

UNIVERSIDADE DE LISBOA

FACULDADE DE PSICOLOGIA  
E CIÊNCIAS DA EDUCAÇÃO



**DEZ ANOS DE INVESTIGAÇÃO EM AVALIAÇÃO DAS APRENDIZAGENS:  
REFLEXÕES A PARTIR DA ANÁLISE DE DISSERTAÇÕES DE MESTRADO.**

**CARLA MARIA LOPES MARTINS**

MESTRADO EM CIÊNCIAS DA EDUCAÇÃO  
Área de Especialização: Avaliação em Educação

Ano 2008

UNIVERSIDADE DE LISBOA

FACULDADE DE PSICOLOGIA  
E CIÊNCIAS DA EDUCAÇÃO



**DEZ ANOS DE INVESTIGAÇÃO EM AVALIAÇÃO DAS APRENDIZAGENS:  
REFLEXÕES A PARTIR DA ANÁLISE DE DISSERTAÇÕES DE MESTRADO.**

**CARLA MARIA LOPES MARTINS**

MESTRADO EM CIÊNCIAS DA EDUCAÇÃO  
Área de Especialização: Avaliação em Educação

Dissertação orientada pelo Prof. Doutor **DOMINGOS FERNANDES**

Ano 2008

## **Resumo**

A partir dos anos 1990, assiste-se, a nível internacional, a uma significativa produção de trabalhos de reflexão e de investigação centrados na avaliação das aprendizagens, fruto de uma tomada de consciência da sua importância na melhoria das aprendizagens dos alunos. No nosso país, a importância crescente da avaliação das aprendizagens verificou-se, essencialmente, a três níveis: políticas educativas, produção de materiais e investigação.

No presente estudo fez-se uma caracterização de dissertações de mestrado produzidas em Portugal entre 1993/94 e 2002/2003, sobretudo ao nível dos contextos em que se realizaram, dos problemas e questões de investigação, bem como das estratégias e dos procedimentos metodológicos utilizados.

A nível metodológico, seguiram-se os processos de uma síntese de investigação, na linha de Suri e Clarke (1999), utilizando-se critérios de inclusão/exclusão de carácter conceptual e não valorativos, com recurso a categorias de análise, leitura focada e sínteses.

Os resultados desta investigação evidenciam que a maioria dos estudos tiveram como objectivo o estudo das concepções e práticas de avaliação de professores ou apenas as práticas de avaliação. No entanto, a principal fonte de informação foi baseada em entrevistas a professores, com pouco recurso à observação de aulas. Desta forma, na grande maioria dos estudos as práticas de avaliação dos professores não foram efectivamente observadas. O contexto disciplinar mais utilizado tem sido o da disciplina de matemática, sendo o secundário o nível de ensino mais estudado. O estudo demonstrou ainda que o primeiro ciclo foi o nível de ensino menos investigado.

**Palavras-chave:** avaliação das aprendizagens; avaliação formativa; avaliação sumativa; síntese de dissertações; dissertações mestrado.

## Summary

Since 1990, at the international level, an expressive production of works of reflection and research centred on learning evaluation has emerged, as a result of the awareness of the importance of evaluation for the improvement of the students' learning. In our country, the growing importance of learning evaluation has essentially occurred at three levels: educational policies, production of educational materials and research.

In this study, a characterisation of “mestrado” dissertations produced in Portugal between 1993/94 and 2002/2003 was made, mainly in what concerns contexts in which they took place, problems and research questions, as well as the strategies and the methodological procedures used.

At the methodological level, procedures of a research synthesis were followed, according to Suri and Clarke (1999), using criteria of inclusion/exclusion of conceptual character rather than criteria of value, having recourse to analysis categories, focused reading and synthesis.

Findings of this research show that the objective of most studies either has been constituted by conceptions and practices of evaluation, or just evaluation practices. However, the main source of information was based on interviews to teachers with little resource to class observation. Thus, as a matter of fact, in the majority of the studies, teachers' evaluation practices were not observed. The most used disciplinary context has been mathematics and the secondary has been the most studied educational level. This study has further shown that the primary school was the less investigated educational level.

**Key-words:** learning evaluation, formative evaluation; summative evaluation; dissertations synthesis; “mestrado” dissertations.

## Résumé

À partir des années 1990, on assiste, au niveau international, à une significative production de travaux de réflexion et de recherche centrés sur l'évaluation des apprentissages, fruit d'une prise de conscience de son importance pour l'amélioration des apprentissages des élèves. Dans notre pays, l'importance croissante de l'évaluation des apprentissages s'est vérifié essentiellement à trois niveaux: politiques éducatives, production de matériaux et recherche.

Dans cette étude, on a procédé à une caractérisation de mémoires de D.E.A. produites au Portugal entre 1993/94 et 2002/2003, surtout au niveau des contextes où ils se sont réalisés, des problèmes et questions de recherche, aussi bien que des stratégies et des procédures méthodologiques utilisées.

Au niveau méthodologique, les procédures d'une synthèse de recherche ont été suivies, dans la ligne de Suri e Clarke (1999), en utilisant des critères d'inclusion/exclusion de caractère conceptuel et non pas valorisants, faisant appel à des catégories d'analyse, lecture focalisée et synthèses.

Les résultats de cette recherche mettent en évidence que la majorité des études ont eu comme objectif l'étude des conceptions et pratiques d'évaluation de professeurs, ou seulement les pratiques d'évaluation. Néanmoins, la source principale d'information s'est basée sur des interviews à des enseignants, ayant peu de recours à l'observation de classes. Donc, dans la plupart des études, les pratiques d'évaluation des enseignants n'ont pas été, effectivement, observées. Le contexte disciplinaire le plus utilisé a été celui de la discipline de mathématique, le secondaire étant le niveau d'enseignement le plus enquêté. En plus, l'étude a démontré que le premier cycle a été le niveau d'enseignement le moins enquêté.

**Mots-clés:** évaluation des apprentissages; évaluation formative; évaluation sommative; synthèse de dissertations; dissertations de D.E.A..

## **Agradecimentos**

Agradeço, em primeiríssimo lugar, ao meu orientador que foi incançável, sempre presente, sempre disponível, crítico, construtivo, objectivo, claro, e incisivo. De uma enorme generosidade, ao emprestou-me bibliografia mais recente que ia adquirindo a título pessoal, e disponibilizado-se para, em qualquer momento, trocar ideias e discutir o percurso do meu trabalho. Foi também meu amigo, mostrando a sua capacidade humana de ver os alunos não como apenas alunos, mas como pessoas com vidas próprias, tantas vezes difíceis de conciliar com o trabalho e o estudo. Foi uma honra poder usufruir da sua orientação, do seu conhecimento e de todo o tempo que me disponibilizou.

Também agradeço à minha grande família, aos que estão comigo das mais diferentes formas. Assim, o primeiro é o meu avô materno, eterno companheiro de todos os dias, passados, presentes e futuros. Em seguida à minha MÃE, como só podia, não é mãe?! Depois as irmãs, Ângela e Joana, o Super-Pai, e todos os que ajudaram a tomar conta da Laurinha ao fim-de-semana e/ ou deram uma força para eu não desistir.

Finalmente, os amigos foram muito importantes, desde os colegas de curso com quem se trocaram ideias e palavras de apoio mútuo, até aqueles que nos salvaram dos sempre inesperados problemas de informática, Ana, Carla, Aníbal, Miguel e Guida, Rui, Zé e Bruno.

## **Dedicatória**

Dedico este trabalho, humilde, mas resultante de muito esforço meu, esforço esse partilhado por quem me esteve mais de perto, a minha filha, que tantas vezes me perguntou se ainda faltava muito para acabar o mestrado, à minha filha exactamente. Desejo que ela tenha a mesma sede de saber e conhecer que sempre tive e que possa ir bem mais longe do que eu alguma vez venha a chegar.

Para ti, Laura, da mãe.

## Capítulo I

INTRODUÇÃO.....	7
Orientação Para o Problema .....	7
Problema e Sua Justificação .....	12
Objectivos e Questões de Investigação .....	13
Limitações .....	14

## Capítulo II

REVISÃO DA LITERATURA .....	15
Introdução.....	15
Positivismo, Aprendizagens e Avaliação .....	15
Construtivismo, Aprendizagens e Avaliação.....	17
Mudanças Conceptuais na Avaliação das Aprendizagens.....	21
Funções da Avaliação.....	25
A Avaliação em Contexto de Sala de Aula .....	27
A Avaliação Formativa ou Avaliação <i>Para</i> a Aprendizagem .....	28
A Avaliação Sumativa ou Avaliação <i>da</i> Aprendizagem .....	32
Tensões entre a Avaliação Formativa e a Avaliação Sumativa.....	33
Práticas Avaliativas Mais Comuns .....	36
Avaliação Externa .....	39
Exames e Provas Aferidas .....	40
Principais Críticas aos Exames.....	42

## Capítulo III

METODOLOGIA.....	44
Introdução.....	44
A Síntese de Investigações Como Abordagem Metodológica .....	46
Critérios de Inclusão/Exclusão de Dissertações .....	49
O Processo de Recolha das Dissertações.....	54
Pesquisa Através da Internet .....	54
Pesquisa Através de Contacto Escrito com as Instituições.....	55
Pesquisa Através do Contacto Directo, Presencial, com as Instituições .....	56
A Base de Dados .....	57
Aspectos Tidos em Conta na Análise das Dissertações .....	58
Problemática de uma Investigação .....	58

Breve Síntese de Algumas Questões Metodológicas .....	59
Planos de Investigação .....	63

## **Capítulo IV**

APRESENTAÇÃO E ANÁLISE DOS DADOS .....	68
Introdução.....	68
Caracterização Geral dos Contextos das Investigações Realizadas .....	68
Caracterização das Fontes de Informação Utilizadas .....	73
Contexto Disciplinar das Investigações Analisadas .....	75
Níveis e Anos de Escolaridade .....	78
Zonas do País em que se Realizaram as Investigações .....	82
Síntese da Resposta à Primeira Questão.....	84
Caracterização dos Problemas e Questões de Investigação.....	86
Síntese da Resposta à Segunda Questão.....	95
Caracterização dos Procedimentos Metodológicos .....	97
Número de Participantes e de Escolas.....	97
Procedimentos de Recolha de Informação .....	103
Estratégias de Recolha de Informação.....	111
Síntese da Resposta à Terceira Questão .....	114

## **Capítulo V**

CONCLUSÕES E REFLEXÕES FINAIS .....	116
Introdução.....	116
REFERÊNCIAS BIBLIOGRÁFICAS .....	123
REFERÊNCIAS LEGISLATIVAS .....	130
ANEXOS.....	131



# CAPÍTULO I

## INTRODUÇÃO

### **Orientação Para o Problema**

Nos últimos 20 a 30 anos, temos assistido a desenvolvimentos importantes ao nível das teorias da aprendizagem, dos currícula e das concepções e práticas avaliativas. Podemos dizer que as abordagens mais recentes das teorias da aprendizagem têm em comum a ideia de que aprender é um processo de construção mental e de elaboração de significados, cultural e socialmente condicionado e que se apoia em estruturas de conhecimento já existentes. Considera-se também que o pensamento inteligente envolve uma série de capacidades, incluindo a de saber quando e como utilizá-las. Assim, de acordo com esta abordagem, o aluno deve desenvolver competências ao nível da metacognição e da auto-avaliação, de modo a assumir um papel activo na organização da sua própria aprendizagem (Gipps, 1999; Gipps e Stobart, 2003; Shepard, 2001).

Globalmente, as mudanças curriculares têm vindo a direccionar-se para que o currículo se torne mais flexível, mais dinâmico, mais ligado à vida prática, suscitando uma visão aglutinadora do conhecimento, ao invés da visão compartimentada nos diferentes domínios disciplinares. Estes princípios repercutem-se em propostas que compreendem uma forte componente experimental, incluindo processos de recolha, análise e apresentação de dados por parte dos alunos, passando as aprendizagens a ser vistas de forma mais integrada. Um dos principais objectivos do ensino é o desenvolvimento de processos complexos de pensamento, com recurso à resolução de uma grande diversidade de problemas, confrontando o aluno com tarefas mais elaboradas, conduzindo à integração e mobilização de saberes diversos.

A tomada de consciência, fruto de trabalhos de investigação, das relações entre as práticas de avaliação, as práticas de ensino e as aprendizagens evidenciaram o poder da avaliação na melhoria, não só das próprias aprendizagens, mas também de todo o sistema educativo (Cardinet e Lavenault, 2001; Gipps 1994, 1999; Gipps e Cumming, 2003; Gipps e Stobart, 2003; Sheppard 2001; Wiliam, Lee, Harrisson e Black, 2004). Talvez por tudo isto, nas duas últimas décadas tem-se assistido a uma significativa produção de trabalhos de reflexão e de investigação centrados na avaliação das aprendizagens. Como consequência, recomendam-se práticas avaliativas integradas no processo de ensino e de aprendizagem, mais contextualizadas, e cuja função principal seja inequivocamente, a melhoria das aprendizagens. A avaliação passa a ter uma

natureza eminentemente formativa quer para o professor que, a partir da análise e interpretação da informação obtida, pode reorganizar as suas práticas, quer para o aluno que se torna mais consciente das suas capacidades, dificuldades, motivações, o que lhe confere uma maior autonomia, ao mesmo tempo que lhe atribui uma maior responsabilidade relativamente a todo o processo de aprendizagem. Em síntese, a literatura permite-nos afirmar que a avaliação tende cada vez mais a ser processual, partilhada e até negociada com os alunos, recorrendo-se a tarefas diversificadas e a contextos informais de avaliação (Black e Wiliam, 1998; Cumminng e Maxwell, 1999; Gipps, 1994, 1999; Gipps e Cumming, 2003; Gipps e Stobart, 2003; Shepard, 2001; Wiliam *et al.*, 2004).

Actualmente, predominam essencialmente duas concepções gerais de avaliação das aprendizagens baseadas, respectivamente, no positivismo e no construtivismo (Cole, Ryan e Kick, 1995; Fernandes, 2004, 2005; Gipps, 1994, 1999, Gipps, e Cumming, 2003; Gipps e Stobart, 2003; Guba e Lincoln, 1989; Shepard, 2001). O paradigma dominante ao longo da maior parte do século passado foi o psicométrico, de acordo com o qual a avaliação consistia essencialmente na medição de conhecimentos e sua eventual aplicação a situações práticas, devendo a informação resultante conduzir a um juízo de valor que serviria posteriormente para fundamentar a tomada de decisões, essencialmente de natureza selectiva e certificativa. A partir da década de 80 do século passado, começámos a assistir ao emergir, com maior consistência teórica, do construtivismo. De acordo com esta perspectiva a aprendizagem é uma construção social, feita através da atribuição de significados, dentro de uma rede de valores e crenças e de uma estrutura mental pré-existent, mas sempre em mutação, devido à integração de novos saberes. Este processo ocorre, mais facilmente, em contextos tão próximos do real quanto possível, tornando-se mais eficaz quando envolve o aluno em tarefas de natureza experimental. Neste contexto, a avaliação pode constituir um auxílio à aprendizagem, sobretudo se criar condições para o desenvolvimento de competências de auto-avaliação e metacognição no aluno. De um modo geral, a expressão *avaliação alternativa* é utilizada para referir uma avaliação de natureza mais formativa, mais integrada nos processos de ensino e aprendizagem e com a função principal de melhorar estes processos.

Os trabalhos de natureza teórica, as investigações e as reflexões produzidas nos últimos 20 a 30 anos no domínio da avaliação das aprendizagens poderão ser sistematizados da seguinte maneira:

1. Estudos que se ocupam da reflexão acerca de mudanças conceptuais na avaliação, referindo um progressivo afastamento relativamente às teorias behavioristas e psicométricas, que pressupõem uma avaliação muito associada a uma racionalidade técnica e uma aproximação a abordagens de natureza cognitivista e construtivista, mais associadas a racionalidades críticas que tendem a “ver” a avaliação como uma construção social (Alkin e Christie, 2004; Cardinet e Lavenault, 2001; Cole, Ryan, e Kick, 1995; Coll, Martín e Onrubia, 2001; De Ketele, 2001; Gipps, 1994, 1999; Gipps e Stobart, 2003; Guba, e Lincoln, 1994; Hadji, 1994; Marsh, 1997; Shepard 2001; Torrance, 1993);

2. Estudos que, em consequência da mudança conceptual referida, atribuem à avaliação um papel importante na criação de competências de auto-regulação e de auto-avaliação das aprendizagens nos próprios alunos, que passam a ser sujeitos activos nas tarefas avaliativas (Allal, 2001; Black e Wiliam, 1998; Cumming e Maxwell, 1999; Gipps, 1994; 1999; Gipps, e Cumming, 2003; Gipps e Stobart, 2003; Lee e Gavine, 2003; Perrenoud, 1999; Sadler, 1998; Shepard, 2001; Wiliam, Lee, Harrison e Black, 2004);

3. Estudos de caracterização de ambientes de avaliação em salas de aula, das tarefas propostas e das relações professor/alunos e aluno/aluno (Black e Wiliam, 1998; Cole, Ryan, e Kick, 1995; Crooks, 1988; Gipps e Cumming 2003; Huba e Freed, 2000; McMillan, Myran e Workman, 1999; Stiggings e Conklin, 1992);

4. Estudos de caracterização das relações entre a avaliação formativa e a avaliação sumativa (Biggs, 1998; Harlen, 2006; Shepard, 2001);

5. Estudos acerca das relações entre a avaliação interna e a avaliação externa e acerca das consequências desta nas práticas de sala de aula, nos alunos e nos professores (Gipps, 1994; Herman e Golan, 1993; Kellaghan e Madaus, 2003; Kellaghan, Madaus e Airasian, 1980; McMillan *et al.*, 1999; Madaus, Haney e Kreitzer, 2000; Plomp, Howie e McGaw, 2003; Shepard, 2001; Sturman, 2003; Wiliam, Lee, *et al.* 2004).

Em Portugal, de acordo com Fernandes (2004, 2005), foi sobretudo a partir dos anos noventa do século passado que se assistiu a um desenvolvimento assinalável da avaliação das aprendizagens, podendo destacar-se essencialmente três domínios de análise: políticas educativas; produção de materiais; e investigação.

No domínio das políticas, são de destacar normativos e orientações que, na sequência da Lei de Bases do Sistema Educativo (Lei nº46/86), afirmam claramente que

a avaliação formativa é a modalidade de avaliação por excelência (Despacho nº 162/ME/91; Despacho nº 98-A/92; Despacho nº 338/93; Despacho nº 30/2000; e Despacho nº 1/2005).

Por exemplo, no que se refere ao Despacho 162/ME/91, Fernandes (2004) refere, nomeadamente

Se quisermos, foi um despacho que, até certo ponto, «alinhou» parte dos conteúdos com o «movimento» que, um pouco por todos os países mais desenvolvidos, se tem vindo a designar por *avaliação alternativa* (Gipps e Stobart, 2003), *avaliação autêntica* (Newmann e Archbald, 1992; Tellez, 1996) ou *avaliação educativa* (Gipps, 1994) de inspiração cognitivista e construtivista, em contraste com a avaliação dominante, inspirada nos modelos psicométricos de medição da inteligência e das aptidões (p.36).

Ainda de acordo com Fernandes, o Despacho Normativo nº98-A/92

(...) manteve, no essencial, os princípios e orientações constantes no Despacho anterior. Porventura, sublinhou mais enfaticamente a inspiração cognitivista e construtivista e a concepção fundamentalmente positiva da avaliação das aprendizagens (p.36).

Os despachos e orientações que se seguiram mantiveram, no essencial, as orientações dos anteriores com ligeiras adaptações. Ou seja, o Despacho Normativo nº30/2001 vem concretizar a determinação do Decreto Lei nº6/2001, ou seja, *estabelece os princípios assim como os efeitos da avaliação* (Fernandes, 2005, p.124). Assim, a avaliação passa a ocupar-se de saberes em acção ou em utilização, as chamadas competências, importadas do mundo da economia e empresarial, reforça dimensões como a de diagnosticar, ao nível de cada turma e mesmo de cada aluno, por forma a procederem-se a adaptações curriculares específicas, e sublinha a relevância da auto-avaliação por parte dos alunos, tornando-os co-responsáveis e mais conscientes do processo da aprendizagem. É também defendida uma maior diversidade de estratégias de avaliação e surge a possibilidade de proceder a condições especiais de progressão para alunos com capacidades fora da média. Finalmente, surge o *Processo Individual do Aluno* e o *Projecto Curricular de Turma* como meios de integração da informação ao nível transversal e longitudinal do processo de ensino e aprendizagem.

