autores defendendo a ideia de que tal deverá mesmo acontecer (e.g., Fernandes, 2005a) e outros referindo que poderá acontecer desde que se tomem precauções várias. Estas precauções estão relacionadas com as questões de falta de validade e fiabilidade das avaliações formativas e com os contextos demasiado restritos e *situados* em que ocorrem (Harlen, 2006).

Ao considerar a avaliação formativa como a “(...) bússola orientadora do processo ensino-aprendizagem” Cortesão (1993, p. 13) sublinha o seu imprescindível e fundamental papel na regulação, no desenvolvimento e na melhoria daquele processo. Tal regulação tem que ser posta em prática pelos alunos e pelos professores e, por isso mesmo, se afirma que “(...) a avaliação formativa alternativa pressupõe uma partilha de responsabilidades entre alunos e professores em matéria de avaliação e de regulação das aprendizagens” (Fernandes, 2005a, p. 65).

Finalmente, após alguns exemplos de questões teóricas que decorrem da análise das conceptualizações feitas por alguns autores, destaca-se uma afirmação de Vieira e Moreira (1993) relativamente à questão da exequibilidade da avaliação formativa nas salas de aula: “(...) quaisquer que sejam as circunstâncias de ensino/aprendizagem (das orientações oficiais ao estilo de ensino dos professores ou ao número de alunos por turma), é sempre possível realizar uma avaliação formativa centrada nos processos de aprendizagem” (p. 10).

As abordagens utilizadas pelos autores à avaliação formativa são variadas e sugerem a seguinte agregação da informação analisada:

1. Livros que têm a natureza ou características de manuais em que a ênfase é mais na apresentação de conceitos e definições e na sua explicação e menos na sua discussão teórica (e.g., Damião, 1996; Domingos *et al.*, 1981; Lemos *et al.*, 1992; Ribeiro e Ribeiro, 1989; Ribeiro, 1989; Valadares e Graça, 1998; Vila Nova, 1997);
2. Livros em que se apresenta o conceito tal como aparece definido nos normativos legais e se discutem e/ou sugerem procedimentos de concretização prática (e.g., Lemos *et al.*, 1992; Machado, 1994);
3. Livros em que se desenvolvem implicações pedagógicas e didáticas associadas às práticas de avaliação formativa (e.g., Cortesão, 1993; Cortesão e Torres, 1981, 1983; Fernandes, 2005a; Vieira e Moreira, 1993);
4. Livros em que os autores fazem uma discussão mais ou menos profunda do conceito procurando enquadrá-lo conceptualmente com base nas teorias da aprendizagem, do currículo, da cognição, da comunicação ou outras (e.g., Alves, 2004; Fernandes, 2005a; Pacheco, 1994);
5. Livros em que os autores assumem a relevância política do investimento na avaliação formativa na promoção do sucesso escolar de todos os alunos (e.g., Cortesão, 1993; Cortesão e Torres, 1981, 1983; Baptista, 1999; Fernandes, 2005a).

Considerações e Reflexões Finais

A análise e a síntese da informação que integrou o *corpus* desta investigação suscita um alargado conjunto de considerações e de reflexões. Para efeitos deste artigo seleccionaram-se apenas as que se consideraram mais relevantes.

DA AUSÊNCIA DE UMA AGENDA DE INVESTIGAÇÃO

A maioria dos autores dos livros analisados são docentes e investigadores do ensino superior. Este facto poderia indiciar que os livros pudessem traduzir, implícita ou explicitamente, uma agenda de investigação no domínio da avaliação das aprendizagens. Dito de outra forma, poderia esperar-se que pelo menos parte dos livros emanasse do trabalho empírico desenvolvido pelos autores. Na verdade, são muito escassos os livros em que os autores fazem referência a investigação empírica da sua responsabilidade e autoria com o fim de fundamentar alguma ideia ou teoria. O facto da comunidade científica nesta área ser relativamente recente pode explicar parcialmente esta situação. Além disso, para vários autores, a investigação no domínio da avaliação das aprendizagens não será propriamente a primeira prioridade.

Em todo o caso, tal como se tem vindo a afirmar (Fernandes, 1994; 2004; 2005a), parece ser necessário definir uma agenda de investigação que possa orientar os esforços dos investigadores, tornando o seu trabalho mais útil, pertinente e oportuno. Talvez esta falta de agenda possa explicar o facto das iniciativas da administração educativa ou do Ministério da Educação constituírem a *raison d'être* de um número significativo de livros. Naturalmente que tais iniciativas devem ser sistematizadas, analisadas e objecto de interpretação, mas não deverão constituir o *motivo* ou o *pretexto* principal para que se produza um número tão significativo de publicações por parte de professores e investigadores do ensino superior.

DA ÊNFASE NA DESCRIÇÃO E NA PRESCRIÇÃO

Uma outra característica de uma parte substancial dos livros analisados é a sua natureza essencialmente descritiva e prescritiva. Ou seja, há sobretudo uma apresentação e descrição de um conjunto de factos, definições e, por vezes, de perspectivas teóricas que aparecem como um corpo acabado e completo de conhecimentos na área da avaliação das aprendizagens; uma apresentação *do que há* ou *do que está feito* sem que, verdadeiramente, se produza uma reflexão que mobilize e integre um conjunto de contribuições provenientes de domínios tais como a sociologia, a psicologia social, a antropologia, as teorias da comunicação, as teorias curriculares ou as teorias da aprendizagem. Sente-se que é necessária mais elaboração teórica e maior problematização que contribuam para transformar e melhorar as práticas de avaliação.