No domínio da produção de materiais, Fernandes (2005) faz uma resenha onde, ao nível dos materiais de apoio/orientação, destaca os que foram produzidos no âmbito do Instituto de Inovação Educacional (IIE) no início da década de 90. Estes documentos foram organizados no dossier “Pensar a avaliação, melhorar a aprendizagem”, que foi distribuído por todas as escolas do ensino básico e abordavam seis áreas distintas no domínio da avaliação das aprendizagens - Perspectivas de Avaliação, Avaliação

Formativa, Avaliação Sumativa, Pedagogia Diferenciada e Apoios Educativos, Projectos de Avaliação e Avaliação Aferida - destinadas a apoiar a formação dos professores.

São ainda referidos trabalhos de reflexão teórica (*e.g.* , Cortesão e Torres, 1983 e 1984; Benavente, Alaiz, Barbosa, Campos, Carvalho e Neves, 1995; Valadares e Graça, 1998; e Ramalho, 2003) e de natureza empírica (*e.g.*, Lobo, 1996 e Pais 2000).

São referidos ainda um conjunto de publicações em que prevalecem orientações de natureza prática relativamente aos processos de ensino e avaliação (*e.g.*, Ribeiro e Ribeiro, 1989; Pais e Monteiro, 1996; Leite e Fernandes, 2002; e Abrantes e Araújo (Orgs.) de 2002), bem como livros (Estrela e Nóvoa (Orgs.), 1992 e Estrela, Marmoz, Pires e Pereira (Orgs.), 1998) e artigos de divulgação em geral que se reuniram em periódicos no âmbito da educação, inteiramente dedicados à avaliação (Revista Inovação, nº 3, Vol. 4 de 1990; Revista Noesis, nº 23, de 1992 e Revista Noesis, nº 40, de 1996).

É também a partir dos anos 90 que surgem os primeiros cursos de pós-graduação no domínio da avaliação. Os primeiros mestrados em Ciências da Educação com especialização em Avaliação surgem na primeira metade dos anos 90 na Universidade Católica Portuguesa e na Faculdade de Psicologia e Ciências da Educação de Lisboa, conduzindo à elaboração de um conjunto de dissertações sobre a avaliação das aprendizagens. Por outro lado, ao longo desta mesma década, mestrados com especialização no ensino de disciplinas específicas levaram à produção de dissertações centradas também na avaliação das aprendizagens, como é o caso de Rafael (1998) e Fonseca (1998). No domínio da avaliação das aprendizagens verifica-se a existência de dissertações centradas no estudo: a) das concepções de avaliação de professores (*e.g.* Alves, 1997; Santo, 2002); b) da relação entre as concepções e as práticas de avaliação dos professores (*e.g.* Alaiz, 1996; Campos, 1996; Santos, 2003); c) das concepções avaliativas dos alunos (*e.g.* Alves, 1998; Lobo, 1996); d) da história da avaliação (*e.g.* Ferreira, 1994; Rodrigues, 2001); e) das estratégias avaliativas e/ou os domínios objecto de avaliação (*e.g.* Campos, 1996; Santos, 2002); e f) da avaliação externa (*e.g.* Teixeira, 1994).

Constata-se assim que a tomada de consciência da necessidade de criação de cursos de mestrado e de programas de doutoramento dedicados especificamente à

avaliação em contexto educativo surge numa fase de profundas reformas, inovações, reflexões e estudos neste domínio, a par da consequente produção de uma certa diversidade de dissertações acerca da temática da avaliação em geral e da avaliação das aprendizagens em particular.

### **Problema e Sua Justificação**

Tendo em conta os recentes desenvolvimentos no sistema educativo - com implicações ao nível da reorganização da carga lectiva, das directrizes programáticas, da implementação de novas disciplinas e áreas curriculares não disciplinares, entre outras – tendo em conta o desfasamento da escola face às expectativas e interesses de muitos jovens, que abandonam precocemente os estudos, tendo em conta a falta de flexibilidade do sistema que continua a conduzir a um elevado nível de insucesso logo desde o primeiro ciclo, e à falta de alternativas para os chamados alunos com necessidades educativas especiais e não só, a área da avaliação das aprendizagens é uma prioridade de investigação. Por outro lado, o volume de estudos já produzidos, justifica que se faça um ponto de situação do trabalho desenvolvido neste domínio, reflectindo acerca dele, sistematizando-o e pondo-o à disposição de outros investigadores e da comunidades educativa em geral.

A distância que separa os desenvolvimentos teóricos, fruto de trabalhos científicos no campo da avaliação educativa, das implementações dos seus resultados e recomendações a nível prático para as escolas e salas de aula, pode ser maior ou menor de acordo com a forma como cada país valoriza e divulga o conhecimento científico que é produzido. Assim, para além da esmagadora maioria das dissertações académicas feitas no nosso país não serem publicadas, vemos que, apesar da acima referida proliferação de trabalhos quer ao nível teórico, quer no domínio da investigação, durante os últimos anos, não se conhecem reflexões credíveis que tenham procedido a uma análise e sistematização dos seus resultados e recomendações (Fernandes, 1992, 1994, 2004). Tais sínteses são consideradas fundamentais pois podem permitir a identificação de problemas e eventuais formas de solução e de desenvolvimentos teóricos e práticos que abram novos caminhos no domínio da avaliação das aprendizagens (Cooper e Lindsay, 1997; Fernandes, 1992, 1994, 2004).

Tendo em conta o que acima se explicitou, parece relevante produzir trabalhos de investigação que sistematizem e interpretem os referidos desenvolvimentos e lhes

dêem significado. Assim, este estudo visa proceder a uma análise e a uma síntese de investigação produzida no âmbito de programas de mestrado organizados por universidades portuguesas com assento no Conselho de Reitores das Universidades Portuguesas (CRUP), de acordo com a composição deste organismo no ano lectivo de 2002/2003.

Assim, foi decidido analisar as dissertações de mestrado na área da avaliação das aprendizagens, que se concluíram entre os anos lectivos de 1993/1994 e 2002/2003, inclusive. A escolha deste período prende-se com o facto de ter sido a partir de 1992 que tiveram início os primeiros mestrados em avaliação e a opção de 2002/2003 como limite máximo prende-se com o facto de uma década de anos lectivos parecer um período razoável para se fazer este ponto de situação, e também por questões de natureza prática relacionadas com o tempo necessário para as dissertações estarem acessíveis ao público em geral, através das bibliotecas das diferentes instituições. Outros critérios de selecção das dissertações incluídos nesta investigação são apresentados no Capítulo III.

### **Objectivos e Questões de Investigação**

Podemos nomear como principais objectivos deste trabalho, os seguintes:

1. Sintetizar a investigação produzida;
2. Procurar pontos de convergência e divergência entre a investigação analisada e a revisão da literatura referente à investigação realizada noutros países;
3. Identificar prioridades de investigação futura.

Não é objectivo do presente trabalho emitir qualquer juízo de valor acerca da qualidade das dissertações revistas. Essencialmente pretende-se descrever, tão fielmente quanto possível, os problemas, as metodologias, os resultados e conclusões de um corpo significativo de investigação realizado nas universidades portuguesas, nos dez anos acima referidos.

Nestas condições, formularam-se as seguintes questões orientadoras deste trabalho de investigação:

1. Como é que se poderão caracterizar as investigações realizadas?

2. Como se poderão descrever a natureza dos problemas e das questões de investigação das dissertações analisadas?
3. Quais foram as principais estratégias de investigação e os procedimentos metodológicos utilizados pelos investigadores?

### **Limitações**

Embora se tenha envidado o maior esforço possível para conseguir obter a totalidade de dissertações que respektassem os requisitos referidos para o presente estudo, é possível que, por razões devidamente explicitadas na metodologia, não estejam incluídas todas as dissertações com condições para que tal inclusão se realizasse. No entanto, a estratégia de pesquisa utilizada permite referir que, na eventualidade de isso acontecer, será com um número muito reduzido de dissertações. Consequentemente, as conclusões e reflexões finais desta dissertação não serão afectadas de forma significativa por esta eventualidade.



## **Capítulo II**

### **REVISÃO DA LITERATURA**

#### **Introdução**

No presente capítulo, pretende-se fazer uma breve síntese das concepções de aprendizagem e de avaliação segundo as abordagens positivista e construtivista, tomando-as como extremos de um contínuo evolutivo do debate paradigmático que tem decorrido no âmbito das ciências sociais em geral e das ciências da educação em particular. De seguida, procede-se a uma breve reflexão acerca das principais e mais recentes mudanças conceptuais, bem como das funções que a avaliação vem assumindo, resultantes das mudanças ao nível paradigmático.

O contexto de sala de aula é essencial para uma melhor compreensão dos factores condicionantes da aprendizagem, nomeadamente as práticas de ensino e de avaliação. Por isso, é tendo em conta esse contexto que se produz uma reflexão sobre a avaliação formativa e a avaliação sumativa. Além disso, analisam-se os principais efeitos e características das avaliações externa e interna, e as relações existentes entre estes dois tipos de avaliação.

#### **Positivismo, Aprendizagens e Avaliação**

Ao longo do séc. XIX e grande parte do séc. XX, as ciências sociais tiveram como paradigma dominante o positivismo, fruto da tentativa de se tornarem credíveis cientificamente, no seguimento do esforço iniciado, ainda no séc. XIX, para aplicação do método científico a este domínio do conhecimento.

De acordo com este paradigma, a realidade é objectiva, regida por leis e mecanismos independentes dos contextos e dos momentos, sendo passível de ser apreendida, explicada e generalizável. Nesta abordagem, a apreensão da realidade é feita com recurso à metodologia experimental, herdada das ciências naturais, delineando-se hipóteses a serem, posteriormente, testadas. A relação do observador ou investigador com o conhecimento produzido é considerada neutra, devido às estratégias metodológicas utilizadas, concebidas no sentido de minimizar os riscos de “contaminação” dos resultados (Gipps, 1999, e Stobart 2003; Guba e Lincoln, 1989).

A abordagem positivista reduz a aprendizagem praticamente à reprodução, por parte do aluno, dos conhecimentos e capacidades que lhe foram transmitidos. As teorias da aprendizagem dominantes até ao final dos anos 1930 foram as associacionistas, com os trabalhos de Thorndike como principal referência e, a partir dos anos quarenta, as teorias comportamentais, com os trabalhos de teóricos e investigadores tais como Hull, Skinner e Gagné (Sprinthal e Sprinthal, 1990). Estas duas teorias caracterizam-se por considerarem, implícita ou explicitamente, que, à entrada da escola, o aluno constitui uma espécie de *tábua rasa*, ao nível dos conhecimentos; as capacidades individuais são inatas e inalteráveis ao longo da vida, independentemente do ensino e das interações sociais que se desenvolvem e, por último, os indivíduos são diferenciáveis ao nível das suas capacidades inatas. Torna-se natural falar em “níveis de inteligência” e em “níveis de aprendizagem”, e construir meios de os determinar por medição objectiva. Paralelamente, assiste-se à impregnação do taylorismo do mundo industrial no sistema de ensino, com uma concepção utilitarista da aprendizagem, de acordo com a qual o currículo passa a ser diferenciado para se adequar aos papéis sociais pré-determinados e às capacidades individuais de cada aluno (Birenbaum, 1996; Coll, Martín e Onrubia, 2001; Guba e Lincoln, 1989; Shepard, 2001).

Com o contributo essencial de Tyler, nos anos trinta e quarenta do séc. XX, surge a “aprendizagem por objectivos”, que consiste na concepção do currículo cuidadosamente segmentado em parcelas pré-determinadas, a serem alcançadas pelos alunos através da aprendizagem. Neste contexto, a avaliação consiste em medir a posição de cada aluno face aos objectivos definidos e a aprendizagem é feita por transmissão de nova informação por parte do professor, de modo cumulativo, sequencial e com carácter hierárquico. Trata-se de uma perspectiva cumulativa, que assenta na necessidade de se adquirir capacidades básicas para alcançar outras mais complexas, sendo ainda a aprendizagem coadjuvada pela associação de reforços positivos e/ou negativos às atitudes dos alunos em processo de ensino e aprendizagem (Coll, Martín e Onrubia, 2001; Shepard, 2001; Sprinthal e Sprinthal, 1990).

Tendo em conta as teorias de aprendizagem dominantes e a implementação do método científico nas ciências sociais e humanas, nomeadamente nas ciências da educação, assiste-se a uma grande preocupação no sentido de que a avaliação respeite princípios do conhecimento científico tais como: possibilitar a comparação dos resultados de diferentes indivíduos de forma neutra e objectiva; permitir a previsão de

resultados em situações semelhantes às estudadas, bem como viabilizar a criação de leis generalizáveis sobre a aprendizagem, independentemente dos contextos sociais, culturais e temporais. Para o efeito, implementa-se uma avaliação que pretende ser estandardizada, quer no conteúdo quer nas condições de aplicação, e cujos resultados são quantificáveis e descontextualizados. O instrumento, por excelência, que serve estes objectivos é o teste escrito, que cumpria os propósitos de uma concepção de avaliação como *medida isenta de valores* (Scriven, 2003) e que atinge, nos anos vinte e trinta do século passado, uma fase áurea nos Estados Unidos da América (EUA) sob a forma de avaliação externa estandardizada, tendo-se mantido neste país e disseminado para outros, ocidentais e não só, para servir fins certificativos e selectivos a nível do ensino e também de empregabilidade (Birenbaum, 1996; Gipps, 1999; Guba e Lincoln, 1994; Sturman, 2003). Em síntese, a avaliação de natureza positivista ou psicométrica procura, então, medir aprendizagens através de instrumentos de formato fechado (e.g.o testes, exames) sendo os resultados de natureza quantitativa. A sua implementação é estandardizada, de modo a dar garantias de fiabilidade que permitam comparações e generalizações dos resultados, de forma supostamente neutra, objectiva, rigorosa e justa.

No entanto, são várias as críticas apontadas a uma avaliação com estas características, uma avaliação que recorre, de forma generalizada, a modos não diversificados de avaliação, tendo como principal suporte o teste escrito, cujos resultados, expressos de forma quantitativa, transmitem uma falsa ideia de neutralidade, objectividade e equidade. Na verdade, facilitando, embora, a comunicação escola-família, este tipo de avaliação baseia-se numa comunicação vazia de conteúdo, uma vez que a quantificação dos resultados, por si só, não permite apreender as reais aprendizagens efectuadas pelos alunos (Perrenoud, 1999).

### **Construtivismo, Aprendizagens e Avaliação**

No final da década de oitenta do séc. XX, Guba e Lincoln (1989) consideram que a avaliação se encontra numa fase de transição ao nível do paradigma que lhe serve de base, defendendo para esta uma natureza mais interpretativa, por dever ter em conta, ao longo de todo o processo, as necessidades e interesses dos implicados. A avaliação passa a assumir, ainda, uma natureza construtivista, na medida em que é considerada uma construção social. Ou seja, a avaliação resulta muito das interações sociais que se estabelecem nas salas de aula. Para Shepard (2001), este paradigma encontra-se numa

fase emergente, pois não está, completamente desenvolvido ao nível teórico e os seus reflexos práticos são ainda muito reduzidos.

A noção de conhecimento como resultado de uma construção individual e social da realidade, a partir da atribuição de significados, ganha especial relevo na segunda metade do séc. XX. Simultaneamente, iam-se acumulando as evidências teóricas e práticas da falta de adequação de alguns princípios da psicometria, quando aplicados às ciências sociais e humanas, bem como das estratégias metodológicas utilizadas na pesquisa (Gipps, 1994, 1999). Os trabalhos de Bruner - nomeadamente a criação da teoria da aprendizagem pela descoberta e dos seus quatro princípios fundamentais: motivação, estrutura, sequência e reforço – e os de Bandura, - na defesa da aprendizagem por modelagem e na criação da teoria da aprendizagem social - nos anos sessenta e setenta do séc. XX, introduziram a ideia de que qualquer criança pode aprender, desde que lhe sejam facultados os meios e a motivação adequados (Sprinthal e Sprinthal, 1990). Com estes importantes contributos, a aprendizagem passou a estar progressivamente mais virada para a compreensão, tornando-se o desenvolvimento do aluno o foco principal do processo de ensino. A aprendizagem passou a ter um papel capacitante, procurando-se alcançar o melhor desempenho possível de cada aluno, em vez de um desempenho estandardizado. De acordo com Sprinthal e Sprinthal (1990), na segunda metade da década de setenta, o conceito de “zona próxima do desenvolvimento”, introduzido por Vigotsky, veio apoiar uma visão do ensino e da aprendizagem em que o professor é um recurso, trabalhando em colaboração com o aluno para que este atinja o seu máximo desenvolvimento. Em concordância com os mesmos autores, uma década depois, Resnick chama a atenção para o facto da aprendizagem resultar de inter-relações complexas entre factores cognitivos, sócio-culturais e afectivos. Além disso, já nos anos noventa, considerou-se que a aprendizagem é diferenciada pelo contexto em que se realizam o ensino, a aprendizagem e a avaliação (Coll, Martín e Onrubia, 2001; Gipps, 1999; Wiggins, 1993).

Ainda nos anos 90, conclui-se que a aprendizagem passa por um processo de atribuição de significados, por parte do sujeito aprendente. Mais concretamente, o paradigma construtivista encara a aprendizagem como a construção de conhecimento e a atribuição individual de significados, envolvendo processos diversificados e complexos (Gipps, 1999; Gipps e Cumming, 2003; Gipps e Stobart, 2003; Shepard,

2001). Shepard (2001) considera mesmo ter-se assistido a uma revolução cognitiva, assente nos princípios de que: aprender é um processo activo de construção mental e de elaboração de significados; as capacidades intelectuais são desenvolvidas de forma social e culturalmente contextualizada; as estruturas de conhecimento existentes, condicionadas pelos valores e crenças pessoais, vão viabilizar ou não as novas aprendizagens; os processos mentais complexos envolvem a metacognição e a auto monitorização da aprendizagem e do pensamento, bem como o discernimento acerca de quando e como usar as diferentes aptidões; o conhecimento profundo consiste numa forma de pensar e representar problemas de forma coerente e organizada, sendo passível de transferência para diferentes situações e, por último, o desempenho cognitivo depende de predisposições e características da personalidade individual.

Em suma, assistiu-se a uma mudança progressiva ao nível das teorias da aprendizagem, que passaram a considerar o aluno no lugar central do processo de ensino-aprendizagem, uma vez que, como já foi referido, a aprendizagem é fortemente condicionada por factores intrínsecos ao aluno – cognitivos e afectivos – bem como extrínsecos – sócio-culturais, de contexto do processo de ensino, aprendizagem e avaliação (Gipps, 1999 e 2003; Gipps e Stobart, 2003; Shepard, 2001). A esta visão de aprendizagem resultante de uma construção de significados está associada a necessidade de uma avaliação do produto, mas também de todos os processos que lhe são inerentes. Neste sentido, a avaliação tem de ser diversificada, procurando caracterizar em profundidade a estrutura, a qualidade da aprendizagem e a compreensão dos alunos. A focalização do ensino e aprendizagem no aluno implica que este passe a conhecer-se como aprendente, tornando-se importante a chamada meta-aprendizagem, para além da sua co-responsabilização em todo o processo de ensino, aprendizagem e avaliação, aspectos que passam a ser indissociáveis e a funcionar como uma tríade.

As principais consequências do que acima se referiu vão no sentido da tomada de consciência de que: as capacidades de cada aluno vão mudando ao longo do tempo, fruto de aprendizagens escolares e resultantes de interacções sociais e culturais; a avaliação não é uma medida absoluta; os resultados da avaliação dos alunos estão relacionados com as estratégias utilizadas e com o momento e contexto de implementação e, finalmente, a construção da realidade, assim como a aprendizagem, dependem do sistema individual de crenças e de valores do aprendente, pelo que a avaliação tem de ser social e culturalmente contextualizada (Gipps, 1994; Guba e

Lincoln, 1989). Torna-se necessário pôr em prática uma avaliação que reflecta as aprendizagens realizadas pelo aluno, bem como o seu empenho nas tarefas, incidindo nos aspectos mais importante do currículo, no que foi ensinado e aprendido, mas também no modo como esta aprendizagem se efectuou. A avaliação deve, então, constituir-se como parte integrante do processo de ensino e de aprendizagem.

Deste modo, o currículo vai sofrer alterações, passando a ter em conta que a aprendizagem se faz com base na construção de significados, sendo retida mais fácil e longamente através da participação do aluno em actividades práticas de manipulação e análise dos dados, e da sua aplicação a novas situações. A metacognição passa a ser considerada um aspecto chave para o desenvolvimento do pensamento complexo, devendo o aluno ser capaz de monitorizar a sua própria aprendizagem. O objectivo é preparar indivíduos autónomos e responsáveis pela sua aprendizagem, que terá lugar ao longo da vida, como resultado das constantes inovações e novos desafios propostos pelas sociedades actuais. Procura-se que o currículo se torne mais flexível, não só ao nível de uma visão menos estanque das matérias leccionadas, criando-se espaços multidisciplinares, mas também ao nível dos implicados no processo de ensino e aprendizagem, passando a ser tidas em conta as necessidades, interesses e características sócio-culturais e pessoais dos alunos. Este currículo deve ser dinâmico, de modo a adaptar-se à vida real, ocupando-se não só com conteúdos académicos, mas também com o desenvolvimento de domínios como o das relações sociais, das situações de trabalho de grupo, exigindo-se ao aluno não apenas capacidades de liderança, mas também de cooperação em equipa. As actividades de aprendizagem passam a assumir uma natureza mais interactiva; recorre-se, com maior frequência, à manipulação de dados, desde a sua recolha até à sua análise, contempla-se a resolução de uma grande diversidade de problemas, exige-se a realização de trabalhos de projecto, implicando os alunos de forma empenhada nas actividades de sala de aula (Coll, Martín e Onrubia, 2001; Fernandes, 2004 e 2005; Gipps e Cumming, 2003; Shepard, 2001). Democratização, flexibilidade, experimentação ou manuseamento, recurso a pensamento complexo, e resolução de problemas mais ligados à realidade prática vão ser factores chave de um novo currículo, que se quer mais dirigido a todos os alunos, proporcionando igualdade de oportunidades, em termos de oferta de ensino de elevada qualidade, bem como formas de implementação concordantes com os aspectos sociais e

culturais, de modo a proporcionar aprendizagens autênticas ligadas à realidade quotidiana, ou seja, contextualizando-as.

### **Mudanças Conceptuais na Avaliação das Aprendizagens**

À medida que se foram constatando inovações ao nível de como se aprende e de como o currículo se deve adequar a essas descobertas, foram-se desenvolvendo diferentes concepções de avaliação, para as quais também contribuíram fortemente os conceitos de socioculturalidade (Gipps, 1999) e a abordagem interpretativa (Shepard, 2001). Figari (2001) sistematizou as principais inovações operadas nas concepções de avaliação em geral, nos últimos quinze anos, a três níveis:

1. As concepções que se concentram na avaliação como uma actividade, onde se observa um interesse crescente no que sucede ao longo do acto avaliativo. Por um lado, o interesse e questionamento da avaliação centram-se na situação e nos fenómenos que aí se desenvolvem. Por outro, esta avaliação implica uma tal colaboração entre avaliador e avaliado, que leva à construção recíproca dos papéis de cada um. À medida que termina o séc. XX, a avaliação procura cada vez mais caracterizar em profundidade a estrutura e a qualidade da aprendizagem, tal como a compreensão dos alunos, tornando-se uma tarefa cada vez mais interactiva, que tem em consideração aspectos individuais do aluno, tais como o crescimento, o desenvolvimento, o esforço, a reflexão, a iniciativa e a mudança. As novas formas de avaliação recorrem a tarefas mais abertas e diversificadas como sejam tarefas em pequeno grupo e de experimentação, incluindo a recolha, análise e tratamento dos dados. Recorrem também a técnicas e meios mais interactivos e informais como ensaios, entrevistas, observação, inventários e avaliação de desempenhos. Uma vez que os critérios, as tarefas e as expectativas avaliativas são, desejavelmente, negociados e pré-estabelecidos, a avaliação exige ao aluno parte da responsabilidade pelos seus resultados. Espera-se, pois, que cada aluno aja de forma autónoma, procedendo à auto-avaliação e à auto-correcção de forma implicada. A avaliação assume um carácter mais dinâmico, deixando de ser feita *para* o aluno, para ser feita *com* o aluno. As novas abordagens avaliativas pretendem ocupar-se dos processos tanto como dos produtos e permitem avaliar actividades mais complexas de pensamento, desenvolvendo as competências dos alunos ao nível da metacognição e da reflexão sobre a sua própria aprendizagem (Cardinet e Laveault, 2001; Cole, Ryan e Kick, 1995; Marsh, 1997; Madaus, Haney e Kreitzer, 2000; Fernandes, 2005; Figari, 2001; Gipps e Cumming, 2003; Gipps e Stobart, 2003; Shepard, 2001).

2. As concepções que têm por base a avaliação como componente da cognição, visão presente nas preocupações crescentes com a centralização dos alunos nas aprendizagens. A avaliação torna-se um processo através do qual os alunos assumem um papel activo na organização das suas próprias aprendizagens. São capacidades basilares a auto-reflexão relativamente à forma como aprendem, como se processa todo o seu pensamento no decurso de uma tarefa. Paralelamente, os alunos devem tornar-se parceiros da negociação que conduz à avaliação, uma vez que os critérios e referenciais de avaliação devem ser partilhados e definidos relativamente a cada situação escolar particular. Este conhecimento, a par das capacidades acima referidas, permitem que o aluno possa proceder a uma auto-avaliação que se transforma em evidência quando implementa a auto-regulação da sua aprendizagem. A capacidade de pensar sobre a aprendizagem, ou seja, a metacognição, revela-se uma competência de elevada utilidade para o aluno. Os processos de ensino, aprendizagem e avaliação ocupam-se em procurar o melhor desenvolvimento do indivíduo e não um desempenho estandardizado. Finalmente, esta avaliação, tal como a aprendizagem, é social e culturalmente contextualizada (Allal, 2001; Fernandes, 2005; Figari, 2001; Gipps e Cumming, 2003; Gipps e Stobart, 2003; Shepard, 2001).

3. As concepções em que a avaliação é vista como forma de prestação de contas e de integração social, onde se destacam aspectos tais como: o papel das normas sociais; o recurso à noção de competência; a tomada em consideração das aquisições profissionais e ainda o reforço do papel certificativo da avaliação. As preocupações ao nível da justiça e equidade fazem apelo a uma avaliação mais aberta ao nível do “design” das tarefas avaliativas, dos valores e preconceitos associados ao processo de construção das tarefas, do debate cultural e social e da relação entre avaliador e aluno (Gipps, 1999). A procura de novas formas de avaliação, que contemplem a diversidade cultural e social, pretende também fornecer um meio para os que tradicionalmente são vistos como menos competentes poderem expor as suas capacidades (Gipps, 1999). Em tudo isto permanece a consciência de que a avaliação tem também uma dimensão política, nomeadamente porque é através das políticas educativas que, por exemplo, se pode apoiar mais ou menos o desenvolvimento da avaliação formativa em sala de aula.

No que respeita às conceptualizações de avaliação das aprendizagens, constata-se que, como se referiu, nos últimos quinze a vinte anos, se assistiu à introdução de consideráveis mudanças ao nível das teorias da aprendizagem, da sua natureza, do currículo e, conseqüentemente, da avaliação, passando esta última a ser vista integrada



com o ensino e a aprendizagem. Idealmente, há que efectuar um balanço constante entre os objectivos da aprendizagem, as actividades de ensino, os processos de aprendizagem e os procedimentos avaliativos. Desenvolvem-se concepções de avaliação mais preocupadas com: a recolha de informação detalhada sobre cada aluno; o apoio do ensino e da aprendizagem através do desenvolvimento de capacidades metacognitivas; a análise dos resultados e de todo o processo, para melhor compreender os mecanismos de aprendizagem e de pensamento do aluno; e o acompanhamento do processo de ensino e de aprendizagem, de forma a poder proporcionar uma retroacção, por parte do professor, mais imediatamente aplicável. Assim, há que recorrer a um conjunto diversificado de tarefas e estratégias de avaliação onde se incluem listas de verificação, relatórios, inventários, gravações em vídeo e áudio, entrevistas, avaliações de desempenho, a observação, a conferência e o debate de ideias, processos escritos, auto e hetero-avaliação e o próprio teste. Este tipo de avaliação exige opções pedagógicas, em que o aluno se torna co-responsável do processo avaliativo, compreendendo claramente a tarefa a realizar, bem como aquilo que se espera dele na referida actividade (Archbald e Newmann, 1992; Cole, Ryan e Kick, 1995; Cumming e Maxwell, 1999; Wiggins, 1989).

Em suma, os desenvolvimentos teóricos permitem que a natureza da avaliação se caracterize sucessivamente por: implicar o aluno em todo o processo de ensino, aprendizagem e avaliação, o que requer uma reorientação das práticas de sala de aula; recorrer a tarefas de carácter eminentemente prático; ter em conta não só os produtos mas todo o percurso de aprendizagem dos alunos; reforçar as funções pedagógica, reguladora e formativa da avaliação, de forma a obter-se uma maior coerência e continuidade entre as decisões de ordem pedagógica e social que lhe estão associadas; atribuir uma maior autonomia aos professores e às instituições educativas na planificação e implementação da avaliação (Coll, Martin e Onrubia, 2001).

À medida que se foi tomando consciência das lacunas da avaliação tradicional e que se foram difundindo as teorias cognitivistas, construtivistas e socioculturais da aprendizagem, a avaliação foi assumindo novos formatos que, embora tendo em comum as características da avaliação de natureza cognitivista e construtivista, conduziram, essencialmente a partir dos anos noventa do século passado, à coexistência de diferentes terminologias, que diferem entre si na ênfase que dão a aspectos específicos dessa avaliação.

Assim, a expressão *avaliação dinâmica* é usada por autores como Newman, Griffin e Cole, e Lunt (citados por Gipps e Stobart, 2003) no sentido de acentuar a sua constante adaptação, com vista a acompanhar os processos de ensino e aprendizagem, explorar o modo como o aluno aprende e pensa, dar informações ao professor acerca das estratégias e processos de aprendizagem do aluno e, indirectamente, fornecer pistas para a melhoria do ensino.

Embora vários autores refiram a importância do contexto em que se realizam quer os processos de ensino e de aprendizagem, quer os de avaliação, Gipps (1999) foi mais longe, considerando mesmo que esta última não pode, de forma alguma, ser vista fora do seu contexto social, económico e político, uma vez que as respostas dos alunos são social e culturalmente condicionadas. A expressão *avaliação autêntica* tem origem na grande ênfase atribuída ao contexto em que a avaliação tem lugar, que se deve aproximar tanto quanto possível de situações reais e, por isso, complexas. O desempenho do aluno torna-se o foco da avaliação, devendo este integrar conhecimentos multidisciplinares de modo funcional, na realização de tarefas diversificadas e complexas. Os critérios de avaliação têm de estar de acordo com as competências visadas e ser conhecidos dos alunos, que têm conhecimento claro daquilo que deles se espera na realização da tarefa avaliativa. Finalmente, a avaliação autêntica inclui tarefas de auto-avaliação e exige do aluno estratégias não apenas cognitivas, mas também metacognitivas (Fernandes, 2005). Trata-se de um tipo de avaliação que ocorre integrada no contexto da sala de aula, de forma a apoiar a aprendizagem do aluno, assumindo uma função essencialmente formativa. Alguns autores propõem estratégias e meios concretos para o exercício desta avaliação, onde se incluem: listas de verificação, relatórios, inventários, vídeos, cassetes áudio e entrevistas (Cole, Ryan e Kick, 1995).

Cardinet e Laveault (2001) utilizam a expressão *comunicação facilitadora* para se referirem a uma avaliação que se baseia na comunicação entre professor e aluno, no sentido de arte último saber claramente o que dele se espera. Esta terminologia surgiu na sequência de uma síntese da investigação realizada entre 1986 e 1996 na Europa e nos EUA, onde os autores concluem que a avaliação em educação vinha adquirindo uma componente comunicacional crescente, com o intuito de tornar visível, para o aluno, as expectativas sobre ele colocadas. De acordo com esta abordagem, a negociação entre o professor e o aluno conduz a aproximações sucessivas até ao alcance do desenvolvimento máximo do aluno.

Em 1997, Marsh utilizou a expressão *avaliação alternativa*, pela primeira vez, em contexto educativo, como oposta à avaliação tradicional, focalizada na medição de competências de nível pouco complexo, através de uma diversidade de testes estandardizados (Marsh, 1997). Mais recentemente, também Gipps e Stobart (2003) e Fernandes (2004) utilizam a expressão *avaliação alternativa* para se referirem a uma abordagem relativamente recente da avaliação das aprendizagens com as características acima referidas. Já em 2005, Fernandes, por considerar a expressão “avaliação alternativa” demasiado vaga, introduz a designação “avaliação formativa alternativa”, com o intuito de sublinhar a natureza formativa que lhe está inerente, salientando que este tipo de avaliação é alternativo à avaliação formativa de natureza *behaviourista* e à “avaliação de intenção formativa”. Isto é, à avaliação a que se referem muitos professores como sendo formativa mas que, na realidade, não o é de facto, tratando-se apenas de uma intenção ou de uma vontade que não passa para a prática.

Em síntese, apercebemo-nos de que as terminologias associadas à avaliação têm tentado retratar aquela que acham ser a principal função que a avaliação deve ter – apoiar a aprendizagem e regular o ensino, devendo ser praticada numa visão de ensino, aprendizagem e avaliação como fases interligadas de um processo.

Interessa agora fazer uma reflexão acerca das diferentes funções da avaliação de acordo com os contextos em que se realiza.

### **Funções da Avaliação**

Embora não pretendendo fazer uma análise exaustiva das diferentes funções que a avaliação pode ter, desde o nível do aluno ao nível do sistema educativo, importa perceber a origem de algumas dessas funções. Referem-se sobretudo as que, na segunda metade do séc. XX, se vêm afirmando como as mais recorrentes.

Na segunda metade da década de cinquenta do século passado, em pleno período da chamada *guerra fria*, a procura de supremacia ao nível da competitividade tecnológica leva, nos EUA, à tomada de decisões bastante drásticas, nomeadamente ao nível da equipa de produção dos programas do *National Science Foundation* e do *Office of Education*. Na área da educação, verifica-se a premência em rever e avaliar programas, instituições e reformular concepções de avaliação, de acordo com a finalidade de trazer maior desenvolvimento para o país, através de uma formação e

educação de elevada qualidade. É neste contexto de revolução educativa que Scriven, em 1967, aplica, pela primeira vez, os termos *avaliação formativa* e *avaliação sumativa*, no âmbito da avaliação de programas. O autor recorre a estas denominações, imbuídas da teoria *behaviourista* dominante na altura, para referir duas grandes funções da avaliação: a formativa, que estaria ao serviço do bom desenvolvimento da actividade em curso, e a sumativa, que recolheria informação para a prestação de contas, certificação e selecção. Só em 1971, estas funções entram no contexto da avaliação das aprendizagens, com Bloom, Hastings e Madaus (Gipps, 1999; Guba e Lincoln, 1989; Nevo, 1986).

De acordo com Nevo (1986), em 1972, Stufflebeam considera dois tipos de avaliação com funções próximas, mas cuja denominação dá mais ênfase ao momento da sua aplicação: a proactiva, associada à tomada de decisões, e a retroactiva, com a função de prestação de contas. Ainda na segunda metade dos anos setenta, com autores como Patton, House e Cronback, a avaliação passa a ser utilizada para chamar a atenção de determinados sectores da sociedade para uma actividade em curso e para motivar comportamentos desejados nos indivíduos avaliados. Nevo (1986) denomina esta dupla função de psicológica e sócio-política. Ainda no referido período, Dornbusch e Scott referem a função de exercício de autoridade que a avaliação permite ao superior quando a realiza sobre um seu subordinado (Nevo, 1986). Assim, Nevo considera as funções referidas como as quatro grandes finalidades que a avaliação pode servir: a formativa; a sumativa; a psicológica ou sócio-política e a administrativa.

Nas décadas de 1980 e 1990, diferentes teóricos e investigadores referem funções como as de informar sobre as aquisições realizadas e controlar a realização de objectivos fixados para o ensino (Noizet e Caverni, 1985), as de diagnosticar, prever, orientar e controlar as aprendizagens dos alunos por parte dos professores (Ramos, 1989), diagnosticar, classificar e prever as aprendizagens do aluno, motivá-lo e ainda diagnosticar o ensino (Marsh, 1997). No entanto, ao nível das aprendizagens, a maioria dos teóricos considera que as funções fundamentais da avaliação são a formativa e a sumativa.

Constata-se que a evolução histórica das funções da avaliação começou com a função selectiva seguida da função certificativa de indivíduos e, mais tarde, a de controlo do currículo e do ensino. Esta evolução prende-se com factores económicos,

mais concretamente com a globalização, que conduziu a uma necessidade de medir os resultados do ensino, numa perspectiva de avaliação com função de melhorar o sistema educativo. De facto, a visão de crescimento económico ligado à educação levou a que países como os EUA, a Nova Zelândia, a Austrália e a Inglaterra recorressem à avaliação para ajudar a melhorar os currículos, impor parâmetros de desempenho elevados e encorajar a competição entre instituições de ensino. A este respeito, Gipps (1999, p.384) afirma:

A testagem é agora usada para controlar o currículo e o ensino. A avaliação é pois algo de poderoso. Existe também um interesse crescente nos resultados da avaliação como indicadores de desempenho, tal como pressão política e nacional por avaliação 'objectiva' altamente fiável de forma a permitir o ranking e graduação do desempenho dos indivíduos e das escolas (e até de países).

Deste modo, as funções socializadora e de promoção do desenvolvimento pessoal estarão na base da dupla vertente social e pedagógica da avaliação que é usada para verificar se é cumprido o projecto social e cultural que preside e guia a educação escolar (Coll, Martín e Onrubia 2001).

Recentemente, e ao nível da avaliação das aprendizagens, os conceitos de *avaliação alternativa* e de *avaliação de sala de aula* acrescentam às finalidades referidas a de apoio do aluno na sua aprendizagem, exigindo o desenvolvimento de capacidades de auto-avaliação, auto-monitorização da aprendizagem e competências de conceptualização de problemas da vida real, de forma organizada e assertiva. Estas funções são fortemente condicionadas pelos propósitos que a avaliação de sala de aula pretende servir.

### **A Avaliação em Contexto de Sala de Aula**

A avaliação em sala de aula assume essencialmente três funções distintas, caracterizadas pelo seu momento de aplicação, pelo seu conteúdo e, sobretudo, pelo objectivo que leva à sua implementação. Assim, no início de um processo de ensino e aprendizagem, com o objectivo de identificar as competências dos alunos face às novas informações curriculares que se vão implementar, pode efectuar-se uma avaliação designada de diagnóstica, inicial ou “predictiva”. A partir dela, o professor obtém informações de modo a delinear as estratégias de ensino a implementar relativamente à situação concreta dos alunos. De acordo com a partilha destes dados com os alunos e do

grau de partilha do processo de planificação dos passos subsequentes, esta avaliação pode assumir um carácter mais ou menos formativo.

Com base nas teorias da aprendizagem, constata-se que é necessário acompanhar o seu processo de realização das tarefas de ensino e de aprendizagem, preferencialmente dando respostas atempadas e contextualizadas face à forma como cada aluno está a desenvolver as actividades propostas. Este segundo tipo de avaliação tem um carácter contínuo, regulador e, logicamente, formativo, uma vez que desenvolve nos alunos capacidades de regulação dos seus próprios processos de aprendizagem, e fornece informações mais contextualizadas aos professores para a tomada de decisões atempadas e, por isso, mais efectivas, no sentido de melhorar as actividades propostas e o ensino (Shepard, 2001; Sprinthal e Sprinthal, 1990).