DOS MÉTODOS E DAS TÉCNICAS

Os métodos, técnicas ou tarefas de avaliação ocupam um lugar destacado num número significativo de livros. Em vários casos a maior parte dos capítulos é dedicada à *instrumentação*. Esta ênfase decorre da ideia da avaliação como medida, da necessidade de tornar a avaliação quantificável, e tem naturalmente a classificação dos alunos como preocupação subjacente. Dada a importância destas e doutras questões relacionadas, parece relevante que se discuta a questão do enquadramento epistemológico, mas também didáctico e pedagógico, dos chamados *instrumentos de avaliação*. Também aqui, na maioria dos livros, os *instrumentos* são simplesmente apresentados, indicando-se por vezes as suas vantagens e desvantagens, mas não são propriamente discutidas questões tais como o tipo de informação que proporcionam, a utilização que se pode fazer dessa informação ou a sua relação com as actividades da sala de aula.

DOS REFERENCIAIS UTILIZADOS PELOS AUTORES

Como também já se referiu a maioria das referências utilizadas pelos autores são de investigadores francófonos e anglo-saxónicos havendo, nalguns casos, referências a autores espanhóis. É natural e compreensível que assim seja e tal procedimento garante bons enquadramentos teóricos para as questões que se pretendem discutir. Mas parece positivo que se alargue o espectro de referências não ignorando outras tradições teóricas e práticas começando, por exemplo, por analisar mais trabalhos de autores portugueses.

Na verdade, ao nível da Europa central e do norte há importantes contribuições de autores holandeses (e.g., Kleijne e Schuring, 1993; Lange, 1987, 1993), alemães (e.g., Koller, 2005), dinamarqueses (e.g., Jensen, 1993; Townshend, Moos e Skov, 2005) e finlandeses (e.g., Voogt e Kasurinem, 2005). Nos países do sul da Europa há também uma variedade de contributos de autores espanhóis (e.g., Rico, 1993; Rosales, 1984, 2000; Santos Guerra, 1993, 2005), italianos (e.g., Bazzini, 1993; Looney, Laneve e Moscato, 2005) ou portugueses (e.g., Cortesão, 1993; Fernandes, 2005a; Leal e Abrantes, 1990; Leite e Fernandes, 2002; Pacheco, 1994; Ribeiro, 1989; Rosado e Colaço, 2002; Vilar, 1992).

Também têm emergido importantes contributos teóricos em países da América do Sul como é o caso do Brasil (Demo, 1996; Hoffmann, 2005; Luckesi, 1995; Sousa, 1995, 2003; Vasconcellos, 2005).

DA AVALIAÇÃO FORMATIVA E DA SUA TEORIA

No domínio mais específico da avaliação formativa há três aspectos que merecem ser referidos nesta altura e que se discutem sucintamente a seguir.

Na grande maioria dos livros analisados é dado relativamente pouco desenvolvimento ao tema da avaliação formativa e ao seu enquadramento sócio-educativo e pedagógico. A tendência é para definir o conceito, compará-lo com a avaliação sumativa e, por vezes, com

outras modalidades consideradas e pouco mais. Ou seja, fica-se com a sensação de que os autores procuram uma certa operacionalização do conceito que o torne *pronto a usar*. São escassos os livros em que se analisa com alguma profundidade as relações da avaliação formativa com outros elementos relevantes do processo educativo como é, por exemplo, o caso das aprendizagens, do *feedback*, da regulação, do papel dos professores ou da avaliação sumativa.

Apenas em três livros os autores se referem a investigação empírica realizada sob a sua responsabilidade na área da avaliação formativa. Esta questão merece ser reflectida pois parece relevante que a comunidade dos investigadores desta área apoie, cada vez mais, as suas reflexões e recomendações em evidências empíricas. Além disso, é substancialmente diferente apoiar recomendações para as práticas com base em deduções de natureza teórica ou em evidências provenientes das salas de aula.

O *feedback* e a regulação das aprendizagens são processos reconhecidamente indissociáveis do desenvolvimento de quaisquer práticas de avaliação formativa (e.g., Allal, 1986; Black e Wiliam, 1998, 2006a; Gipps e Stobart, 2003; Perrenoud, 1998a). Porém, são também escassos os livros em que é feita uma discussão que discuta aqueles dois conceitos e o seu lugar no processo de avaliação formativa.

O desenvolvimento da avaliação formativa nas salas de aula exige um significativo esforço de elaboração de natureza teórica que permita construir um quadro conceptual sólido. Sem este quadro dificilmente se poderão criar abordagens suficientemente sólidas que apoiem a melhoria das práticas.

Autores como Scriven (1994, 2000, 2003) têm referido que a dificuldade da consolidação de uma teoria da avaliação em geral radica em boa medida na área prática da avaliação das aprendizagens e, em particular, na clarificação de conceitos tais como *avaliação formativa* e *avaliação sumativa* ou *subjectivo* e *objectivo*. Outros autores (e.g., Black e Wiliam, 2006b, 2006c; Gipps, 1994; Stobart, 2006) têm vindo a contribuir para a construção teórica no domínio da avaliação formativa abordando questões que vão desde a qualidade (e.g., validade,

fiabilidade) até ao papel dos professores e alunos na regulação das aprendizagens.

A construção de uma teoria da avaliação formativa deverá basear-se em três elementos essenciais: a) a compreensão dos processos de desenvolvimento do currículo nas salas de aula e a sua relação com os processos de avaliação; b) a compreensão dos papéis de alunos e professores nos processos de ensino, aprendizagem e avaliação; e c) a compreensão dos contextos, dinâmicas e ambientes de ensino, aprendizagem e avaliação nas salas de aula (Fernandes, 2006).

Esta síntese interpretativa mostrou que muito há ainda a fazer no domínio da avaliação das aprendizagens e, muito particularmente, no que se refere à avaliação formativa. Mas também revelou que há um caminho que foi percorrido ao longo de 25 anos que não devemos ignorar. O que se procurou fazer nesta investigação foi dar-lhe sentido e significado para que se possam prosseguir os esforços já realizados e contribuir para que a *avaliação para aprender* seja uma realidade cada vez mais presente nas escolas e nas salas de aula. Mas tudo isto é indissociável do desenvolvimento de uma agenda forte no domínio da investigação empírica que possa abrir caminho à construção teórica e à transformação e melhoria das práticas.