Finalmente, é necessário determinar até que ponto os alunos adquiriram ou não as competências pretendidas para o que se implementa uma avaliação final em termos de projectos ou de segmentos mais ou menos amplos do processo de ensino-aprendizagem. Pelo seu carácter relativamente globalizante, em termos curriculares, este tipo de avaliação é designada de sumativa (Ribeiro, 1997).

Importa agora analisar as modalidades de avaliação usadas em sala de aula, eventuais tensões existentes e as estratégias a que cada modalidade recorre.

### **A Avaliação Formativa ou Avaliação *Para* a Aprendizagem**

A avaliação formativa tem intenções reguladoras, de desenvolvimento e de melhoria das aprendizagens. Deste modo, trata-se de uma avaliação *para* aprender e não uma avaliação daquilo que se aprendeu. No entanto, a escolha da modalidade, tarefas e estratégias avaliativas deve estar de acordo com o objectivo visado em cada situação concreta, não sendo, pois, possível afirmar a existência de tipos de avaliação melhores ou piores, por si só (Nevo, 1986).

A partir dos anos oitenta do séc. XX, parecem coexistir duas grandes abordagens da avaliação formativa. Uma, mais tradicional, baseada nas correntes *behaviouristas* ligadas à educação, está presente no chamado ensino para a mestria, onde a avaliação pressupõe a definição de objectivos claros, de curto prazo, sucedendo à avaliação uma retroacção, por parte do professor, contendo informação precisa e detalhada de como o aluno pode melhorar os seus resultados. Existe ainda uma segunda, que se apoia no

paradigma construtivista, fundamentado no cognitivismo e nas teorias socioculturais, que dão atenção especial àquilo que os alunos podem aprender a ser capazes de fazer e enfatizam a importância que os recursos e o tipo de tarefa têm para o desenvolvimento dos diferentes processos de aprendizagem dos alunos (Torrance e Harry, 1993).

Actualmente, na prática, continuam a coexistir ambas as concepções acima referidas.

As abordagens mais recentes da avaliação formativa fazem transparecer um conjunto complexo de aspectos que contribuem para a transformar numa avaliação que está essencialmente ao serviço da aprendizagem. Assim, as definições diferenciam-se com base em características como: o momento e frequência da implementação de tarefas avaliativas; o seu tipo; sobre que aspectos da tarefa recai a avaliação; o tipo de retroacção e o momento em que é dada pelo professor; os domínios avaliados; a argumentação e fundamentação das respostas dadas pelos alunos; e o papel do professor e do aluno na avaliação (Black e Wiliam, 1998; Shepard, 2001).

De acordo com as teorias mais recentes de aprendizagem e de avaliação das aprendizagens, para que a avaliação formativa seja uma avaliação *para* a aprendizagem, ela deve ocorrer ao longo de todo o processo de aprendizagem, bem como no seu final, tendo uma frequência tão elevada quanto possível, aproximando-se de uma avaliação contínua. Por outro lado, é necessária uma preocupação em diversificar o mais possível as situações de avaliação, incluindo apresentações orais, construção de textos e de relatórios, desenvolvimento de experiências e resolução de problemas complexos. Procura-se que tenham um carácter tão realista quanto possível, preparando o aluno para situações quotidianas e dando oportunidade a que cada um demonstre o seu desempenho, através de tarefas diversas, de modo a contemplar a diversidade de características de personalidades dos alunos, que reagem de diferentes formas às diferentes tarefas avaliativas. As tarefas devem ser globalizantes em termos de domínios, de conteúdos e de competências. A avaliação não se pode centrar apenas no produto final de uma actividade, mas ter em conta todo o processo de aprendizagem, bem como as estratégias de pensamento a que o aluno recorre para resolver os problemas. Há ainda que considerar o momento e o tipo de retroacção fornecida pelo professor. Com efeito, a retroacção (da máxima relevância para a aprendizagem) deve ser imediata, relativamente ao que se pretende ver alterado, deve também ser específica

ao aluno e à tarefa em curso, uma vez que acompanha o processo avaliativo, permitindo que o aluno rapidamente integre este apoio de forma a melhorar o desempenho. Esta retroacção ou *feedback* não deve consistir na transmissão de qual seria a resposta correcta, mas antes, traçar uma direcção para o auto-questionamento do aluno (Black e Wiliam, 1998; Cumming e Maxwell, 1999; Fernandes, 2005).

Duas capacidades fulcrais no desenvolvimento integral do aluno numa perspectiva construtivista são as de auto-avaliação e as de metacognição. Embora na prática a maioria das actividades metacognitivas recaiam sobre tarefas de finalização de aprendizagem ou de resolução de problemas, sempre ao nível cognitivo, Allal (2001) afirma que as pesquisas mais recentes mostram a necessidade de ligar as dimensões cognitiva e metacognitiva às dimensões motivacional e volitiva. Neste sentido, a metacognição deve procurar perceber/conhecer as competências cognitivas do aluno, bem como as suas estratégias motivacionais (Fernandes, 2005; Sprinthal e Sprinthal, 1990).

O desenvolvimento de capacidades de auto-avaliação está relacionado com a definição de competências metacognitivas. A metacognição, que envolve uma forma de pensamento de segunda ordem, inclui uma variedade de processos de auto consciência para ajudar a planificar, monitorizar, implementar e controlar a própria aprendizagem. Para tal, os alunos usam estratégias como o auto questionamento, a procura de interligações e de conflitos e a elaboração de juízos. A autonomia dos alunos como aprendentes exige que tenham experiência em pôr em questão e melhorar a qualidade do seu trabalho, bem como em auto-avaliação, que inclui uma compreensão daquilo que se espera que ele seja capaz, bem como dos critérios através dos quais vai ser avaliado. O aluno deve ser capaz de usar estratégias que lhe permitam estabelecer relações e conflitos com aquilo que já conhecem e ainda julgar qual o conhecimento necessário para a realização de uma determinada tarefa (Crooks, 1988; Cumming e Maxwell, 1999; Fernandes, 1994).

A importância das capacidades metacognitivas é traduzida na valorização da argumentação e da fundamentação das soluções propostas pelos alunos às tarefas, de discutir, analisar e de emitir juízos fundamentados. A necessidade de preparar alunos para a vida activa exige ainda o desenvolvimento de capacidades de trabalho em equipa e do recurso adequado a instrumentos de apoio às actividades, como dicionários,



calculadoras, materiais de consulta, entre outros. A avaliação, que tradicionalmente era algo em que o aluno não tinha qualquer papel para além de executar as tarefas que lhe eram solicitadas, passa a integrar estes recursos, além dos encarregados de educação e, logicamente, os professores, embora com diferentes competências. Defende-se uma cada vez maior colaboração entre professor e alunos na planificação de todo o processo de ensino, aprendizagem e avaliação como um todo, estabelecendo-se, à partida, aquilo que se deseja ver alcançado pelos alunos, bem como os critérios para a sua avaliação. Deste modo, o papel do professor cada vez mais se reporta ao de um recurso que o aluno tem à disposição para realizar as suas aprendizagens (Black e Wiliam, 1998; Coll, Martins e Onrubia, 2001; Gipps e Stobart, 2003; Harlen, 2006; Sadler, 1998).

Em síntese, a avaliação *para* a aprendizagem deve ter uma natureza mais dinâmica, multifacetada e interactiva. Deve envolver uma grande diversidade de actividades, de modo a proporcionar diferentes oportunidades de aprendizagem; deve envolver uma relação próxima entre o professor e o aluno, sendo as tarefas a desenvolver perfeitamente claras para este último; deve permitir a diferenciação pedagógica e deve permitir respostas de formato diversificado e não apenas as de tipo escrito (Coll, Martin e Onrubia, 2001; Fernandes, 2005; Gipps, 1994).

O teste escrito tradicional, quando construído com base em questões abertas, pode constituir um bom meio para os alunos demonstrarem as suas capacidades de usar evidência histórica ou científica para sustentar as suas ideias. As questões de resposta curta, de escolha múltipla ou de preenchimento de espaços também podem ser úteis para verificar determinados tipos de conhecimento processual. No entanto, é importante certificar-se que as capacidades solicitadas para a sua resolução estejam representadas de forma proporcional à sua importância no currículo, em vez de serem utilizadas de acordo com a sua maior ou menor facilidade de medição (Ribeiro e Ribeiro, 1989).

Já os portfólios constituem uma outra forma, recentemente muito popularizada, de avaliação. Em avaliação em sala de aula podem proporcionar uma estrutura organizada para o desenvolvimento da capacidade crítica e de auto-reflexão do aluno, apoiado pelo professor; a preocupação da estrutura do portfólio deve servir de ajuda aos professores e alunos para que se tornem mais sistemáticos na análise e aprendizagem a partir do trabalho do aluno. Através da construção do portfólio, o aluno traça objectivos de aprendizagem, revê o seu trabalho e desenvolve critérios para a selecção de

determinados documentos relativamente a outros, aprende a avaliar os aspectos em que o seu trabalho apresenta pontos fortes e fraquezas, e adquire experiência de comunicação dos seus objectivos e juízos aos outros. Os portfólios podem ser usados com fins de aprendizagem e avaliação, integrando documentos que retratem aquisições, autoavaliação, fases processuais e análise da experiência de aprendizagem. É, no entanto, controverso até que ponto o portfólio pode ser usado com fins de avaliação para a aprendizagem e, simultaneamente, de avaliação da aprendizagem (Gipps e Stobart, 2003).

Os principais problemas apontados a qualquer tipo de avaliação são a dificuldade em encontrar um equilíbrio entre a diferenciação pedagógica e a justiça na igualdade de tratamento, e ainda o tempo que esta avaliação exige para ser implementada.

### **A Avaliação Sumativa ou Avaliação da Aprendizagem**

A avaliação sumativa assume tradicionalmente dimensões certificativas, selectivas e classificativas a níveis pedagógico, social e político. Já a avaliação formativa tem por objectivo servir de apoio à aprendizagem, exigindo um grau de confiança elevado dos alunos para com os professores, de forma a poderem explicitar claramente as suas dificuldades sem receios de que as suas dúvidas possam ser vistas como evidências de falta de competência e não como uma fase do processo de aprendizagem. Dadas as diferentes funções destes tipos de avaliação e o facto de o principal agente de ambas ser a mesma pessoa, o professor, tende a haver uma dificuldade de articulação e de destrição entre a avaliação sumativa e a formativa.

Dadas as finalidades atrás referidas e cujas consequências têm efeitos a nível da progressão ou retenção dos alunos, ou mesmo da sua selecção para níveis superiores de ensino ou/e postos de trabalho, a avaliação sumativa deve ser criterial e não normativa, integrando todo o currículo que se pretende alcançar e não apenas parcelas deste. O momento da sua implementação deve ser integrado hierarquicamente no processo de ensino, aprendizagem e avaliação, podendo, para autores como Biggs (1998) e Coll, Martins e Onrubia (2001) ser altamente informativa e formativa para os processos de ensino subsequentes. A abordagem psicométrica da avaliação sumativa permanece ainda nas concepções de muitos professores e decisores, que enfatizam a quantificação das diferenças entre os alunos.

Dadas as funções selectiva e certificativa da avaliação sumativa, é importante que a tarefa de avaliação garanta a fiabilidade dos resultados, permita determinar desempenhos de diferentes níveis e, eventualmente, tirar ilações acerca do sistema educativo. As tarefas usadas neste tipo de avaliação tendem a ser mais fechadas, implementadas em momentos pré-determinados e assumir a forma de testes e outras tarefas escritas. Procura-se, com isto, uma avaliação estandardizada que permita ver como reagem diferentes alunos a uma mesma situação (Biggs, 1998; Fernades, 2004 e 2005).

Ambas as formas de avaliação, formativa e sumativa, são necessárias. Na maioria das vezes, porém, os seus objectivos entram em conflito, pelo que se torna, de certo modo, inevitável a existência de tensões entre boas práticas em sala de aula e o recurso a testes a nível de escola, do país ou mesmo internacionalmente (Gipps e Cumming, 2003).

### **Tensões entre a Avaliação Formativa e a Avaliação Sumativa**

Shepard (2001) opta por defender que a avaliação com carácter formativo está associada à avaliação de sala de aula e à avaliação interna, enquanto que a avaliação com carácter sumativo está próxima da avaliação externa, aplicada em menor ou maior escala. Também para esta autora, a avaliação formativa e a avaliação sumativa servem propósitos distintos, mas ambos necessários, pelo que se pode afirmar possuírem uma relação de complementaridade. No entanto, para Shepard (2001), é da máxima importância que ambas as modalidades se baseiem nos mesmos *standards* curriculares, envolvam os alunos no mesmo tipo de investigações e demonstrações de proficiência e permitam que estes sejam avaliados com base em critérios previamente partilhados, de forma a ser feito um juízo sobre trabalhos de elevada qualidade.

Surgem, então, discussões sobre formas de ultrapassar as tensões decorrentes de avaliações com propósitos, em alguns aspectos, antagónicos. Por outro lado, Gipps (referida por Shepard, 2001) defende que a avaliação selectiva, de prestação de contas, pode também servir fins de apoio à aprendizagem. No entanto, Shepard (2001) considera que dificilmente a avaliação selectiva ou certificativa colocará a questão certa no momento certo, de modo a ter um papel útil no processo de ensino e aprendizagem. Esta autora defende que os alunos se devem familiarizar com práticas avaliativas, formativas e sumativas, da mesma forma que se familiarizam com a apresentação da

versão final de um trabalho ou com a discussão de uma parte do seu processo de elaboração. Comparativamente com a avaliação sumativa externa, Shepard (2001) considera que a avaliação sumativa interna fornece informação mais válida e que a sua função é, entre outras, a de dar *feedback* para fora da escola, nomeadamente para os encarregados de educação. Já a avaliação externa deve ser uniforme em termos de conteúdo, deve ser implementada solicitando procedimentos estandardizados, mas deve ser utilizada apenas para avaliações em larga escala.

Harlen (2006) considera que, em vez de uma dicotomia entre avaliação formativa e avaliação sumativa, é possível ver estas duas funções da avaliação como dimensões de uma só avaliação. As evidências recolhidas ao longo das aulas, através de tarefas escritas tradicionais, como os testes, mas também da observação, do questionamento, da escuta de discussões informais entre os alunos, da correcção de trabalhos e da auto-avaliação permitem obter uma visão da progressão do aluno. Esta visão será tanto mais abrangente quanto maior o tempo ao longo do qual se realiza, e deve ainda ser expressa em termos de indicadores. Quando esta informação tem fins puramente formativos exige uma interacção integral entre o professor e o aluno, tomando este último um papel activo no progresso, sendo a informação obtida reportada em relação àquele aluno particular, ou seja, tem como referencial as capacidades desse aluno. Se a informação recolhida tiver propósitos sumativos, há que proceder à integração da informação recolhida reportando-a a critérios mais abrangentes e que são idênticos para todos os alunos, de forma a produzir uma graduação destes. Quando as finalidades da avaliação são puramente sumativas, então a avaliação visa a prestação de contas relativamente à aquisição de determinadas competências por parte dos alunos, sendo da responsabilidade última do professor. Neste caso, a avaliação compara o nível de cada aluno face a critérios ou *standards* previamente definidos. Harlen (2006) considera, então, que a informação recolhida com fins formativos pode servir fins sumativos, sendo para isso necessário proceder à descrição da progressão dos alunos e dos níveis alcançados em termos de categorias.

Um aspecto a ter em conta na análise da tensão entre a avaliação formativa e a avaliação sumativa está na dificuldade de um mesmo actor avaliativo, o professor, adoptar dois papéis distintos: o de recurso para a aprendizagem, com quem os alunos partilham as suas dúvidas e dificuldades; e o de juiz, perante o qual cada aluno pretende demonstrar o melhor desempenho possível, tentando dissimular as suas inseguranças e

dúvidas (Perrenoud, 1999). Para ultrapassar esta situação de papéis quase antagónicos para o professor, Perrenoud propõe que as duas modalidades avaliativas sejam implementadas por diferentes sujeitos, indo mesmo ao ponto de sugerir que os encarregados de educação sejam decisores acerca da transição ou retenção do seu educando, a partir da informação fornecida pelo professor, baseada na avaliação que fez.

Decorrente das tensões entre a função formativa e reguladora e a função selectiva e certificativa da avaliação, constata-se que os professores têm tendência, por pressão das avaliações externas ou das avaliações sumativas com elevado peso social e certificativo, a reduzir o currículo ensinado àquele que vai ser avaliado, a não diversificar as tarefas de sala de aula, implementando actividades próximas daquelas que constituem a avaliação sumativa externa. Desta forma, o processo de ensino passa a ser fortemente condicionado, quer em termos de conteúdo, quer em termos das formas de implementação utilizadas.

De acordo com Harlen (2006) os trabalhos de Maxwell e de Black *et al.*, referem-se a formas de utilizar a avaliação sumativa como apoio à aprendizagem, tais como o questionamento, o *feedback* através da simples atribuição de uma classificação, além da avaliação pelos pares e da auto-avaliação. Também são indicadas três formas principais de utilização do teste escrito em tarefas formativas de sala de aula: a revisão da matéria dada e a resolução de testes anteriormente realizados de forma a encontrar pontos fracos a ser mais trabalhados; a elaboração por parte dos alunos de testes, bem como a atribuição das cotações que cada questão deve ter e ainda a hetero-correcção dos testes.

No sentido de minimizar as tensões entre a avaliação sumativa e a avaliação formativa Coll, Martín e Onrubia (2001) defendem que no ensino obrigatório a avaliação sumativa tenha por objectivo a obtenção de informações úteis e relevantes para a tomada de decisões pedagógicas. Para tal deve ser realizado no final de processos de ensino-aprendizagem pouco extensos, utilizar os resultados para regular o processo de ensino e facilitar as estratégias de auto-regulação e controlo das aprendizagens pelos alunos. Os seus resultados devem estar pois sistematicamente vinculados à modificação e melhoria do ensino. No que respeita às tarefas e situações de implementação desta avaliação sumativa, elas devem implicar, tanto quanto possível, a participação dos

alunos, estar adequadamente contextualizadas e ser definidas de forma conjunta pelo professor e os alunos. Esta avaliação deve ainda assumir uma natureza criterial e não normativa, devendo incorporar formas de autoavaliação associadas à aquisição de estratégias de auto-regulação. A função comunicativa da avaliação sumativa deve procurar formas alternativas de transmitir os resultados aos encarregados de educação e outros elementos interessados, de modo a diminuir a componente certificativa da avaliação. Já a avaliação sumativa com carácter certificativo deve ser implementada apenas fora da escolaridade obrigatória podendo assumir o papel de uma avaliação externa (Coll, Martín, e Onrubia, 2001).

### **Práticas Avaliativas Mais Comuns**

Tomando por base duas consagradas revisões da literatura acerca da avaliação das aprendizagens, Crooks (1988) e Black e Wiliam (1998), ambas muito referenciadas na literatura, e ainda a leitura de outros autores, nomeadamente Stiggins e Conklin (1992), discute-se práticas e ambientes de avaliação nas salas de aula.

Para efeitos da revisão da literatura que efectuou acerca da avaliação em sala de aula, em 1988, Crooks considerou que esta se caracterizava por se basear em actividades que os alunos desenvolviam como parte integrante dos programas educativos, podendo ser efectuada dentro e fora da sala de aula. Incluem-se assim, tarefas como testes feitos pelos professores, testes baseados no currículo, questões orais dos alunos, uma vasta gama de actividades de *performance* cognitiva e psicomotora, abrangendo a avaliação aspectos motivacionais e atitudinais, além dos de aprendizagem.

Nessa altura, o autor constatou a existência de padrões na avaliação que se praticava em sala de aula: a) as actividades dos alunos sobre as quais recaía a avaliação constituíam uma parte substancial do tempo de aula - esse tempo era menor no 1º ciclo, aumentando progressivamente até ao seu máximo, no ensino secundário; b) coexistiam actividades avaliativas quer de carácter formal quer informal, constatando-se, no entanto, a supremacia dos testes escritos formais; c) fazia-se recurso a diferentes actividades avaliativas que variavam de acordo com o nível de escolaridade e com a disciplina; d) os aspectos afectivos eram avaliados de modo informal; e e) grande parte dos professores apresentava falta de formação em avaliação e os que a tinham não a consideravam útil em sala de aula, especialmente no 1º ciclo, onde recorriam essencialmente à observação e outras tarefas que não o teste.

Do seu trabalho, Crooks retirou conclusões a diferentes níveis. Por um lado, no que respeita às estratégias usadas pelos alunos, elas constituíam poderosos elementos preditivos dos resultados educacionais, pelo que é importante a forma como se procede à sua selecção e aplicação. No que respeita aos domínios avaliados, os professores davam a primazia aos conhecimentos e conteúdos académicos, seguindo-se-lhe a compreensão e o desempenho de rotinas e, apenas em terceiro lugar, a resolução de problemas, que implica, por parte do aluno, a transferência de conhecimentos e capacidades a novas situações. Estes aspectos pareceram indicar a necessidade e importância de ajudar os professores a implementar outras tarefas avaliativas além do teste escrito, que consistia na estratégia de recolha de informação avaliativa mais recorrente, apesar das suas limitações em abranger a diversidade de aprendizagens desejáveis para o desenvolvimento global do aluno. Finalmente, parecia haver uma ambiguidade entre expressões muito utilizadas, como *questões de elevado nível*, *competências de pensamento*, *resolução de problemas*, ambiguidade que resultava de uma definição pouco consistente.

Uma década mais tarde, em 1998, Black e Wiliam procederam a uma revisão da literatura sobre a avaliação formativa, tendo em comum com o trabalho de Crooks o facto de se basearem em situações vivenciadas em sala aula. Os autores constataram que esta avaliação se praticava de forma pouco frequente e frágil, fruto de razões várias, destacando-se o recurso a tarefas que incentivavam aprendizagens rotineiras e superficiais baseadas na memorização, a falta de reflexão sobre a avaliação entre os professores, a sobrevalorização da função classificativa e sub-valorização da função pedagógica da avaliação, além da tendência para usar critérios normativos, que promovem a competição, relativamente aos criteriosais, mais adequados ao desenvolvimento pessoal do aluno. Os autores concluíram ainda que as tarefas de avaliação solicitavam aos alunos capacidades de baixo nível de complexidade, principalmente de memorização, em vez de se proceder à especulação e reflexão crítica, que envolve actividades cognitivas complexas. A sobrevalorização da função sumativa da avaliação faz muitas vezes com que os alunos se habituem a ultrapassar as tarefas da forma menos arriscada possível, atribuindo, os professores, mais valor à quantidade do que à qualidade dos trabalhos realizados. A avaliação não era objecto de uma conceptualização pacífica, uma vez que os professores manifestavam conflitos pessoais e sociais com ela, tendo dificuldade em falar acerca da aprendizagem e sentindo-se

fortemente pressionados pela avaliação externa, o que tinha como resultado que as práticas avaliativas de sala de aula fossem muito influenciadas pelos requisitos locais e nacionais de certificação e de prestação de contas.

No que respeita ao meio mais utilizado para recolha de informação sobre a avaliação – o teste escrito – já em 1988, Crooks apresenta as seguintes críticas: a) um grande número de questões do nível mais baixo da taxionomia de Bloom; b) oitenta por cento do conteúdo baseado em conhecimentos; c) níveis de maior exigência, na taxionomia de Bloom, presentes em testes das disciplinas de matemática e francês.

Também Coll *et al.* (2001) referem dois efeitos negativos associados ao teste escrito. O efeito de estereotipia, em que o professor tem um juízo de valor relativamente ao aluno, que actua sobre a forma como o professor interpreta os resultados da avaliação, bem como na previsão que faz dos mesmos. E ainda o efeito de halo, em que a avaliação do professor é influenciada por aspectos de natureza emocional ou afectiva, isto é, pelas “impressões” que o professor tem acerca de cada aluno, muitas vezes baseadas em preconceitos culturais, sociais e outros.

Estes autores chamam ainda a atenção para o facto de a ordem de correcção de um teste ter influência sobre a mesma, ou seja, após a correcção de uma prova excepcional são mais evidentes as falhas durante a correcção da prova seguinte, havendo um efeito negativo acrescentado na correcção da segunda prova. O contrário, ou seja a sobrevalorização dos resultados de uma prova que se segue à correcção de uma prova fraca, também acontece.

Resumindo a opinião de vários autores, constata-se que o instrumento de avaliação ainda hoje privilegiado, o teste escrito, apresenta limitações para avaliar níveis complexos das taxionomias do desenvolvimento cognitivo, além de se mostrar ineficaz para avaliar os restantes domínios do desenvolvimento integral do aluno. A sobrevalorização deste meio de recolha de informação avaliativa terá por base uma concepção psicométrica ainda associada às actuais práticas avaliativas. O currículo tende, deste modo, a ser reduzido ao que vai ser avaliado pelo teste, as actividades de sala de aula são pouco interactivas, uma vez que a preocupação central dos professores é preparar os alunos para um formato de questionamento similar ao do teste escrito, cujos resultados constituem a base da avaliação final do aluno. Constata-se, pois, que a sobrevalorização do teste escrito na avaliação psicométrica condiciona o currículo



ensinado, a forma como é ensinado, limita a diversidade de actividades de ensino e aprendizagem e é muito limitada na determinação das aquisições realizadas pelos alunos, na medida em que não permitem avaliar níveis complexos de pensamento. O carácter normativo, em geral associado a esta avaliação, faz com que se promova mais a competição entre alunos do que o desenvolvimento individual global, podendo afirmar-se que a psicometria sobrevaloriza a função social – selectiva, certificativa e de classificação – relativamente à função pedagógica – de controlo, regulação e melhoria – da avaliação (Coll, Martín e Onrubia, 2001; Gipps e Cummings, 2003; Gipps e Stobart, 2003; Guba e Lincoln, 1989; Fernandes, 2005; e Shepard, 2001).

### **Avaliação Externa**

Como se depreende do que foi referido, a avaliação assumiu, desde cedo, uma função certificativa das aprendizagens realizadas, podendo ainda acumular a função selectiva para a transição a níveis superiores do ensino, e ainda a de controlo e, eventualmente, de regulação ao nível de todo o sistema de ensino. No entanto, a investigação tem mostrado que o carácter certificativo e selectivo não deve presidir em contexto de sala de aula, onde é necessária uma avaliação que actue como coadjuvante do ensino e das aprendizagens. Em termos dos contextos de abrangência, a avaliação pode ser implementada, pelo menos, nos quatro seguintes: o da sala de aula; o da escola; o do país e o de um grupo de países.

A avaliação que é implementada a nível de um país, ou de um conjunto de países é normalmente produzida e controlada por entidades com funções de supervisão, exteriores à escola, pelo que se costuma denominar de avaliação externa. Esta caracteriza-se por servir fins de certificação, selecção, motivação e controlo ao nível das escolas e/ou dos sistemas educativos. Muitas destas avaliações, não têm um carácter obrigatório, a não ser para a transição a níveis superiores de ensino, e o seu conteúdo, bem como os resultados, costumam ser tornados públicos. O suporte desta avaliação abrange áreas curriculares com recurso a tarefas escritas, incluindo, com frequência, questões de resposta curta ou fechada e ainda questões de escolha múltipla (Kellaghan e Madaus, 2003).

Para cumprir funções de certificação e de selecção, a avaliação externa tem de satisfazer *standards* de justiça e de comparabilidade elevados, ao mesmo tempo que o seu conteúdo deve representar adequadamente os diferentes domínios do currículo. A

avaliação externa é, em geral, de natureza sumativa. Já a avaliação interna é aquela que se faz a nível da escola e, em geral, os seus conteúdos, formato, critérios de correcção e situações de implementação, variam de escola para escola. Nesta acepção mais abrangente, que é aquela que adoptamos para efeitos do presente estudo, a avaliação interna inclui aquilo a que, no sistema de ensino português, se designa por *provas globais*, além de todo o tipo de avaliação levada a cabo em contexto de sala de aula. Na avaliação interna incluem-se avaliações de natureza diagnóstica, formativa e sumativa.

Constata-se que a avaliação assume carácter sumativo, quer na avaliação externa, quer na avaliação em contexto de sala de aula. No entanto, a investigação defende que em sala de aula deve presidir o carácter formativo da avaliação, embora seja também necessário obter resultados que se possam traduzir em classificações dos alunos. Simultaneamente, a avaliação externa, com fins certificativos, selectivos e de controlo ou prestação de contas, deve ter um carácter estandardizado, de forma a garantir justiça e credibilidade às decisões e comparações feitas a partir dos resultados. Pelo menos aparentemente, estas duas facetas da avaliação, sumativa e formativa, tal como já se referiu, parecem não ser fáceis de operacionalizar por um mesmo sujeito, o professor, sendo sempre necessário ter como base as funções que a avaliação pretende servir em cada momento.

### **Exames e Provas Aferidas**

A avaliação externa, caracterizada pela sua natureza sumativa, vem sendo implementada há mais de um século, em diferentes países do mundo, com vista à obtenção de informação que permita a monitorização das aprendizagens em larga escala, à prestação de contas por parte das escolas e à certificação dos níveis de proficiência dos alunos.

No caso nacional, os exames começaram a ser implementados ainda no final do século XIX, sendo na primeira metade do século XX que adquiriram um estatuto de avaliação *rigorosa*, através de mudanças que foram sofrendo ao longo do tempo, nomeadamente ao nível de quem os concebia e corrigia, da garantia ou não de anonimato dos alunos e da uniformização dos conteúdos (Fernandes, 2005). Parece certo que, a partir de 1947, a legislação exigiu que se implementasse uma avaliação de carácter externo e obrigatório, estandardizada e baseada no currículo nacional,

garantindo o anonimato dos alunos. No entanto, na década de 1970, assistiu-se à abolição desta avaliação externa, sob a forma dos chamados *exames nacionais*, que só voltaram a ser postos em prática em 1996, para os alunos do final do ensino secundário e, em 2002, no final do 9º ano, nas disciplinas de Língua Portuguesa e Matemática (Fernandes, 2005).

Dadas as suas funções, a avaliação externa tem sempre de ser estandardizada para garantir elevados graus de justiça e de equidade, independentemente do grau de abertura e diversidade das tarefas que a integram. É, por isso, muito importante assegurar a uniformização do conteúdo, da forma e das condições temporais e físicas da sua aplicação. Esta avaliação deve ainda ser bastante consistente com aquela que ocorre em sala de aula, uma vez que se deseja que ambas estejam alinhadas com o currículo. No que respeita às actividades que a constituem, Shepard (2001) chama a atenção para a necessidade de a avaliação externa dar ênfase especial às capacidades de raciocínio e de resolução de problemas, em busca de informação rigorosa e de resultados comparáveis, através de tarefas tecnicamente válidas e fiáveis, já que, a partir dos seus resultados, são tomadas decisões importantes. A mesma autora considera ainda que, embora seguindo procedimentos distintos, quer a avaliação em contexto de sala de aula, quer a avaliação externa devem ser guiadas pelos mesmos padrões de exigência e integrar os alunos no mesmo tipo de questionamento e de demonstração de proficiência, havendo partilha dos critérios de correcção das tarefas, de modo a aumentar a qualidade do que é avaliado (Shepard, 2001).

Outra avaliação de tipo externo utilizada em diferentes países é a chamada avaliação aferida. No nosso país, a sua criação legal deu-se em 1991. No entanto, por vicissitudes várias, só acabaram por começar a ser administradas a todos os alunos do quarto ano de escolaridade no ano de 2000 (Fernandes, 2005).

Este tipo de avaliação pode servir funções importantes ao nível da verificação da adequação e cumprimento do currículo; do controlo da qualidade do sistema educativo; da credibilização social dos diplomas escolares e da moderação externa das avaliações internas. Os seus resultados servem não apenas aos decisores políticos, mas ainda às escolas, aos professores, bem como aos investigadores. A aplicação da avaliação aferida é feita desde há décadas em diferentes países, merecendo os seus resultados bastante atenção por parte dos diferentes interessados. No entanto, a utilidade desta avaliação no

sentido da melhoria do sistema educativo depende de uma clarificação do seu papel, funções e natureza, por parte dos decisores políticos (Fernandes, 2005). Ainda de acordo com Fernandes,

Se bem enquadrada, devidamente organizada e levada muito a sério, a avaliação aferida pode ser uma forma relativamente barata de se obter informação de elevada qualidade sobre vários desempenhos do sistema e uma alavanca importantíssima para, entre outros efeitos, induzir boas práticas de ensino e de avaliação, promover investigação ou apoiar a tomada de decisões políticas e administrativas aos níveis local, regional e nacional (Fernandes, 2005, pp.124 e 125).

### **Principais Críticas aos Exames**

A avaliação externa tem sido, durante dezenas de anos, usada em muitos países para verificar a eficácia dos professores, das escolas e dos sistemas educativos. Assim, tem permitido quer a recolha de informação ao nível de inovações educativas e de programas em fase de experimentação, quer a identificação de escolas com um maior número de alunos desfavorecidos, para que estas recebam os apoios necessários e possíveis. No entanto, a avaliação externa tem ainda sido utilizada para a tomada de decisões acerca dos alunos, individualmente, com o intuito de fornecer orientações a professores e alunos em sala de aula, sendo este último aspecto o que tem sido alvo de mais críticas (Kellaghan, Madaus e Airaisian, 1980).

O uso de exames com consequências importantes no progresso escolar, na selecção ou na certificação dos alunos, tem apresentado consequências claras no contexto de sala de aula. Como principais contras deste tipo de avaliação podem ser referidos, por um lado o facto do currículo ensinado se restringir ao que é avaliado pela avaliação externa – constata-se que o uso dos chamados *exames* (meio de recolha de informação utilizado na avaliação externa) tem levado a que a maioria dos professores enfatize as capacidades académicas em detrimento daquelas que estão mais ligadas à realidade quotidiana, que são mais complexas, além de menos estanques em termos disciplinares. O currículo ensinado procura ser mais abrangente em termos de extensão, em detrimento da profundidade com que os assuntos são tratados e as actividades de preparação para o exame substituem, na maioria dos casos (Sturman, 2003), as actividades de ensino e aprendizagem da sala de aula (Kellaghan e Madaus, 2003; McMillan, Myran e Workman, 1999). Por outro lado, no que respeita à avaliação praticada em sala de aula, ela passa a consistir também num treino para os exames, tendo como recurso exames anteriormente implementados ou exames publicados. A

concepção de avaliação é aquela que reduz esta actividade à verificação de aquisições de conhecimentos, deixando de fora os restantes domínios, cujo desenvolvimento é necessário para um desenvolvimento integral do indivíduo. Simultaneamente, são vários os autores que referem a pobreza intelectual desta avaliação, que geralmente faz recurso a níveis de desenvolvimento intelectual baixos (Gipps, 1999; Landsheere, 1976; Shepard, 2001; Sturman, 2003). De acordo com Landsheere (1976) as aquisições feitas para dar resposta a este tipo de estratégia avaliativa são perdidas em cerca de 80%, nos quinze dias subsequentes à implementação da avaliação. Outro aspecto menos positivo desta avaliação é o facto de nos exames se verificar um predomínio de objectivos associados a estratégias de estudo e de desempenho em que o aluno se preocupa essencialmente com o juízo que fazem do seu desempenho, podendo chegar a procurar ocultar as suas dificuldades, relativamente a objectivos associados a estratégias de auto-regulação, de resolução de problemas e de aprendizagem efectiva (Kellaghan e Madaus, 2003). Ao nível dos alunos e das famílias as repercussões negativas do sistema de avaliação externa são geralmente o de gerarem uma grande ansiedade, uma vez que condicionam a transição de nível de escolaridade ou a obtenção de um certificado (Landsheere, 1976). Ainda ao nível dos alunos, o efeito possível de motivar pode ser substituído, naqueles que já obtiveram resultados pouco satisfatórios em situações anteriores, pela adopção de atitudes de desinteresse, raiva e alienação relativamente à avaliação, ou ainda a adopção de práticas que alterem os resultados para que estes sejam superiores àqueles que realmente traduzem as suas capacidades (Kellaghan e Madaus, 2003). Também ao nível dos professores se constatam efeitos negativos da avaliação, com uma perda progressiva da liberdade, autonomia e criatividade profissionais, transformando-se a profissão numa tarefa mecânica de transmissão de conhecimentos de forma não diferenciada (McMillan, Myran e Workman, 1999).

No entanto, registam-se aspectos positivos resultantes da implementação da avaliação externa. Podem-se referir os efeitos do “ensinar para o exame” como o recurso a actividades de revisão da compreensão dos alunos acerca do trabalho de investigação, de pesquisa e de projecto desenvolvidos, além da sua capacidade em aplicar os conhecimentos. Sturman (2003) constatou mesmo que a adopção desta abordagem, por parte dos professores, está relacionada com melhores resultados dos alunos.

## **Capítulo III**

### **METODOLOGIA**

#### **Introdução**

Este é um estudo de carácter descritivo, onde se pretende fazer uma síntese da informação existente num vasto conjunto de documentos, para que outros autores possam recorrer à leitura de um único documento síntese, em vez de se verem obrigados a ler um vasto conjunto de obras. Trata-se de um tipo de trabalho que vem ocorrendo cada vez com mais frequência, dada a quantidade de conhecimento produzido e a necessidade dos investigadores, e do público em geral, de estar a par do que se vem fazendo em diferentes domínios do saber através de textos de divulgação.

Deste modo, começa-se por explicitar o que se entende por uma síntese de investigação, através da análise de trabalhos de alguns dos autores mais significativos neste domínio, chegando-se a uma abordagem baseada nas concepções emergentes dessa leitura. Trata-se, então, de apresentar a metodologia utilizada neste trabalho, seguindo os diferentes passos que se consideram integrantes de uma síntese de investigações. Não se procederá a qualquer juízo sobre a qualidade das dissertações em análise, na linha de Suri e Clarke (1999), quando defendem a utilização de critérios de inclusão/exclusão de carácter conceptual e não valorativo. Posteriormente, serão explicitados, de forma detalhada, os passos dados na recolha das dissertações incluídas neste estudo, até à constituição do conjunto de dissertações que serão alvo desta síntese. Finalmente, proceder-se-á à leitura e análise focada das diferentes secções das dissertações, de modo a dar resposta a cada uma das questões de investigação, bem como ao problema.

Uma vez que a problemática das dissertações constituiu um critério fundamental para a inclusão ou não das mesmas no conjunto daquelas que foram objecto de análise e que, por outro lado, uma das questões de investigação consiste exactamente na caracterização das problemáticas emergentes da literatura das dissertações em estudo, procedeu-se, no presente capítulo, a uma reflexão sobre a definição de problema de investigação e suas especificações. De seguida, dado que outra das questões de investigação consiste na caracterização das metodologias utilizadas nas várias dissertações em estudo, é feita uma análise dos pontos de vista de alguns dos autores

que se destacam ao nível da investigação em ciências sociais, procedendo-se a uma síntese das principais ideias e à definição de conceitos e suas interpretações.

Para dar resposta às questões de investigação, procedeu-se à análise do conteúdo dos capítulos referentes à problemática e à metodologia, por vezes complementada com a leitura de outras secções das dissertações. Daí resultou a construção de tabelas de duas entradas, como se exemplifica no Quadro 1.

**Quadro 1. Esquema geral da análise feita às dissertações em estudo.**

Item em análise \ Dissertações em estudo	Dissertação A	Dissertação B	Dissertação C	
Observação	Registo	Registo	Registo	Síntese transversal relativa à observação
Questionário	Registo	Registo	Registo	Síntese transversal relativa ao questionário
Entrevista	Registo	Registo	Registo	Síntese transversal relativa à entrevista
Análise de conteúdo	Registo	Registo	Registo	Síntese transversal relativa à análise de conteúdo
	Síntese da dissertação A	Síntese da dissertação B	Síntese da dissertação C	Síntese global

Como pode ser observado, para cada dissertação e para cada item em análise, foi elaborada um pequeno registo que permite uma leitura comparativa entre as diferentes dissertações relativamente ao mesmo indicador e uma leitura de caracterização de cada dissertação relativamente à totalidade dos itens analisados.