Referências Bibliográficas

- ABRANTES, P. e ARAÚJO, F. (orgs.) (2002). *Avaliação das aprendizagens: Das concepções às práticas*. Lisboa: Departamento de Educação Básica do Ministério da Educação.
- ABRECHT, R. (1991). *L'évaluation formative: Une analyse critique*. Bruxelles: De Boeck.
- ALLAL, L. (1979). Stratégies d'évaluation formative: Conceptions psychopédagogiques et modalités d'application. In Allal, L.; Cardinet J. e Perrenoud, Ph. (eds.). *L'évaluation formative dans un enseignement différencié*. (Actes du colloque à l' Université de Genève, Mars 1978.). Berne: Peter Lang, pp. 153-183.
- ALLAL, L. (1986). Estratégias de avaliação formativa: Concepções psicopedagógicas e modalidades de aplicação. In Allal, L.; Cardinet, J. e Perrenoud, Ph. (orgs.). *A avaliação formativa num ensino diferenciado*. Coimbra: Almedina, pp. 175-209.

- ALLAL, L.; CARDINET, J. e PERRENOUD, PH. (eds.) (1979). *L'évaluation formative dans un enseignement différencié*. (Actes du colloque à l' Université de Genève, Mars 1978). Berne: Peter Lang.
- ALLAL, L., CARDINET, J. e PERRENOUD, PH. (orgs.) (1986). *A avaliação formativa num ensino diferenciado*. (Actas do colóquio realizado na Universidade de Genebra, Março 1978). Coimbra: Almedina.
- ALLAL, L. e LOPEZ, L. (2005). Formative assessment of learning: A review of publications in French. In OECD (ed.). *Formative assessment: Improving learning in secondary schools*. Paris: OECD publishing, pp. 241-264.
- ALVES, M. P. (2004). *Currículo e avaliação: Uma perspectiva integrada*. Porto: Porto Editora.
- ALVES, F. e FORMOSINHO, J. (1992). *Contributos para uma outra prática educativa: O novo sistema de avaliação dos alunos do ensino básico*. Porto: ASA.
- BAPTISTA, J. A. (1999). *Aprender por medida*. Porto: Porto Editora.
- BARREIRA, C. e PINTO, J. (2006). *A investigação em Portugal sobre avaliação das aprendizagens dos alunos (1990-2005)*. Disponível em: <http://www.educ.fc.ul.pt/docentes/msantos/spce.pdf> (11 - 2006).
- BAZZINI, L. (1993). The teaching/learning process and assessment practice: Two intertwined sides of mathematics education. In Niss, M. (ed.). *Cases of assessment in mathematics education: An ICMI study*. Dordrecht: Kluwer, pp. 99-106.
- BENTO, J. (1989). *Planeamento e avaliação em Educação Física*. Lisboa: Horizonte.
- BERLAK, H. (1992a). The need for a new science of assessment. In Berlak, H.; Newmann, F.; Adams, E.; Archbald, D.; Burgess, T.; Raven, J. e Romberg, T. (eds.). *Toward a new science of educational testing and assessment*. Albany, NY: State University of New York Press, pp. 1-22.
- BERLAK, H. (1992b). Toward the development of a new science of educational testing and assessment. In Berlak, H.; Newmann, F.; Adams, E.; Archbald, D.; Burgess, T.; Raven, J. e Romberg, T. (eds.). *Toward a new science of educational testing and assessment*. Albany, NY: State University of New York Press, pp. 181-206.
- BLACK, P. e WILLIAM, D. (1998). Assessment and classroom learning. *Assessment in Education: Principles, Policy & Practice*, 5, 1, pp. 7-74.
- BLACK, P. e WILLIAM, D. (2006a). Assessment for learning in the classroom. In Gardner, J. (ed.). *Assessment and learning*. London: Sage, pp. 9-25.
- BLACK, P. e WILLIAM, D. (2006b). Developing a theory of formative assessment. In Gardner, J. (ed.). *Assessment and learning*. London: Sage, pp. 81-100.
- BLACK, P. e WILLIAM, D. (2006c). The reliability of assessments. In Gardner, J. (ed.). *Assessment and learning*. London: Sage, pp. 119-131.
- BLOOM, B. (1956). *Taxonomy of educational objectives*. New York: McKay.