O Quadro 2 apresenta, de uma forma sistematizada, a metodologia usada no presente trabalho. Da análise do quadro referido, constata-se a existência de cinco grandes etapas ao longo do presente trabalho. Uma primeira, cujo principal objectivo foi a constituição do conjunto de dissertações a ser alvo da análise e que se baseou em três tipos de pesquisa complementares, de modo a garantir a maior exaustividade de pesquisa e a abrangência da totalidade, ou quase totalidade, das dissertações que preenchessem os requisitos para integrar esta investigação. Num segundo momento procedeu-se à leitura de diferentes secções das dissertações de forma a obter a informação necessária à criação de categorias de uma base de dados, que constituiu a terceira fase deste estudo. A base de dados foi fundamental para a organização da informação de acordo com cada categoria, relativamente às diferentes dissertações, e de acordo com cada dissertação, relativamente a cada categoria. Assim, o passo seguinte

foi o de analisar e interpretar os dados já organizados, desta vez procurando a existência de relações ou não entre diferentes categorias e a emergência de aspectos de carácter relevante de acordo com as questões do estudo e outras que se vieram a identificar ao longo de todo o processo. Finalmente, a quinta fase do trabalho consistiu na elaboração de sínteses globais da informação analisada, que permitiram uma leitura quer de cada dissertação, transversalmente às diferentes categorias de análise, quer de cada categoria, transversalmente às diferentes dissertações abrangidas pelo estudo.

#### **Quadro 2. Esquema geral da metodologia utilizada na presente síntese de dissertações.**

Definição clara dos critérios de inclusão das dissertações a analisar na presente síntese	
Pesquisa das dissertações e constituição do conjunto de dissertações a analisar através de	pesquisa via Internet, nos sites de cada universidade com assento no Conselho de Reitores das Universidades Portuguesas (CRUP), e faculdades/institutos constituintes, e no site da Biblioteca Nacional (BN)
	contacto escrito através de uma carta tipo enviada ao responsável pelos serviços de documentação de cada instituição
	pesquisa através de contacto directo e presencial com as bibliotecas da Faculdade de Psicologia e Ciências da Educação (FPCE) da Universidade de Lisboa (UL), da Reitoria da UL, da Faculdade de Letras da UL, da Faculdade de Motricidade Humana (FMH) da Universidade Técnica de Lisboa (UTL) e na BN
Recolha, análise e interpretação dos dados através da leitura das seguintes secções de cada dissertação	leitura do título, do resumo e do índice
	leitura do capítulo onde se define o problema e se formulam as questões e/ou os objectivos da investigação
	leitura focada relativa aos procedimentos metodológicos
	leitura focada relativa à apresentação e discussão dos dados
Constituição de uma base de dados com categorias abrangentes (ao nível de aspectos bibliográficos, da problemática e da metodologia) que emergiram das leituras feitas.	
Análise e interpretação dos dados através de	apresentação de dados em quadros e gráficos
	análise focada do conteúdo dos dados sintetizados
	elaboração de quadros-síntese resultantes de cruzamentos da informação obtida
	análise resultante da integração e articulação da informação considerada relevante para as finalidades da investigação
Elaboração de sínteses globais da informação, de modo a dar resposta ao problema e às questões de investigação, através da integração de dois tipos de sínteses	transversais às diferentes dissertações e relativamente a cada categoria de análise
	de cada dissertação ao nível das diferentes categorias de análise

#### **A Síntese de Investigações Como Abordagem Metodológica**

Cada vez mais os cientistas sociais confiam nas sínteses de investigação dentro de um domínio específico, na medida em que é difícil um único estudo fornecer respostas generalizáveis e permitir a detecção de factores causais de um determinado efeito. Assim, as sínteses de literatura permitem a identificação de eventuais padrões, além de constituírem um meio de disseminação do conhecimento, emergente a partir de



um conjunto de estudos, constituindo também um exemplo de estruturação para pesquisas subsequentes (Marzano, 1998; Suri e Clarke, 1999). O valor deste tipo de investigação está patente nas palavras do matemático David Hilbert, (citado por Glass, McGaw e Smith, 1981) “A importância de um trabalho científico pode ser medida pelo número de publicações anteriores cuja leitura torna desnecessária” (p.11).

No domínio da tríade ensino, aprendizagem e avaliação, as sínteses de Natriello (1987), Nevo (1986), Crooks (1988) e Black e Wiliam (1998) são as mais referenciadas, servindo de base para trabalhos científicos e constituindo prova inequívoca da importância deste tipo de trabalho para o investigador. Nevo, na sua síntese de literatura de 1986, afirma expressamente que o tipo de revisão que faz é útil para diferentes abordagens teóricas e práticas, quer pelo esquema organizador que lhe serve de base, que se adequa a futuras revisões, quer ainda pelo conteúdo, que permite a leitura integrada de uma vasta quantidade de informação.

Na medida em que as sínteses de investigação assumem um papel cada vez mais presente na disseminação do conhecimento, a metodologia que lhes serve de base é de importância crucial. Nesse sentido, vários autores, como Bangert-Downs e Rudner (1991), Cooper e Lindsay (1998), Marzano (1998), Glass, McGaw e Smith (1981) e Suri e Clarke (1999), têm vindo a preocupar-se com a criação de metodologias específicas para a elaboração destas sínteses. De acordo com Marzano (1998), as revisões de literatura vêm-se fazendo desde há muito tempo, geralmente numa abordagem narrativa, em que o investigador procura sintetizar os resultados de um conjunto de estudos centrados num determinado domínio, procurando padrões emergentes. No entanto, o mesmo autor, corroborado por Cooper e Lindsay (1998), alerta para o facto da credibilidade das conclusões poder ser posta em causa por eventuais tendências do investigador tais como: concentrar a sua análise numa parte não representativa da totalidade dos estudos; basear-se em representações erróneas e falíveis dos resultados dos estudos; não tornar clara a metodologia usada na revisão, de modo a que o leitor possa fazer um juízo da validade das conclusões da investigação; e na falta de sentido crítico com que analisam as metodologias usadas em cada estudo. Já em 1974 Traveggia argumentava que os principais pontos de eventuais fraquezas associados ao método das sínteses de investigação se situavam nos seguintes níveis: selecção da pesquisa; recolha, indexação e codificação dos estudos; análise e

comparabilidade dos dados; acumulação dos dados comparáveis; análise da distribuição obtida; e apresentação dos resultados (Cooper e Lindsay, 1998).

A preocupação com a promoção da qualidade científica das revisões de literatura levou Glass, nos anos setenta do século passado a propor a expressão *meta-análise* como sendo um método de integração de pesquisa que consistia

(...) na integração e análise quantitativas dos resultados de todos os estudos empíricos relevantes num domínio e fazer emergir uma agregação quantitativa. Não só quantifica a dimensão do efeito, mas também identifica as suas potenciais variáveis moderadoras (Suri e Clarke, 1999, p.4).

O objectivo final era o de tornar revisões de literatura tão sistemáticas quanto uma investigação primária, de forma que a leitura dos seus resultados traduzisse a distribuição dos resultados individuais (Rudner, Glass, Evartt e Emery, 2002). Já para Bangert-Drowns e Rudner (1991) a *meta-análise* surge no sentido de ultrapassar as dificuldades de comparação de trabalhos de investigação que envolvam diferentes contextos, metodologias e estratégias metodológicas. Publicações como *Primary, secondary and meta-analysis* (Glass, 1976), *Relationship of class-size to classroom processes, teacher satisfaction and pupil affect: a meta-analysis* (Smith e Glass, 1977), *Meta-analysis - cumulating research findings across studies* (Hunter, Schmidt e Jackson, 1982), *Research synthesis and Meta-analysis* (Cooper e Lindsay, 1998), *Meta-analytic procedures for social research* (Rosenthal, 1984) e *Revisiting Methods of Literature Syntesis* (Suri e Clarke, 1999) tiveram, a partir da segunda metade dos anos setenta do século passado, uma grande importância no desenvolvimento metodológico das sínteses de investigações primárias. De um modo geral, estes autores concluíram que a síntese tradicional não permitia um trabalho de integração da literatura de forma rigorosa. A concepção de *meta-análise*, proposta por Glass em 1976, terá vindo ultrapassar as dificuldades que as sínteses apresentavam, através do recurso a meios estatísticos.

Para Suri e Clarke (1999), existem quatro métodos de síntese de investigação: as revisões narrativas tradicionais; as *meta-análises*; as sínteses das melhores evidências; e os métodos de síntese de investigação qualitativa.

À medida que a síntese de pesquisa e a *meta-análise* se foram afirmando, passaram a assumir características próprias que as distinguem uma da outra. A preocupação com a identificação dos pontos fracos, quer das sínteses de literatura, quer

da *meta-análise*, bem como a definição das características de cada abordagem metodológica, levaram Cooper e Lindsay (1998) a propor os passos a seguir numa e noutra, tal como se apresentam no Quadro 3.

**Quadro 3. Etapas da revisão de pesquisas e de meta-análises (adaptado de Lindsay e Cooper, 1998).**

Etapas da Revisão de literatura	Etapas da Meta-análise
<ul style="list-style-type: none"> <li>• Formulação do problema</li> <li>• Recolha do material</li> <li>• Avaliação do material</li> <li>• Análise e interpretação</li> <li>• Apresentação dos resultados</li> </ul>	<ul style="list-style-type: none"> <li>• Contagem das diferenças entre médias dos estudos em análise</li> <li>• Combinação das probabilidades entre os diferentes estudos</li> <li>• Estimativa da dimensão do efeito</li> <li>• Análise das diferenças entre as dimensões do efeito</li> </ul>

Para autores como Suri e Clarke (1999), a síntese de investigações é um processo dialéctico, hermenêutico, indutivo e interpretativo. Os critérios de inclusão de investigações numa síntese devem ser de carácter conceptual e, após a recolha das investigações primárias, deve proceder-se à sua codificação. Este último procedimento deve ter por base um esquema abrangente, de modo a adequar-se à natureza dos dados e aos resultados mais relevantes de cada investigação. Numa segunda fase, cada uma das categorias em estudo deve ser agrupada de acordo com aspectos sobre os quais se concentra. O processo de síntese decorre da inter relação das sínteses dos diferentes conjuntos e subconjuntos de itens em análise, através de uma abordagem indutiva e dinâmica. Foi essencialmente na linha da proposta de Suri e Clarke (1999) que se procedeu à presente síntese de dissertações de mestrado, cuja metodologia foi já explicitada e esquematizada no Quadro 1.

### **Critérios de Inclusão/Exclusão de Dissertações**

Como já foi referido no primeiro capítulo, o período temporal escolhido pretende abranger uma década, período de tempo considerado suficiente para a obtenção de um número significativo de dissertações de mestrado. Os seus limites, inicial e final, foram definidos de acordo com aspectos práticos ligados ao tempo necessário para que as dissertações estejam disponíveis ao público em geral, e ao tempo que seria necessário para proceder à sua análise e posterior organização do presente estudo. A dificuldade em estabelecer um critério para a selecção das instituições onde foram produzidas as dissertações em análise foi ultrapassada com a escolha da constituição do CRUP em Novembro de 2003. Assim, foram consideradas as dissertações no âmbito de mestrados em Universidades com assento no CRUP à data referida. Como se sabe, cada

Universidade abarca um conjunto mais ou menos vasto de Institutos e Faculdades, pelo que a pesquisa se estendeu, via Internet, às bibliotecas gerais de cada Universidade, às bibliotecas dos diversos Institutos e Faculdades que as compunham, tendo posteriormente sido enviada uma carta tipo ao responsável dos serviços de documentação de cada Universidade. O referido documento dava conta do objectivo do presente estudo, da pesquisa já efectuada através da Internet e solicitava apoio no sentido de ser colmatada alguma lacuna devido, por exemplo, a atrasos no processo de catalogação e disponibilização para consulta das obras. Nem todas as instituições responderam a este pedido e as cartas recebidas em resposta variaram em conteúdo e natureza, algumas com respostas estranhas e lamentáveis, como a da Universidade de Coimbra, mas outras com respostas muitíssimo positivas como a da Universidade Técnica de Lisboa ou a da Universidade da Madeira.

É de referir que, no último ano e meio, houve alguma mudança no que respeita à disponibilização das obras de cada biblioteca universitária através da Internet, existindo actualmente mais bibliotecas de centros de documentação de universidades *on-line*, e uma tendência para uniformizar os formatos das bases de dados das diferentes instituições. Tal não acontecia no momento em que foi realizada a pesquisa para a presente dissertação. A este respeito parece oportuno citar Estrela, Esteves e Rodrigues (2002) que ao procurarem fazer um levantamento dos estudos de mestrado e doutoramento na área da formação de professores, referiram que

Querendo ser objectivas, confrontámo-nos, por um lado, com a diversidade das bases de dados das bibliotecas e, por outro, com o facto das mesmas nunca serem completas e terem sempre largos meses de atraso nos seus registos. Há ainda a referir a diversidade de critérios de indexação ... (p.9).

De facto, através de conversa informal, obteve-se a informação de que a indexação das obras numa biblioteca é feita, na maioria das vezes, sem ter por base palavras-chave dadas pelos autores ou orientadores, e mesmo sem critérios claramente definidos. Desta forma, se compreendem as discrepâncias de catalogação das obras, de universidade para universidade, e até mesmo dentro de uma mesma instituição.

Numa síntese de investigações é da máxima importância a definição clara dos critérios de inclusão de cada obra, de forma a dar ao leitor elementos para que possa averiguar a validade da síntese. Deste modo, a presente análise teve como objecto todas as dissertações que obedecessem cumulativamente às seguintes condições:

1. Ser dissertações realizadas no âmbito de cursos de mestrado;
2. Terem sido concluídas entre os anos 1994 a 2003, inclusive;

3. Terem sido realizadas em universidades com assento no Conselho de Reitores das Universidades Portuguesas (CRUP), com os membros efectivos que este organismo apresentava em 2003, a saber: Universidade Aberta (UAB), Universidade dos Açores (UAÇ), Universidade do Algarve (UAL), Universidade de Aveiro (UAV), Universidade da Beira Interior (UBI), Universidade Católica Portuguesa (UCP), Universidade de Coimbra (UCO), Universidade de Évora (UEV), Universidade de Lisboa (UL), Universidade da Madeira (UMA), Universidade do Minho (UMI), Universidade Nova de Lisboa (UNL), Universidade do Porto (UP), Universidade Técnica de Lisboa (UTL) e Universidade de Trás-os-Montes e Alto Douro (UTAD);

4. Ter como tema central a avaliação das aprendizagens entendida como

(...) todo e qualquer processo deliberado e sistemático de recolha de informação, mais ou menos participado e interactivo, mais ou menos negociado, mais ou menos contextualizado, acerca do que os alunos sabem e são capazes de fazer numa diversidade de situações. Normalmente, este processo permite a formulação de apreciações por parte de diferentes intervenientes (incluindo os próprios alunos), acerca do mérito ou valor do trabalho desenvolvido pelos estudantes, o que, em última análise, deverá desencadear acções que regulem os processos de aprendizagem e de ensino. Ou seja, acções que contribuam decisivamente para que os alunos ultrapassem eventuais dificuldades e aprendam com mais gosto e com mais autonomia. Acções que ajudem os alunos a desenvolver processos de auto-avaliação e de auto-regulação, relativamente ao que é necessário aprenderem. Assim, neste sentido amplo, a avaliação das aprendizagens inclui a avaliação de conhecimentos, de desempenhos, de capacidades, de atitudes, de procedimentos ou de processos mais ou menos complexos de pensamento. Se quisermos, trata-se da avaliação de competências, ou da avaliação dos saberes em utilização ... (Fernandes, 2005, p.16)

Partindo desta definição ampla de avaliação das aprendizagens, foi decidido que as concepções e as práticas de professores, o desenvolvimento e análise de instrumentos, a análise histórica de práticas institucionais de avaliação, bem como a análise das concepções dos alunos constituem exemplos das áreas que foram contempladas nesta investigação. Ou seja, uma dissertação seria seleccionada para análise desde que o seu problema e/ou as suas questões de investigação tivessem implícita ou explicitamente fundamentalmente relacionadas com a avaliação das aprendizagens dos alunos. Ainda dito de outra maneira, dir-se-ia que, para responder às questões e/ou ao problema de investigação seria necessário ter em conta a avaliação das

aprendizagens dos alunos quer porque forneceria uma conceptualização indispensável para o estudo, quer ainda porque seria indispensável para a análise dos dados obtidos.

5. As dissertações deveriam ainda obedecer, cumulativamente, às seguintes condições:

5.1. Centrar-se na avaliação das aprendizagens ao nível do ensino básico e/ou secundário;

5.2. Incluir apenas as que envolvessem professores em exercício profissional efectivo excluindo, por isso, estudos que envolvessem de forma primordial professores estagiários e/ou professores em formação profissional;

5.3. Terem sido realizados no contexto do ensino público;

5.4. Terem sido realizados em Portugal;

A definição de avaliação das aprendizagens aqui utilizada é muito ampla, incluindo aspectos tão diversos como a participação dos alunos no processo avaliativo, conceitos associados à avaliação, conhecimentos e competências dos alunos, juízos de valor, processos de regulação do ensino e da aprendizagem, auto-avaliação e auto-regulação das aprendizagens, bem como processos avaliativos.

A selecção inicial das dissertações foi feita, na maioria dos casos, a partir do título, do problema e questões de investigação e outras fontes de informação (p. e., resumo, conhecimento pessoal dos autores e/ou dos orientadores, artigos ou livros de divulgação das dissertações). Esperava-se que este procedimento não oferecesse dificuldades especiais no processo inicial de selecção. Na verdade, na maioria dos casos assim aconteceu. Por exemplo, os títulos das dissertações de Campos (1996) e de Araújo (2002) ilustram, à partida e respectivamente, dois casos em que a selecção inicial é feita sem quaisquer dúvidas: *Concepções e práticas de professores sobre avaliação das aprendizagens – dois estudos de caso*; e *A avaliação no ensino básico – o que pensam alunos e professores*.

Do mesmo modo, a maioria das dissertações que, numa fase inicial, não foram incluídas no estudo foram excluídas com base no título e no problema de investigação. Num segundo momento, a selecção ou exclusão das dissertações foi ainda baseada numa análise mais cuidadosa dos conteúdos dos resumos, do enunciado do problema e das questões da investigação. Neste segundo momento foram encontradas situações em

que nem sempre foi fácil seleccionar claramente certas dissertações, de acordo com os critérios definidos. Nestes casos, foi necessário ler parte da dissertação, a fim de verificar até que ponto a avaliação das aprendizagens constituía ou não, um conceito central na referida dissertação. A título de exemplo, a dissertação de Sousa (2000) suscitou dúvidas devido ao seu título – *Avaliação das Dificuldades de Aprendizagem da Leitura: o difícil consenso de critérios*. Neste caso concreto, por se verificar uma preocupação central com os aspectos relacionados com as metodologias do ensino e aprendizagem da leitura, optou-se por não incluir a dissertação no estudo. Por outro lado, o título da dissertação de Figueiredo, (1997) *A formação integral do aluno no âmbito da disciplina de português: concepções e práticas de professores do ensino básico – estudo de caso*, parece indicar que a avaliação das aprendizagens seria, de algum modo, uma questão periférica. Na verdade, quer a apresentação do problema quer a formulação das questões de investigação, também apontavam nesse sentido. Todavia, após a leitura de algumas secções da dissertação, optou-se por inclui-la neste estudo, uma vez que, na verdade, a avaliação das aprendizagens acaba por ter uma centralidade de algum modo inesperada e que o título da dissertação não indiciava. Outra dissertação que levantou dúvidas quanto à sua inclusão ou não foi a de Santo (2002) por inquirir professores e formadores da formação em alternância nos centros de formação de professores. No entanto, dado que o número de professores participante (333) foi significativo, foi decidido seleccioná-la.

Uma outra dificuldade no processo de selecção foi o facto das dissertações não obedecerem a um formato comum, tanto em instituições diferentes como dentro de uma mesma instituição, podendo, por exemplo, o problema e questões de investigação serem formulados a meio da dissertação, no capítulo das metodologias, ou até mesmo na última secção do estudo.

Em suma, embora o critério basilar de selecção das dissertações tenha sido o de a avaliação das aprendizagens, na sua acepção mais ampla aqui definida, constituir parte substancial da dissertação, sendo isso claramente explicitado no problema e nas questões de investigação, trata-se de um processo marcado por alguma subjectividade, uma vez que os critérios podem ter interpretações distintas de leitor para leitor. No sentido de minimizar o problema da subjectividade no processo de selecção das dissertações elaborou-se uma tabela contendo o enunciado dos problemas e das questões de investigação. Esta tabela foi analisada por um painel de Professores da área das

Ciências da Educação. As respectivas opiniões permitiram chegar a uma selecção final, consensual, das dissertações que seriam objecto de análise nesta investigação (ver Anexo A).

### **O Processo de Recolha das Dissertações**

O processo de recolha das dissertações que satisfaziam as condições acima referidas foi efectuado com recurso a três tipos de pesquisa bibliográfica: a) Através da Internet; b) Através do contacto escrito com as instituições; e c) Através do contacto directo, presencial, com as instituições.

#### **Pesquisa Através da Internet**

Entre Junho de 2003 e Março de 2005, foram realizadas diversas pesquisas nos *sites* das universidades seleccionadas para o estudo. Em todos os casos, era através do *site* da universidade que se tentava ter acesso às bibliotecas ou aos serviços de documentação. Em algumas instituições não existia uma biblioteca geral da universidade, pelo que se tornou necessário consultar, primeiro, o *site* de cada uma das faculdades/institutos que integravam a universidade e, posteriormente, procurar aí a biblioteca ou serviço de documentação (a UL foi disto um exemplo). Noutros casos, a universidade é constituída por agrupamentos de faculdades/institutos, algumas vezes localizados em diferentes cidades do país, tendo cada um desses “pólos” uma biblioteca associada. São exemplos disso a UMI e a UCP, casos em que a pesquisa foi efectuada no *site* de cada uma dessas bibliotecas.

Foram, ainda, encontradas situações de instituições cujo *site* não referia a existência de qualquer serviço *on-line* de pesquisa bibliográfica, tendo-se procedido a pesquisas ulteriores, sempre dentro do período de cerca de dois anos, acima referido.

Dois importantes obstáculos podem não ter garantido uma pesquisa exaustiva. Por um lado, os programas de pesquisa bibliográfica que diferem de instituição para instituição, não permitindo uma uniformização de processos para os vários casos. Por outro lado, parâmetros como “palavras em assunto” ou “palavras-chave” serem definidos pelos funcionários dos serviços de documentação de cada instituição, o que



levanta, igualmente, problemas de uniformização de pesquisa, por não existir um critério geral para este tipo de catalogação.

Apesar destas dificuldades que, num certo sentido, são incompreensíveis, o processo de pesquisa através da Internet permitiu recolher a quase totalidade das dissertações incluídas neste estudo.

As palavras utilizadas na pesquisa em “palavras em assunto” e em “palavras-chave” foram *mestrado, educação, avaliação, ensino, aprendizagem, aprendizagens e concepções* de forma adjacente nas diferentes combinações possíveis, com duas ou mais palavras, no sentido de garantir que a pesquisa fosse exaustiva.

### **Pesquisa Através de Contacto Escrito com as Instituições**

Diversos contactos informais revelaram que o processo de catalogação e disponibilização das obras numa base de dados *on-line* decorre num período de tempo mais ou menos longo que, variando, embora, de instituição para instituição, pode mesmo ultrapassar os três anos (caso da Biblioteca Nacional, organismo que recolhe toda a produção ao nível de dissertações de mestrado e de teses de doutoramento do país). A fim de ultrapassar este obstáculo, procedeu-se ao contacto escrito com as universidades em estudo. Assim, a nove de Março de 2004, foi enviada uma *carta-tipo* registada, ao responsável pelos serviços de documentação de cada uma das universidades. Solicitou-se, então, informação acerca da existência, ou não, de dissertações de mestrado cujo tema se debruçasse sobre avaliação das aprendizagens e que, por demora no processo de catalogação, ainda não estivessem disponíveis na pesquisa *on-line*. Com esta carta pretendeu-se, adicionalmente, obter referência bibliográfica de dissertações defendidas em instituições que não tivessem disponível pesquisa bibliográfica *on-line* (casos da UBI e da UTAD).

Das quinze universidades contactadas foram obtidas as seguintes respostas:

- Por carta escrita (6) – UAÇ, UCO, UEV, UMI, UTL e UTAD.
- Por *e-mail* (4) – UAL, UAV, UMA e UNL.
- Não foi obtida qualquer resposta (5) - UAB, UBI, UCP, UL e UP.

É de referir que das dez instituições que responderam três não foram ao encontro do que era solicitado, parecendo ter havido alguma falha de comunicação, uma vez que remetiam a investigadora para a pesquisa no respectivo *site*, ou local, sendo, qualquer das opções, ineficaz, uma vez que se pretendia saber da existência de obras que ainda se encontrassem em processo de catalogação. Deve destacar-se a forma como a UTL acedeu ao pedido, através de uma carta, conjuntamente com o envio dos catálogos “Doutoramentos e Mestrados – Resumos das teses” relativos aos anos de 2001 e 2002 e de um CD-ROM com informação sobre os doutoramentos e mestrados concluídos nesta instituição entre 1989 e 2000. A preocupação em produzir estes catálogos parece ser um indicador de que a UTL tem, relativamente às restantes, uma maior preocupação em tornar acessível a investigação que aí se produz, no âmbito de doutoramentos e mestrados. Em contrapartida, a resposta da parte dos serviços da Universidade de Coimbra foi de algum modo surpreendente, uma vez que não só não deu resposta à questão colocada na carta enviada, como fez sugestões de pesquisa, à investigadora, que poderão, em última análise, questionar as reconhecidas vantagens da chamada sociedade da informação e do conhecimento.

Este procedimento de recolha, apesar das variações observadas na cooperação obtida por parte das diferentes instituições, permitiu confirmar e melhorar a selecção de dissertações feita com os procedimentos anteriores.

### **Pesquisa Através do Contacto Directo, Presencial, com as Instituições**

A recolha local e presencial teve lugar nas bibliotecas da Faculdade de Psicologia e Ciências da Educação da UL, da Reitoria da UL, da Faculdade de Letras da UL da Faculdade de Motricidade Humana da UTL, na Biblioteca da Faculdade de Ciências da Universidade de Lisboa e na Biblioteca Nacional. Neste último caso, foi constatado que, ao longo de 2004, as dissertações disponíveis, concluídas após 2000, eram em número muito reduzido. Através de conversa informal com alguém exercendo funções nesta instituição, soube-se que, por falta de recursos humanos, se tinha tomado a decisão de não catalogar dissertações recebidas a partir de 2000. Em 2005, apresentou-se uma reclamação, por escrito, no livro amarelo desta instituição onde, entre outros aspectos, foi referida a morosidade do processo de catalogação deste tipo de obras, cuja divulgação é da máxima importância para o desenvolvimento da investigação. Em resposta, recebeu-se uma carta, onde, não tendo sido feita qualquer

referência à morosidade do processo de catalogação, se remetia a investigadora para a recolha junto de cada uma das várias universidades do país.

Em suma, e como já foi referido, através destas formas de recolha bibliográfica foi possível obter listas de dissertações que, desde logo, foram objecto de uma primeira selecção com base no título. Assim, verificou-se que existiam dissertações realizadas no âmbito de mestrados, em universidades pertencentes ao CRUP e concluídas no período temporal definido para este estudo, cujo título indiciava que não cumpriam os critérios definidos para a sua integração no conjunto de estudos a analisar. Adicionalmente, foi necessário proceder à consulta de cada uma das dissertações que não tinham sido eliminadas pelo sistema acima descrito. Por fim, obteve-se a lista definitiva de dissertações que foram objecto de análise nesta investigação e que se apresenta no Anexo A.

### **A Base de Dados**

A informação recolhida a partir de cada dissertação foi organizada numa base de dados construída com o *software Microsoft Access*.

Procedeu-se ao levantamento de informação dentro das seguintes categorias:

1. *Aspectos Bibliográficos* que incluía, nomeadamente, o ano de conclusão do mestrado, a instituição de acolhimento, o curso de mestrado no âmbito do qual tinha sido realizada a dissertação, o autor, o orientador, o título e a cota atribuída em diferentes bibliotecas;
2. *Aspectos relativos ao Problema e às Questões de Investigação*, onde se identificaram o problema e as questões de cada dissertação, recolhidos sob a forma de citação ou, no caso de ultrapassarem os 255 caracteres, na forma de síntese;
3. *Aspectos relacionados com a Metodologia*, onde se caracterizou a amostra, através do número de indivíduos e/ou documentos em estudo, o nível de ensino (ciclo e ano de escolaridade) e abrangência geográfica. Também nesta categoria, foram identificadas as estratégias e procedimentos adoptados para a recolha da informação.

## **Aspectos Tidos em Conta na Análise das Dissertações**

### **Problemática de uma Investigação**

O primeiro passo de uma investigação consiste na identificação de um tema que, de alguma forma, suscita o interesse ou constitui uma preocupação para quem o escolhe. Dentro desse tema, é necessário identificar o problema que se pretende abordar, uma vez que este constitui o motor de toda a pesquisa. Para Fraenkel e Wallen (1990)

Um problema de investigação é (...) um problema que alguém gostaria de investigar. Um problema pode ser qualquer coisa que se considera mal resolvida, uma dificuldade de qualquer tipo, uma situação que precisa de ser mudada, qualquer coisa que não esteja a correr tão bem como devia (p. 21).

Gay (1981) acrescenta ainda a necessidade do problema ser investigável através da recolha e análise dos dados, considerando impraticáveis investigações acerca de problemas filosóficos ou éticos. O problema de investigação pode surgir de uma lacuna do conhecimento, ou simplesmente da necessidade da sua melhoria. De acordo com Giddens (2004)

Nenhum trabalho de investigação se encontra isolado. Os problemas de investigação surgem como parte do trabalho que se está a efectuar; um projecto de investigação pode facilmente conduzir a outro, pois sugere assuntos que o investigador não tinha considerado previamente. (...) também podem surgir com a leitura do trabalho de outros investigadores publicados (...) ou tomada de consciência da existência de tendências específicas na sociedade (p.644)

Para Tuckman (1978) “ A exposição do problema deve identificar, se possível, todas as variáveis independentes, moderadoras e dependentes, e deve indagar, na forma de uma pergunta, acerca das suas relações” (p.311). Este enunciado deve ainda ser feito na secção introdutória do plano de investigação e do seu relatório (Gay, 1981; Tuckman, 1978). Para além das referidas características, a forma como o problema é enunciado tem de obedecer a critérios que variam ligeiramente de autor para autor. De acordo com Van Dalen (1979) e Fraenkel e Wallen (1990), a apresentação do problema pode ser feita sob a forma de questão ou de afirmação e deve haver uma clara referência aos principais aspectos a ser observados.

Um bom problema de investigação tem ainda de ser pertinente, ou seja, valer a pena ser estudado, conduzir a um resultado que constitua uma mais valia em termos de produção de conhecimento, da compreensão de uma preocupação social, da determinação da eficácia de um programa, da implementação e estudo de um novo programa ou da resolução de um problema concreto. Deve ainda ser exequível em

termos de tempo e de recursos materiais, orçamentais e humanos, para que possa ser levado a bom porto. Para além das referidas características, a forma como o problema é enunciado obedece a critérios que variam ligeiramente de autor para autor. A título de exemplo, Tuckman (1978) defende as seguintes características para o enunciado do problema: questionar a relação entre duas ou mais variáveis; ser exposto de forma clara e inequívoca; ser verificável através de métodos empíricos e não conter qualquer posição moral ou ética. Por seu lado, Van Dalen (1979) considera que o problema pode ser de “largo espectro e não directamente testável” (p.179).

Numa síntese da teoria revista, pode dizer-se que um problema deve ter as seguintes características: a) ser claro em termos de linguagem/formulação; b) ser constituído por um número reduzido de frases; c) identificar o(s) objecto(s) de estudo; e d) ser investigável.

### **Breve Síntese de Algumas Questões Metodológicas**

No sentido de proceder à análise das dissertações torna-se importante fazer alguma reflexão em torno das metodologias utilizadas em investigação no domínio das ciências sociais e, mais especificamente, em educação. Proceder-se-á, seguidamente, a uma reflexão sobre as abordagens qualitativa, quantitativa e mistas, bem como sobre as estratégias de recolha de informação mais utilizadas em cada uma delas. O objectivo é o enquadramento de cada dissertação num tipo de design metodológico, a partir da análise dos procedimentos utilizados em cada uma.

Quase invariavelmente, o investigador faz as suas opções metodológicas de acordo com o paradigma com que mais se identifica, que é fortemente condicionado pela sua formação académica. Consciente desta tendência, Patton (1990) afirma:

Eu prefiro o pragmatismo (Patton, 1988c) ao perfilhar de uma única abordagem em termos de paradigma. O meu objectivo ao escrever acerca de paradigmas de investigação alternativos foi aumentar as opções disponíveis para os avaliadores, e não substituir um paradigma limitado por outro igualmente limitado, embora diferente. A importância de compreender paradigmas de investigação alternativos é sensibilizar investigadores e avaliadores para as formas como a sua parcialidade metodológica, derivada das suas experiências de socialização académicas, podem reduzir a sua adaptabilidade e flexibilidade metodológicas. O objectivo de descrever o modo como os paradigmas tradicionalmente operam no mundo real da investigação é libertar os avaliadores do contrato de obediência a um único paradigma (p.38).

De facto, do debate paradigmático feito por inúmeros autores de entre os quais se destacam Gipps, (1999), Gipps e Cumming, (2003), Guba e Lincoln (1989 e 1994) e

Shepard (2001) constata-se a existência de uma dicotomia metodológica, aparentemente associada a uma dicotomia de paradigmas científicos, em que num extremo existe uma concepção positivista, eventualmente emanada e mais adequada no âmbito das ciências ditas exactas e naturais e, num outro extremo, existe uma concepção construtivista e interpretativa, eventualmente mais adequada às práticas de investigação nas áreas das ciências sociais e humanas. A abordagem quantitativa aparece, em geral, associada ao positivismo, enquanto que a qualitativa aparece associada ao construtivismo.

Autores como Bogdan e Biklen (1994), Cohen e Manion (1994), Gay (1981), Guba e Lincoln (1994), Patton (1990) e Tuckman (1978) dedicaram muito do seu trabalho à reflexão sobre as abordagens metodológicas no âmbito das ciências sociais e humanas, pelo que, tal como Patton defende na citação feita no início desta subsecção, se torna pertinente proceder a uma rápida síntese das suas principais ideias.

Bogdan e Biklen (1994) apresentam uma sistematização bastante exaustiva das principais características das abordagens qualitativa e quantitativa, incluindo itens de comparação que vão desde conceitos chave associados a cada uma das abordagens, até a aspectos como o tipo de dados, de amostra e de procedimentos de investigação. O Quadro 4 apresenta as principais características das abordagens qualitativa e quantitativa, de acordo com os autores referidos.

De acordo com Guba e Lincoln (1994), o paradigma positivista está associado a uma metodologia experimental, onde se procuram refutar ou confirmar hipóteses. Na busca de uma adaptação à natureza dos fenómenos sociais, assistiu-se a uma mudança de paradigma no sentido de incorporar, na sua metodologia, a pesquisa em ambientes mais contextualizados, a recolha de informação situacional, a solicitação do ponto de vista émico (do observado). Para estes mesmos autores, o paradigma que mais se opõe ao positivista é o construtivista, associado a metodologias hermenêuticas e dialécticas, através das quais se procura um consenso entre as diferentes participações individuais, chegando-se a uma construção mais informada e sofisticada do que qualquer outra que a tenha precedido (Guba e Lincoln, 1994). Trata-se, pois, de uma metodologia que vai sendo construída à medida que a investigação decorre, com base na constante adaptação do contexto em estudo, sendo por essa razão denominada de *grounded theory*.

**Quadro 4 Principais características das abordagens qualitativa e quantitativa (adaptado de Bogdan e Biklen, 1994).**

<b>QUALITATIVA</b>	<b>QUANTITATIVA</b>
<i>Conceitos chave associados</i>	<i>Conceitos chave associados</i>
<ul style="list-style-type: none"> <li>• Significado</li> <li>• Compreensão</li> <li>• Definição da situação</li> <li>• Construção social</li> </ul>	<ul style="list-style-type: none"> <li>• Variável</li> <li>• Hipóteses</li> <li>• Validade</li> <li>• Significância estatística</li> <li>• Replicação</li> </ul>
<i>Objectivos</i>	<i>Objectivos</i>
<ul style="list-style-type: none"> <li>• Descrever realidades múltiplas</li> <li>• Desenvolver a compreensão</li> </ul>	<ul style="list-style-type: none"> <li>• Descrição estatística</li> <li>• Encontrar relações entre variáveis</li> </ul>
<i>Elaboração das Propostas de Investigação</i>	<i>Elaboração das Propostas de Investigação</i>
<ul style="list-style-type: none"> <li>• Breves</li> <li>• Especulativas</li> <li>• Descrição geral da abordagem</li> </ul>	<ul style="list-style-type: none"> <li>• Extensas</li> <li>• Detalhadas e específicas nos objectivos</li> <li>• Formulação de hipóteses</li> </ul>
<i>Dados</i>	<i>Dados</i>
<ul style="list-style-type: none"> <li>• Descritivos</li> <li>• Documentais pessoais</li> <li>• Notas de campo</li> <li>• O discurso dos sujeitos</li> <li>• Documentos oficiais e outros</li> </ul>	<ul style="list-style-type: none"> <li>• Quantitativos</li> <li>• Codificação quantificável</li> <li>• Contagens, medidas</li> <li>• Estatística</li> </ul>
<i>Amostra</i>	<i>Amostra</i>
<ul style="list-style-type: none"> <li>• Pequena</li> <li>• Não representativa</li> </ul>	<ul style="list-style-type: none"> <li>• Ampla</li> <li>• Selecção aleatória</li> </ul>
<i>Técnicas ou métodos</i>	<i>Técnicas ou métodos</i>
<ul style="list-style-type: none"> <li>• Observação</li> <li>• Estudo de documentos vários</li> <li>• Observação participante</li> <li>• Entrevista aberta</li> </ul>	<ul style="list-style-type: none"> <li>• Inquéritos</li> <li>• Entrevista estruturada</li> <li>• Observação estruturada</li> <li>• Experimentais</li> <li>• Quase experimentais</li> </ul>
<i>Relação com os sujeitos</i>	<i>Relação com os sujeitos</i>
<ul style="list-style-type: none"> <li>• Empatia</li> <li>• Contacto intenso</li> </ul>	<ul style="list-style-type: none"> <li>• Curta duração</li> <li>• Distante</li> </ul>
<i>Instrumentos</i>	<i>Instrumentos</i>
<ul style="list-style-type: none"> <li>• Gravador</li> <li>• Transcrição</li> </ul>	<ul style="list-style-type: none"> <li>• Inventários</li> <li>• Questionários</li> </ul>

Patton (1990) refere como vantagens dos métodos quantitativos o facto de permitirem medir a reacção de um número elevado de indivíduos a um número limitado de efeitos, possibilitando o tratamento estatístico dos dados recolhidos e a apresentação dos resultados de forma sucinta e passível de generalização. Para o efeito, será necessário recorrer a medidas estandardizadas, de modo a traduzir a diversidade de experiências e perspectivas dos indivíduos, num número limitado de categorias, às quais é associada uma numeração. O objectivo final da pesquisa quantitativa será, em última análise, a produção de leis de causa-efeito generalizáveis, que permitam a previsão e o controlo dos fenómenos (Gay, 1981 e Tuckman, 1978). Por outro lado, as principais

vantagens dos métodos qualitativos são, de acordo com Patton (1990), a produção de um conjunto de informação detalhada acerca de um número reduzido de indivíduos ou casos, permitindo uma melhor compreensão das situações estudadas, apesar das perdas ao nível da generalização dos resultados. A abordagem qualitativa permite o estudo em detalhe e profundidade de um número de situações, aproximação ao campo de trabalho sem categorias pré-estabelecidas e a análise contextualizada da informação.