- BLOOM, B. e KRATHWOHL, D. (1956). *Taxonomy of educational objectives: The classification of educational goals by a committee of college and university examiners. Handbook I: Cognitive Domain*. New York: Longmans, Green.
- BLOOM, B.; HASTINGS, J. e MADAUS, G. (1971). *Handbook of formative and summative evaluation of student learning*. New York: Mac Graw Hill.
- BONNIOL, J-J. (1989). Sur les regulations du fonctionnement cognitif de l'élève: Contribution à une theorie de l'évaluation formative. *Atelier de recherche sur l'évaluation des résultats scolaires: Motivations et réussite des élèves* (Liège, 12-15 Septembre 1989). Strasbourg: Conseil de L'Europe.
- BONNIOL, J-J. e VIAL, M. (1997). *Les modèles de l'évaluation: Textes fondateurs avec commentaires*. Bruxelles: De Boeck.
- CABRAL, N. (2003). *Avaliação no ensino básico*. Porto: Porto Editora.
- CARDINET, J. (1986). *Évaluation scolaire et mesure*. Bruxelles: De Boeck.
- CARDINET, J. (1991). L'apport sociocognitif à la régulation interactive. In Weiss, J. (ed.). *L'évaluation: Problème de communication*. Cousset (Fribourg): Delval, pp. 199-213.
- COELHO, C. e CAMPOS, J. (2003). *O portfolio na sala de aula*. Porto: Areal.
- CONSELHO NACIONAL DE EDUCAÇÃO (1995). *A avaliação dos alunos da educação básica e do ensino secundário*. Lisboa: Autor.
- COOPER, H. (1988). Organizing knowledge syntheses: A taxonomy of literature reviews. *Knowledge in Society*, 1, pp. 104-126.
- COOPER, H. (1998). *Synthesising research: A guide for literature reviews* (3rd ed.). London: Sage.
- COOPER, H. (2003). Editorial. *Psychological Bulletin*, 129, 1, pp. 3-9.
- COOPER, H. e LINDSAY, J. (1998). Research synthesis and meta-analysis. In Bickman, L. e Rog, D. (eds.). *Handbook of social research methods*. London: Sage, pp. 315-337.
- CORTESÃO, L. (1993). *A avaliação formativa: Que desafios* Porto: ASA.
- CORTESÃO, L. e TORRES, M. A. (1981). *Avaliação pedagógica I: Insucesso escolar*. Porto: Porto Editora.
- CORTESÃO, L. e TORRES, M. A. (1983). *Avaliação pedagógica II: Perspectivas de sucesso*. Porto: Porto Editora.
- CORTESÃO, L. e TORRES, M. A. (1996). *Avaliação pedagógica: Mudança na escola, mudança na avaliação*. Porto: Porto Editora.
- COSTA, A. (1989). *Tipos de avaliação e de instrumentos utilizados em educação física*. Lisboa: Faculdade de Motricidade Humana.
- CROOKS, T. (1988). The impact of classroom evaluation practices on students. *Review of Educational Research*, 58, pp. 438-481.
- CUNHA, A. (2004). *Reforma do ensino secundário – currículo e avaliações: Síntese e guia prático*. Porto: ASA.
- DAMAS, M. e DE KETELE, J.M. (1985). *Observar para avaliar*. Coimbra: Almedina.

- DAMIAO, M. H. (1996). *Pré, inter e pós acção: Planificação e avaliação em pedagogia*. Coimbra: Minerva.
- DE KETELE, J-M. (1986). A propósito das noções de avaliação formativa, de avaliação sumativa, de individualização e de diferenciação. In Allal, L., Cardinet, J. e Perrenoud, Ph. (orgs.). *A avaliação formativa num ensino diferenciado*. Coimbra: Almedina, pp. 211-218.
- DE KETELE, J-M. (1993). L'évaluation conjuguee en paradigmes. *Revue Française de Pédagogie*, 103, pp. 59-80.
- DE KETELE, J-M. (2001). Évolution des problématiques issues de l'évaluation formative. In Figari, G. e Achouche, M. (eds.). *L'activité évaluative réinterrogée: Regards scolaires et socioprofessionnels*. Bruxelles: De Boeck, pp. 102-108.
- DEMO, P. (1996). *Avaliação sob o olhar propedêutico*. São Paulo: Papirus.
- DOMINGOS, A. M.; NEVES, I. e GALHARDO, L. (1981). *Uma forma de estruturar o ensino e a aprendizagem*. Lisboa: Horizonte.
- EARL, L. (2003). *Assessment as learning: Using classroom assessment to maximize student learning*. Thousand Oaks, CA: Corwin Press.
- ESTRELA, A. e NÓVOA, A. (orgs.) (1992). *Avaliações em educação: Novas perspectivas*. Lisboa: Educa.
- FERNANDES, D. (1992). O tempo da avaliação. *Noesis*, 23, pp. 18-21.
- FERNANDES, D. (1994). Avaliação das aprendizagens: Das prioridades de investigação e de formação às práticas na sala de aula. *Revista de Educação*, 8, pp. 15-20.
- FERNANDES, D. (2004). *Avaliação das aprendizagens: Uma agenda. muitos desafios*. Cacém: Texto Editores.
- FERNANDES, D. (2005a). *Avaliação das aprendizagens: Desafios às teorias, práticas e políticas*. Cacém: Texto Editores.
- FERNANDES, D. (2005b). Para uma ênfase na avaliação formativa alternativa. Editorial. *Educação e Matemática*, 81, p.1.
- FERNANDES, D. (2006). Para uma teoria da avaliação formativa. *Revista Portuguesa de Educação*, 19, 2, pp. 21-50.
- FERNANDES, D. (coord.) (s/d). *Pensar a avaliação. Melhorar a aprendizagem*. Lisboa: Instituto de Inovação Educacional.
- FERNANDES, D.; RAMALHO, G. e LEMOS, V. (1991). *Opiniões dos professores dos ensinos básico e secundário relativamente às medidas constantes no projecto do sistema de avaliação dos alunos*. Lisboa: Instituto de Inovação Educacional.
- FERREIRA, C. (2004). *Avaliação formativa: Conceptualização e orientações para a prática*. Vila Real: Universidade de Trás-os-montes e Alto Douro.
- FIGARI, G. (1996). *Avaliar: Que referencial?* Porto: Porto Editora.
- FIGARI, G. (2001). Us et abus de la notion de référentiel. In Figari, G. e Achouche, M. (eds.). *L'activité évaluative réinterrogée*. Bruxelles: De Boeck, pp. 310-314.

- GARDNER, J. (2006a). Assessment and learning: An Introduction. In Gardner, J. (ed.). *Assessment and learning*. London: Sage, pp. 1-5.
- GARDNER, J. (2006b). Assessment for learning: A compelling conceptualization. In Gardner, J. (ed.). *Assessment and learning*. London: Sage, pp. 197-204.
- GIFFORD, B. e O' CONNOR (eds.) (1992). *Changing assessments: Alternative views of aptitude, achievement and instruction*. Dordrecht: Kluwer.
- GIPPS, C. (1994). *Beyond testing: Towards a theory of educational assessment*. Londres: Falmer.
- GIPPS, C. e STOBART, G. (2003). Alternative assessment. In Kellaghan, T. e Stufflebeam, D. (eds.). *International handbook of educational evaluation*. Dordrecht: Kluwer, pp. 549-576.
- GLASS, G. (1976). Primary, secondary and meta-analysis research. *Educational Researcher*, 5, pp. 3-8.
- GRÉGOIRE, J. (ed.) (1996). *Évaluer des apprentissages: Les apports de la psychologie cognitive*. Bruxelles: De Boeck.
- GRONLUND, N. (1970). *Stating behavioral objectives for classroom instruction*. New York: Macmillan.
- GRONLUND, N. e LINN, R. (1990). *Measurement and evaluation in teaching*. New York: MacMillan.
- GUIMARÃES, H., LEAL, L. e ABRANTES, P. (orgs.) (1991). *Avaliação: Uma questão a enfrentar*. Actas do seminário sobre avaliação. Lisboa: Associação de Professores de Matemática.
- HADJI, C. (1992). *L'évaluation des actions éducatives*. Paris: PUF.
- HARLEN, W. (2005). Teachers' summative practices and assessment for learning: Tensinos and synergies. *Curriculum Journal*, 16, 2, pp. 207-223.
- HARLEN, W. (2006). On the relationship between assessment for formative and summative purposes. In Gardner, J. (ed.). *Assessment and learning*. London: Sage, pp. 103-118.
- HARLEN, W. e JAMES, M. (1997). Assessment and learning: Differences and relationships between formative and summative assessment. *Assessment in education: Principles, policy and practice*, 4, 3, pp. 365-379.
- HARROW, A. (1972). *A Taxonomy of the psychomotor domain. A guide for developing behavioral objectives*. New York: McKay.
- HOFFMANN, J. (2005). *O jogo do contrário em avaliação*. Porto Alegre: Mediação.
- JENSEN, H. (1993). Assessment of primary and lower secondary mathematics in Denmark. In Niss, M. (ed.). *Cases of assessment in mathematics education: An ICMI study*. Dordrecht: Kluwer, pp. 119-128.
- JORRO, A. (2000). *L'enseignant et l'évaluation: Des gestes évaluatifs en question*. Bruxelles: De Boeck.

- KLEIJNE, W. e SCHURING, H. (1993). Assessment of examinations in the Netherlands. In Niss, M. (ed.). *Cases of assessment in mathematics education: An ICMI study*. Dordrecht: Kluwer, pp. 139-154.
- KOLLER, O. (2005). Formative assessment in classrooms: A review of the empirical German literature. In OECD (ed.). *Formative assessment: Improving learning in secondary schools*. Paris: OECD publishing, pp. 265-275.
- KRATHWOHL, D.; BLOOM, B. e MASIA, B. (1964). *Taxonomy of Educational Objectives: The Classification of Educational Goals. Handbook II: Affective Domain*. New York: McKay.
- LANGE, J. de (1987). *Mathematics, insight and meaning*. Utrecht, Netherlands: OweOC.
- LANGE, J. de (1993). Assessment in problem-oriented curricula. In Webb, N. e Coxford, A. (eds.). *Assessment in the mathematics classroom*. Reston, VA: NCTM, pp. 197-208.
- LEAL, L. e ABRANTES, P. (1990). Avaliação da aprendizagem/Avaliação na aprendizagem. *Inovação*, 3, 4, pp. 65-76.
- LEITE, C. (org.) (1993). *Avaliar a avaliação*. Porto: ASA.
- LEITE, C. e FERNANDES, P. (2002). *Avaliação das aprendizagens dos alunos: Novos contextos, novas práticas*. Porto: ASA.
- LEMO, V.; NEVES, A.; CAMPOS, C.; CONCEIÇÃO, J. e ALAIZ, V. (1992). *A nova avaliação da aprendizagem: O direito ao sucesso*. Cacém: Texto.
- LEMO, V. (1986). *O critério do sucesso: Técnicas de avaliação da aprendizagem*. Cacém: Texto.
- LOBO, A. (1998). *A.A.A. (Aprendizagem assistida pela avaliação): Um sorriso difícil para o novo sistema de avaliação*. Porto: Porto Editora.
- LOONEY, J.; LANEVE, C. e MOSCATO, M. (2005). A system in transition. In OECD (ed.). *Formative assessment: Improving learning in secondary schools*. Paris: OECD publishing, pp. 163-176.
- LUCKESI, C. (1995). *Avaliação da aprendizagem escolar*. São Paulo: Cortez.
- MACHADO, F. (1994). *Avaliação em tempos de mudança: Projectos e práticas nos ensinos básico e secundário*. Porto: ASA.
- MIRANDA, F. e BERNARDES, C. (2003). *Portefólio: Uma escola de competências*. Porto: Porto Editora.
- NATRIELLO, G. (1987). The impact of evaluation processes on students. *Educational Psychologist*, 22, pp. 155-175.
- NEVES, A.; JORDÃO, A. e SANTOS, L. (2004). Investigação publicada em Portugal, entre 1971 e 2003, acerca da avaliação das aprendizagens: Seu levantamento e análise. *Revista Portuguesa de Investigação Educacional*, 3, pp. 47-71.
- NOBLIT, G. e HARE, R. (1988). *Meta-ethnography: Synthesising qualitative studies*. London: Sage.

- NUNZIATI, G. (1990). Pour construire un dispositif d'évaluation formatrice. *Cahiers Pédagogiques*, 280, 47-64.
- NUTTALL, D. (Ed.) (1986). *Assessing educational achievement*. London: Falmer.
- OLIVEIRA, I.; PEREIRA, J. e FERNANDES, D. (1993). *Desenvolvimento de instrumentos de avaliação da aprendizagem em Matemática*. Lisboa: Instituto de Inovação Educacional.
- PACHECO, J. (1994). *A avaliação dos alunos na perspectiva da reforma: Propostas de trabalho*. Porto: Porto Editora.
- PACHECO, J. e ZABALZA, M. (orgs.) (1995). *A avaliação dos alunos dos ensinos básico e secundário*. Actas do 1.º colóquio sobre questões curriculares. Braga: Instituto de Educação e Psicologia da Universidade do Minho.
- PAIS, A. e MONTEIRO, M. (1996). *Avaliação: Uma prática diária*. Lisboa: Presença
- PERRENOUD, PH. (1998a). From formative evaluation to a controlled regulation of learning processes: Towards a wider conceptual field. *Assessment in Education: Principles, Policy & Practice*, 5, 1, 85-102.
- PERRENOUD, PH. (1998b). *L'évaluation des élèves. De la fabrication de l'excellence à la régulation des apprentissages*. Bruxelles: De Boeck.
- PHYE, G. (ed.) (1997). *Handbook of classroom assessment: Learning, adjustment, and achievement*. New York: Academic Press.
- RESNICK, L. (1987). *Education and learning to think*. Washington, DC: National Academy Press.
- RIBEIRO, L. (1989). *Avaliação da aprendizagem*. Cacém: Texto.
- RIBEIRO, A. e RIBEIRO, L. (1989). *Planificação e avaliação do ensino-aprendizagem*. Lisboa: Universidade Aberta.
- RICO, L. (1993). Mathematics assessment in the Spanish education system. In Niss, M. (ed.). *Cases of assessment in mathematics education: An ICMI study*. Dordrecht: Kluwer, pp. 9-20.
- ROSADO, A e COLAÇO, C. (orgs.) (2002). *Avaliação das aprendizagens: Fundamentos e aplicações no domínio das actividades físicas*. Lisboa: Omniserviços.
- ROSALES, C. (1984). *Criterios para una evaluación formativa (Segunda edición)*. Madrid: Narcea.
- ROSALES, C. (2000). *Evaluar es reflexionar sobre la enseñanza (Tercera edición)*. Madrid: Narcea.
- SADLER, D. (1998). Formative assessment: Revisiting the territory. *Assessment in Education: Principles, Policy & Practice*, 5, 1, 77-84.
- SANTOS GUERRA, M. (1995). *La evaluación: Un proceso de diálogo, comprensión y mejora*. Málaga: Aljibe.
- SANTOS GUERRA, M. (2003). *Una flecha en la diana: La evaluación como aprendizaje*. Madrid: Narcea.

- SCALLON, G. (2004). *L'évaluation des apprentissages dans une approche par compétences*. Bruxelles: De Boeck.
- SCRIVEN, M. (1967). *The methodology of evaluation* (AERA Monograph series on curriculum evaluation, No. 1). Chicago IL: Rand McNally.
- SCRIVEN, M. (1994). Evaluation as a discipline. *Studies in Educational Evaluation*, 20, pp. 147-166.
- SCRIVEN, M. (2000). Evaluation ideologies. In Stufflebeam, D.; Madaus, G. e Kellaghan, T. (eds.). *Evaluation models: Viewpoints on educational and human services evaluation* (Segunda edição). Boston: Kluwer, pp. 249-278.
- SCRIVEN, M. (2003). Evaluation theory and metatheory. In Kellaghan, T. e Stufflebeam, D. (eds.). *International handbook of educational evaluation*. Boston: Kluwer, pp. 15-30.
- SEBBA, J. (2006). Policy and practice in assessment for learning: The experience of selected OECD countries. In Gardner, J. (ed.). *Assessment and learning*. London: Sage, pp. 185-196.
- SHEPARD, L. (2000). The role of assessment in a learning culture. *Educational Researcher*, 29, 7, pp. 4-14.
- SHEPARD, L. (2001). The role of classroom assessment in teaching and learning. In Richardson, V. (ed.). *Handbook of research on teaching* (4th Edition). American Educational Research Association. New York: Macmillan.
- SILVA, O. (1989). *História da avaliação em educação física*. Lisboa: Faculdade de Motricidade Humana.
- SIM-SIM, I. (1997). *Avaliação da linguagem oral: Um contributo para o conhecimento do desenvolvimento linguístico das crianças portuguesas*. Lisboa: Gulbenkian.
- SLAVIN, R. (1984). Meta-analysis in education. How has it been used? *Educational Researcher*, 13, 8, pp. 6-15, 24-27.
- SLAVIN, R. (1986). Best-evidence synthesis: An alternative to meta-analysis and traditional reviews. *Educational Researcher*, 15, 9, pp. 5-11.
- SLAVIN, R. (2004). Education research can and must address “what works” questions. *Educational Researcher*, 33, 1, pp. 27-28.
- SOBRAL, F. e BARREIROS, M. L. (1980). *Fundamentos e técnicas de avaliação em educação física*. Lisboa: CDI, Instituto Superior de Educação Física.
- SOUSA, S. (1995). Avaliação da aprendizagem nas pesquisas no Brasil de 1930 a 1980. *Cadernos de Pesquisa – Fundação Carlos Chagas*, 94, pp. 43-49.
- SOUSA, S. (2003). Possíveis impactos das políticas de avaliação no currículo escolar. *Cadernos de Pesquisa – Fundação Carlos Chagas*, 119, pp. 175-190.
- STAKE, R. (2006). *Evaluación comprensiva y evaluación basada em estándares*. Barcelona: Graó.
- STIGGINS, R. e CONKLIN, N. (1992). *In teachers' hands: Investigating the practices of classroom assessment*. Albany, NY: State University of New York Press.

- STOBART, G. (2006). The validity of formative assessment. In Gardner, J. (ed.). *Assessment and learning*. London: Sage, pp. 133-146.
- SURI, H. (1999a). *A methodologically inclusive model for research synthesis*. Disponível em: <http://www.aare.edu.au/99pap/sur99673.htm> (09-2006).
- SURI, H. (1999b). *The process of synthesising qualitative research: A case study*. Disponível em: <http://www.latrobe.edu.au/aqr/offer/papers/HSuri.htm> (09-2006).
- SURI, H. (2002). *Essential features of methodologically inclusive research synthesis*. Retirado em 11 de Setembro, 2006, da base de dados ERIC.
- SURI, H. e CLARKE, D. (1999). *Revisiting methods of literature synthesis*. Retirado em 11 de Setembro, 2006, da base de dados ERIC.
- TORRANCE, H. e PRYOR, J. (2001). Developing formative assessment in the classroom: Using action research to explore and modify theory. *British Educational Research Journal*, 27, 5, pp. 615-631.
- TOWNSHEND, J.; MOOS, L. e SKOV, P. (2005). Building on a tradition of democracy and dialogue in schools. In OECD (ed.). *Formative assessment: Improving learning in secondary schools*. Paris: OECD publishing, pp. 117-128.
- TUNSTALL, P. e GIPPS, C. (1996). Teacher *feedback* to young children in formative assessment: A typology. *British Educational Research Journal*, 22, pp. 389-404.
- VALADA, R. (1993). *Guia da avaliação no ensino básico*. Mem Martins: Inquérito.
- VALADARES, J. e GRAÇA, M. (1998). *Avaliando para melhorar a aprendizagem*. Lisboa: Plátano.
- VASCONCELLOS, C. (2005). *Avaliação da aprendizagem: Práticas de mudança por uma práxis transformadora* (7.ª Edição). São Paulo: Libertad.
- VERÍSSIMO, A. (2000). *Registos de observação na avaliação do rendimento escolar*. Porto: Areal.
- VIEIRA, F. e MOREIRA, M. A. (1993). Para além dos testes... A avaliação processual na aula de inglês. Braga: Universidade do Minho.
- VILA NOVA, E. (1997). *Avaliação dos alunos: Problemas e soluções*. Cacém: Texto.
- VILAR, A. (1992). *A avaliação: Um novo discurso?* Porto: ASA.
- VILAR, A. (1993). *A avaliação dos alunos no ensino básico*. Porto: ASA.
- VILHENA, T. (1999). *Avaliar o extracurricular*. Porto: ASA.
- VOOGT, J. e KASURINEM, H. (2005). Emphasising development instead of competition and comparison. In OECD (ed.). *Formative assessment: Improving learning in secondary schools*. Paris: OECD publishing, pp. 149-162.
- WEED, M. (2005). "Meta-Interpretation": A method for the interpretive synthesis of qualitative research. *Forum Qualitative Sozialforschung/Forum Qualitative Social Research* (On-line Journal), 6, 1, Artigo 37. Disponível em: <http://www.qualitative-research.net/fqs-texte/1-05/05-1-37-e.htm> (09-2006).

WEISS, J. (dir.) (1991). *L'évaluation: Problème de communication*. Cousset (Fribourg): Delval.

WIGGINS, G. (1998). *Educative assessment*. San Francisco: Jossey-Bass.

Referências Legislativas

Lei n.º 46/86 de 14 de Outubro. *Diário da República* n.º 237, I Série.

Decreto-Lei n.º 286/89 de 29 de Agosto. *Diário da República* n.º 198, I Série.

Despacho Normativo n.º 98-A/92 de 20 de Junho. *Diário da República* n.º 140, I Série B.

Despacho Normativo n.º 338/93 de 21 de Outubro. *Diário da República* n.º 247, I Série B.

Anexo A – Lista com exemplos de livros não analisados

AUTORES	DATA	TÍTULO	EDITORA
Fernandes, D. (coord.)	s/D	Pensar a Avaliação, Melhorar a Aprendizagem	IIE
Sobral, F. e Barreiros, M. L.	1980	Fundamentos e Técnicas de Avaliação em Educação Física	ISEF
Branco, I. e Gomes, S.	1991	Análise de pareceres de associações e instituições de educação sobre o projecto do sistema de avaliação dos alunos	IIE
Guimarães H., Leal, L. e Abrantes, P.	1991	Actas do seminário sobre avaliação. Avaliação: Uma questão a enfrentar	APM
Fernandes, D., Ramalho, G. e Lemos, V.	1991	Opiniões dos professores dos ensinos básico e secundário relativamente às medidas constantes no projecto do sistema de avaliação dos alunos	IIE
IIE	1992	Avaliar é aprender: O novo sistema de avaliação	IIE
Estrela, A. e Falcão, M. E.	1992	Actas do II Colóquio da Secção Portuguesa da AFIRSE	AFIRSE
Castanheira, C.	1992	Práticas de elaboração, análise e interpretação de testes escolares	GEP-ME
Cristo, F. e Bentes, M. C.	1993	Avaliação da aprendizagem em Língua Portuguesa no contexto dos novos programas: 6º ano de escolaridade	IIE
Oliveira, I., Pereira, J. e Fernandes, D.	1993	Desenvolvimento de instrumentos de avaliação da aprendizagem em Matemática	IIE
Valada, R.	1993	Guia da avaliação no ensino básico	INQUÉRITO
Oliveira, I., Pereira, J. e Fernandes, D.	1994	Seis propostas de avaliação: Matemática	IIE
Estrela, A. e Rodrigues, P. (orgs.)	1994	Para uma fundamentação da avaliação em educação	COLIBRI
Pacheco, J. e Zabalza, M. (orgs.)	1995	A avaliação dos alunos dos ensinos básico e secundário	UNIV. MINHO

Anexo A – Lista com exemplos de livros não analisados (*continuação*)

AUTORES	DATA	TÍTULO	EDITORA
Conselho Nacional de Educação (CNE)	1995	A avaliação dos alunos da educação básica e do ensino secundário	CNE
Castro, J. e Maia, J.	1996	A divergência entre a avaliação contínua e os exames nacionais de Matemática	IIE
Alaiz, V., Gonçalves, C. e Barbosa, J.	1997	Implementação do modelo de avaliação no ensino básico	IIE
Ministério da Educação	2000	Provas de aferição do ensino básico: 4.º ano – 2000	ME
GAVE	2001	Resultados do Estudo Internacional PISA 2000	GAVE
Abrantes, P. e Araújo, F. (orgs.)	2002	Avaliação das aprendizagens: Das concepções às práticas	DEB-ME
GAVE	2002	Conceitos fundamentais em jogo na avaliação da literacia Matemática e competências dos alunos portugueses – PISA 2000	GAVE
GAVE	2002	Contributo para uma melhor compreensão do desempenho dos alunos nos exames do 12.º ano	GAVE
Ministério da Educação	2002	Provas de aferição do ensino básico: 4.º e 6.º anos – 2001	ME
GAVE	2003	Conceitos fundamentais em jogo na avaliação da literacia científica e competências dos alunos portugueses – PISA 2000	GAVE
Ferreira, C.	2004	Avaliação formativa: conceptualização e orientações para a prática	UTAD
GAVE	2004	Supervisão da classificação das provas de exame de Biologia, Matemática e Química: Relatório técnico e de avaliação (2003)	GAVE
Ministério da Educação	2004	Provas de aferição do ensino básico: 4.º, 6.º e 9.º anos – 2003	ME
Ministério da Educação	2004	Provas de aferição do ensino básico: 4.º, 6.º e 9.º anos: Análise comparativa (2001-2003)	ME

Anexo B – Lista dos livros analisados

AUTORES	DATA	TÍTULO	EDITORA
Domingos, A. M., Neves, I. e Galhardo, L.	1981	Uma forma de estruturar o ensino e a aprendizagem	HORIZONTE
Cortesão, L. e Torres, M. A.	1981	Avaliação pedagógica I: Insucesso escolar	PORTO ED.
Cortesão, L. e Torres, M. A.	1983	Avaliação pedagógica II: Perspectivas de sucesso	PORTO ED.
Damas, M. e De Ketele, J..	1985	Observar para avaliar	ALMEDINA
Lemos, V.	1986	O critério do sucesso: Técnicas de avaliação da aprendizagem	TEXTO
Bento, J.	1987	Planeamento e avaliação em Educação Física	HORIZONTE
Ribeiro, A. e Ribeiro, L.	1989	Planificação e avaliação do ensino-aprendizagem	UNIV. ABERTA
Ribeiro, L.	1989	Avaliação da aprendizagem	TEXTO
Lemos, V., Neves, A., Campos, C., Conceição, J. e Alaiz	1992	A nova avaliação da aprendizagem: O direito ao sucesso	TEXTO
Alves, F. e Formosinho, J.	1992	Contributos para uma outra prática educativa: O novo sistema de avaliação dos alunos do ensino básico	ASA
Vilar, A.	1992	A avaliação: Um novo discurso?	ASA
Vieira, F. e Moreira, M. A.	1993	Para além dos testes... A avaliação processual na aula de inglês	UNIV. MINHO
Leite, C. (org.)	1993	Avaliar a avaliação	ASA
Cortesão, L.	1993	A avaliação formativa: Que desafios?	ASA
Vilar, A.	1993	A avaliação dos alunos no ensino básico	ASA
Machado, F.	1994	Avaliação em tempos de mudança: Projectos e práticas nos ensinos básico e secundário	ASA
Pacheco, J.	1994	A avaliação dos alunos na perspectiva da reforma: Propostas de trabalho	PORTO ED.

Anexo B – Lista dos livros analisados (*continuação*)

AUTORES	DATA	TÍTULO	EDITORA
Cortese, L. e Torres, M. A.	1996	Avaliação pedagógica: Mudança na escola, mudança na avaliação	PORTO ED.
Pais, A. e Monteiro, M.	1996	Avaliação: Uma prática diária	PRESENÇA
Damião, M. H.	1996	Pré, inter e pós acção: Planificação e avaliação em pedagogia	MINERVA
Vila Nova, E.	1997	Avaliação dos alunos: Problemas e soluções	TEXTO
Sim-Sim, I.	1997	Avaliação da linguagem oral: Um contributo para o conhecimento do desenvolvimento linguístico das crianças portuguesas	FUNDAÇÃO CALOUSTE GULBENKIAN
Lobo, A.	1998	A.A.A. (Aprendizagem assistida pela avaliação): Um sorriso difícil para o novo sistema de avaliação	PORTO ED.
Valadares, J. e Graça, M.	1998	Avaliando para melhorar a aprendizagem	PLÁTANO
Vilhena, T.	1999	Avaliar o extracurricular	ASA
Baptista, J. A.	1999	Aprender por medida	PORTO ED.
Veríssimo, A.	2000	Registos de observação na avaliação do rendimento escolar	AREAL
Leite, C. e Fernandes, P.	2002	Avaliação das aprendizagens dos alunos: Novos contextos, novas práticas	ASA
Rosado, A e Colaço, C. (orgs.)	2002	Avaliação das aprendizagens: Fundamentos e aplicações no domínio das actividades físicas	OMNISERVIÇOS
Coelho, C. e Campos, J.	2003	O portfolio na sala de aula	AREAL
Cabral, N.	2003	Avaliação no ensino básico	PORTO ED.
Miranda, F. e Bernardes, C.	2003	Portefólio: Uma escola de competências	PORTO ED.
Alves, M. P.	2004	Currículo e avaliação: Uma perspectiva integrada	PORTO ED.
Fernandes, D.	2005	Avaliação das aprendizagens: Desafios às teorias, práticas e políticas	TEXTO