Em síntese, e a partir dos autores referidos, podemos afirmar que os principais pressupostos subjacentes à abordagem quantitativa são:

1. A realidade existe de forma independente relativamente ao observador e pode ser estudada de forma objectiva;
2. O observador deve manter uma postura de distanciamento face ao objecto de estudo;
3. O conhecimento é objectivo, dada a neutralidade do observador, garantida pelas estratégias utilizadas;
4. A investigação parte de hipóteses que serão testadas no sentido de se encontrar leis de relação de causa-efeito;
5. O objectivo último da investigação é a elaboração de teorias gerais que permitam a compreensão, explicação e previsão dos fenómenos.

Por oposição, os principais pressupostos subjacentes à investigação qualitativa são as seguintes:

1. A realidade é situacional e subjectiva, por consistir nas percepções dos diferentes indivíduos;
2. O observador interage com os indivíduos, de forma a diminuir a distância entre observador e observado;
3. O observador tem em conta a impossibilidade de total isenção de valores da pesquisa, recorrendo a formas de a minimizar;



4. Uma boa investigação exige a triangulação de fontes, de investigadores, de estratégias de recolha, de tempos, para além das competências e credibilidade científica do próprio investigador;
5. A investigação é contextualizada;
6. A pesquisa não necessita de ter hipóteses pré-estabelecidas, emergindo categorias de interesse ao longo do contacto com os informantes;
7. O objectivo final da investigação é procurar padrões que ajudem a investigar aquele fenómeno concreto.

Para além destas duas abordagens “puras”, Patton (1990) salienta as vantagens da flexibilidade que as abordagens mistas permitem. Na verdade, de acordo com este autor,

A importância de compreender paradigmas de investigação alternativos é sensibilizar os investigadores e avaliadores para as formas como os seus preconceitos metodológicos, derivados da experiência de socialização das suas disciplinas, pode reduzir a sua flexibilidade e adaptabilidade metodológica (Patton, 1990, p.38).

Este autor considera as seguintes abordagens mistas:

1. Plano experimental – dados qualitativos – análise de conteúdo
2. Plano experimental – dados qualitativos – análise estatística
3. Inquérito naturalista – dados qualitativos – análise estatística
4. Inquérito naturalista – dados quantitativos – análise estatística

### **Planos de Investigação**

Para efeitos do presente estudo, considera-se que o plano de uma investigação consiste na sua estrutura básica, ou seja, ele é dependente da natureza do problema e do objecto de estudo, por um lado e, por outro, condiciona as opções ao nível dos estudos em análise, da existência ou não de grupos de controlo e dos procedimentos e estratégias de investigação a adoptar. Deste modo, podem ser considerados alguns planos básicos, a partir dos quais é possível encontrar inúmeras variantes (Gay, 1981 e Patton, 1990).

De acordo com Tuckman (1978),

Investigação é a tentativa sistemática de fornecer respostas a questões. Tais respostas podem ser abstractas ou gerais como acontece no caso da *investigação pura*, ou podem ser muito concretas e específicas, como acontece na *investigação aplicada*. Em ambos os casos, o investigador apresenta factos e depois formula uma generalização baseada na interpretação desses factos (p.1).

Esta definição revela uma forte tendência positivista, na concepção muito experimental de investigação, onde se procura responder a questões pré-estabelecidas e produzir conhecimento generalizável. O plano da investigação é a estrutura de todo o processo de investigação, que vai depender de opções a diversos níveis. Tuckman defende a existência de cinco planos gerais – pré-experimental, experimental, factorial, quasi-experimental e *ex post facto* – dentro dos quais se podem definir planos mais específicos.

Fraenkel e Wallen (1990) partem da definição de investigação como qualquer tipo de “estudo cuidadoso, sistemático e paciente dentro de um domínio do conhecimento que seja implementado no sentido de determinar factos e princípios” (p.7), contrastando com Tuckman, na medida em que não se prende ao carácter experimental da investigação, de forma exclusiva, dando-se ênfase à produção de conhecimento factual ou geral. De acordo com estes autores, podemos recorrer a seis metodologias de investigação, às quais correspondem diferentes tipos de planos de trabalho, como se apresenta no Quadro 5.

**Quadro 5. Planos de investigação adaptados de Fraenkel e Wallen (1990)**

Metodologia	Características
Investigação Experimental	Procura influenciar variáveis;  Testa hipóteses acerca de relações de causa-efeito.  O controlo por parte do investigador faz-se sentir mais do que nas outras formas de investigação.
Investigação Correlacional	Estuda as relações entre duas ou mais variáveis, sem procurar influenciá-las.
Investigação causal-comparativa	Procura determinar a causa ou consequências de diferenças já existentes entre ou dentro de grupos de indivíduos.
Investigação por Inquérito	O investigador questiona um vasto grupo de indivíduos acerca de um determinado aspecto.
Investigação Qualitativa	Quando se procura um conhecimento em profundidade através de uma abordagem holística da situação particular.  Maior ênfase na descrição detalhada sobretudo o que acontece numa actividade ou situação, ou ao descrever atitudes ou comportamentos.
Investigação Histórica	Centra-se unicamente em acontecimentos passados.

Gay (1981) apresenta uma classificação de investigação de acordo com a sua finalidade, dividindo-a nos seguintes tipos: pura; aplicada; educacional; investigação e desenvolvimento; e investigação-acção. O mesmo autor apresenta ainda uma classificação de investigação de acordo com o método utilizado, que é sintetizada no Quadro 6.

Patton (1990), baseando-se em Cronbach, refere-se à avaliação como uma forma de arte, afirmando:

A arte da avaliação inclui criar um plano e recolher informação que seja adequada a uma situação específica e à tomada de decisões num contexto particular. Na arte não existe um *standard* único e ideal. (...) Não existe um plano que seja único e mais adequado para uma avaliação ... (p.13).

De acordo com esta visão, em vez de propor, *à priori*, planos de investigação que possam ser ou não adequados aos dados recolhidos, só após uma primeira análise dos problemas e questões de investigação, da caracterização da população e das estratégias metodológicas utilizadas serão delineados os planos considerados para efeitos de organização da investigação analisada.

**Quadro 6. Classificação da investigação quanto aos métodos utilizados.**

Método	Em que consiste	Procedimentos possíveis
Histórico	Na recolha sistemática e avaliação objectiva de dados relacionados com eventos passados, com o intuito de testar hipóteses relacionadas a causas, efeitos ou linhas de força desses eventos, que possam ajudar a explicar os acontecimentos actuais e prever acontecimentos futuros.	
Descritivo	Na recolha de dados para testar hipóteses ou dar resposta a questões relacionadas com a situação actual do objecto de estudo.	Questionários Entrevistas
		Gravação áudio e/vídeo Grelhas de registo da observação Questionários Entrevistas Notas de campo
Correlacional	Na recolha de dados para determinar como e até que ponto existe uma relação entre duas ou mais variáveis quantificáveis.	
Causal-comparativo	Descreve condições, causas ou razões para uma determinada situação actual ou passada	
Experimental	No teste de hipóteses de tipo causa-efeito, com vista à criação de conhecimento generalizável.	

(Adaptação do capítulo 6 de Gay, 1981)

Para efeitos do presente trabalho, e com base na reflexão realizada em torno destas questões por Fraenkel e Wallen (1990), Gay (1981) e Patton (1990), foram considerados quatro planos de investigação – estudo de caso, investigação por

sondagem, investigação histórica e outros. O critério fundamental para a categorização do plano de cada dissertação será a explicitação dada pelo próprio autor, sempre que isso acontece. Os restantes critérios característicos de cada plano de investigação são os seguintes:

1. Estudo de caso – amostras de reduzida dimensão, com um número normalmente inferior a dez elementos, recurso a formas de recolha de informação que implicam uma maior proximidade entre o investigador e o “terreno”, como a entrevista e a observação.
2. Investigação por sondagem – amostras numerosas, e recurso a formas de recolha de informação que não exigem um contacto próximo entre o investigador e o “terreno”, como o inquérito por questionário, que pode ser enviado por correio, não existindo, necessariamente, contacto directo entre o investigador e os inquiridos.
3. Investigação histórica – estudos baseados exclusivamente ou quase exclusivamente, na análise de documentos, com vista a traçar percursos do aspecto em estudo, ao longo de um período de tempo.
4. Outros – todos os planos que não se insiram nos anteriormente descritos.

## Capítulo IV

### APRESENTAÇÃO E ANÁLISE DOS DADOS

#### **Introdução**

Neste capítulo, apresentam-se, analisam-se e discutem-se os dados referentes às respostas às três questões de investigação. Com o intuito de caracterizar os contextos das investigações analisadas, procedeu-se a uma abordagem segundo diferentes ângulos de observação, a saber: instituição de acolhimento do curso; cursos de mestrado no âmbito dos quais as dissertações foram realizadas; fontes de informação utilizadas; contextos disciplinares; níveis e anos de escolaridade sobre os quais incidiram as dissertações e a localização geográfica das escolas onde foram recolhidos os dados.

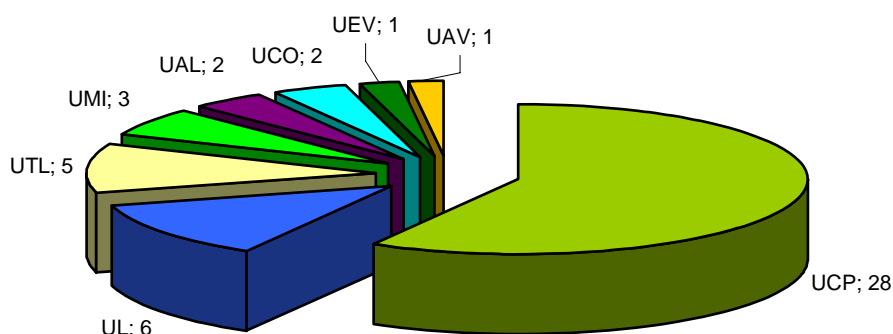
Em seguida, procedeu-se à caracterização dos problemas e questões de cada dissertação. Posteriormente, procedeu-se à caracterização dos procedimentos metodológicos, tendo em conta os seguintes indicadores: participantes; escolas envolvidas; e estratégias de recolha de informação utilizadas. A partir destes dados, foi possível categorizar as dissertações analisadas, de acordo com os métodos de investigação utilizados, discutidos e apresentados no capítulo referente à Metodologia.

#### **Caracterização Geral dos Contextos das Investigações Realizadas**

Para efeitos desta investigação, e com o intuito de dar resposta à primeira questão de investigação - *Como é que se poderão caracterizar as investigações realizadas?* - foram identificadas, recolhidas e analisadas 48 dissertações (Anexo A), cuja distribuição relativamente às instituições onde foram realizadas se encontra no Gráfico1.

Da análise deste gráfico constata-se que 39, ou seja, mais de 80% da totalidade de dissertações analisadas, foram realizadas em universidades localizadas em Lisboa – UCP, UTL e UL. Neste conjunto, verifica-se ainda que mais de dois terços – 28 dissertações – foram realizadas na UCP. Foi também possível concluir que as dissertações elaboradas em cursos da UCP foram, na sua quase totalidade (24 de 28 dissertações), feitas na Faculdade de Ciências Humanas (FCH). As quatro dissertações realizadas na UCP a partir de 2000, inclusive, tiveram como instituição de acolhimento o Instituto de Educação (IE).

**Gráfico 1. Número de dissertações seleccionadas em cada Universidade.**



Das seis dissertações realizadas na UL, quatro foram-no na Faculdade de Psicologia e Ciências da Educação (FPCE), uma em cada um dos anos de 1995, 1999, 2000 e 2001. As duas restantes foram realizadas em 1998 e em 2000 na Faculdade de Ciências (FC) da mesma universidade.

No caso da UTL, as cinco dissertações analisadas neste estudo foram realizadas na Faculdade de Motricidade Humana (FMH), tendo sido concluídas duas em 1997, uma em 1998, uma em 2002 e a última em 2003. Verifica-se, neste caso, um período de quatro anos sem que tenha sido realizada qualquer dissertação no âmbito da avaliação das aprendizagens.

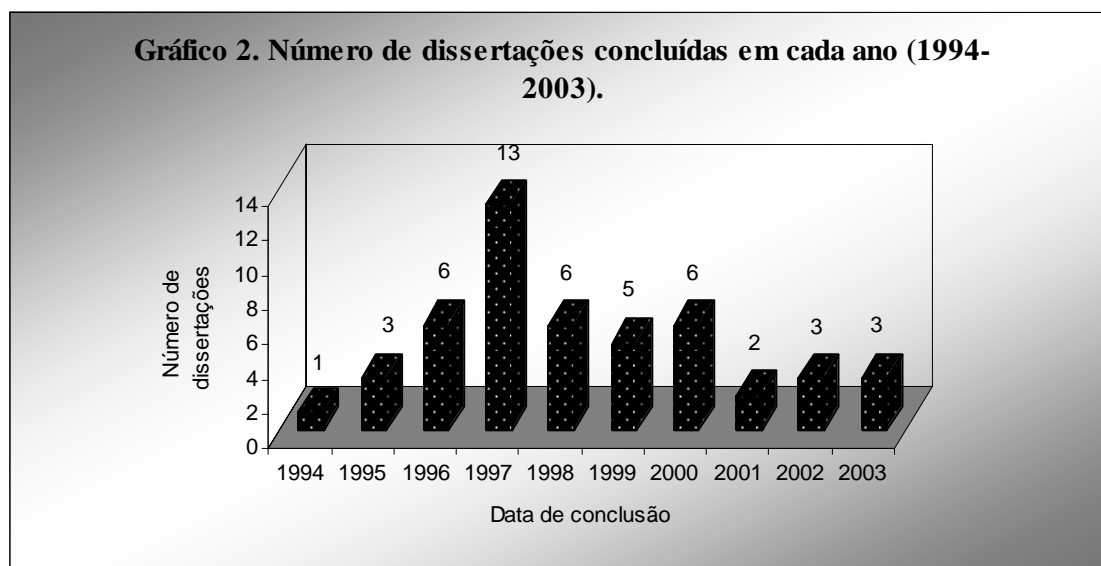
Fora da cidade de Lisboa, a UMI, a UAL, a UAV e a UEV viram dissertações aí concluídas, que se inserem no presente estudo, embora em número menos significativo. Assim, na UMI, em Braga, foram concluídas três dissertações, das quais foi possível saber que uma foi realizada no Instituto de Educação e Psicologia (IEP), tendo como data de conclusão 1999 e outra no Instituto de Estudos da Criança (IEC), concluída em 2001. Quanto à terceira dissertação realizada na UMI, em 2000, não foi possível determinar a escola/instituto onde o respectivo curso de mestrado teve lugar.

Em Faro, na UAL, foram identificadas duas dissertações e em Aveiro e Évora apenas uma em cada universidade, UAV e UEV, respectivamente. Curiosamente, não se conseguiu identificar e recolher nenhuma dissertação na área da avaliação das

aprendizagens que tivesse sido realizada na UP, dentro do período de tempo considerado, apesar desta Universidade ter na sua constituição uma Faculdade de Psicologia e Ciências da Educação.

Embora a larga maioria das dissertações analisadas tenha sido realizada em universidades de Lisboa, verificou-se a realização de dissertações no âmbito deste estudo em mais cinco cidades do país, situadas em regiões bem distintas, de norte a sul. Por outro lado, constata-se também que a composição das universidades é bastante diversa, verificando-se que nas universidades mais antigas existe uma FPCE, sendo aí que se concluíram a maior parte (no caso da UL) ou a totalidade (no caso da UCO) de dissertações integradas neste trabalho. Nas restantes universidades constata-se uma forma de organização e denominação das faculdades/institutos que as compõem bastante diversificada. De facto, e apenas a título de exemplo, na UTL existem quer faculdades quer institutos e a UMI é composta por escolas e institutos. Estes são aspectos que dificultaram, a par da falta de um sistema uniformizado de base de dados bibliográficos e da falta de uma biblioteca geral ou central de cada universidade (constatada na maioria dos casos), a pesquisa com vista à constituição do grupo de dissertações em análise neste estudo.

No Gráfico 2, podemos observar a forma como se distribuíram no tempo as dissertações em análise.



A análise do gráfico permite verificar que a maioria das dissertações analisadas nesta investigação – 36 – foi concluída entre 1996 e 2000, inclusive. Só em 1996, 1997 e 1998 concluíram-se 19 dissertações na área da avaliação das aprendizagens.



O elevado número de dissertações sobre avaliação das aprendizagens nos anos referidos parece estar relacionado, essencialmente, com os seguintes factores:

1. A abertura, pela primeira vez, dos cursos de mestrado em avaliação na UCP, nos primeiros anos da década de noventa;
2. O destaque dado à avaliação das aprendizagens após o Despacho Normativo 98-A/92, nomeadamente ao nível de acções de formação promovidas quer pelo Ministério da Educação, através do IIE, quer através dos Centros de Formação das Associações de Escolas, quer ainda através das Associações de Professores (Fernandes, 2005); e
3. A implementação de legislação com fortes consequências ao nível da avaliação das aprendizagens, especialmente a partir do Despacho Normativo 98-A/92.

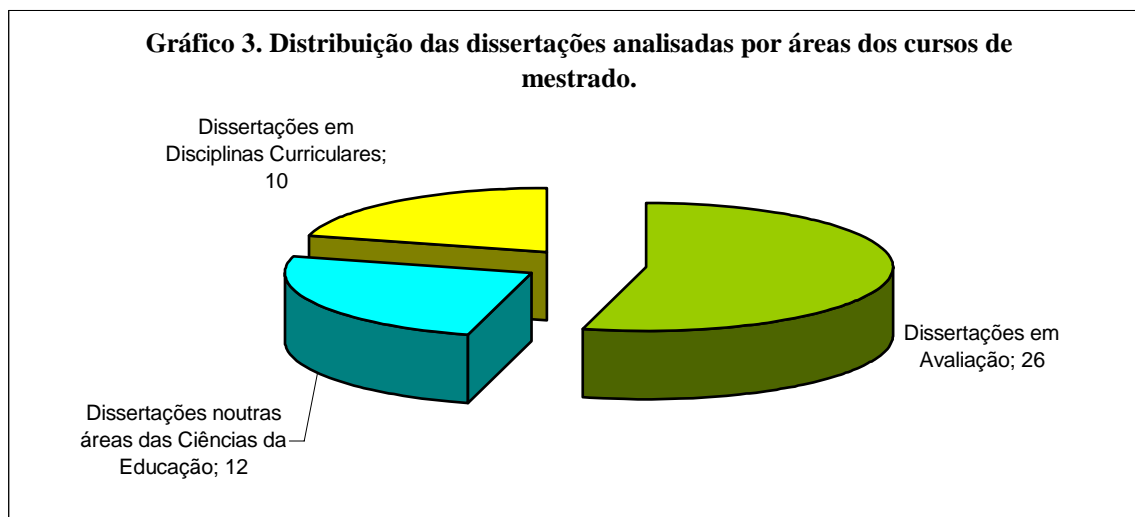
Por outro lado, verifica-se um decréscimo, que se pode considerar acentuado, no número de dissertações realizadas após o ano de 2000. De facto, em 2001, 2002 e 2003 foram identificadas apenas oito dissertações, quando nos três anos anteriores se haviam concluído 17, ou seja, mais do dobro.

As dissertações analisadas nesta investigação foram realizadas no âmbito de uma variedade de cursos de mestrado, a maioria dos quais na área das Ciências da Educação. Para efeitos deste estudo, e tendo em vista facilitar a análise e a leitura dos dados, agruparam-se as dissertações em três categorias, de acordo com os cursos de mestrado no âmbito dos quais se realizaram:

1. Dissertações em cursos de Mestrado com especialidade em Avaliação, como por exemplo: *Avaliação Educacional*; e *Avaliação*;
2. Dissertações em cursos de Mestrado com especialidade em Didácticas ou Metodologias de disciplinas específicas do currículo escolar, como por exemplo: *Ciências da Educação – Metodologia da Educação Física*; *Didáctica da Matemática*; *Metodologia do Ensino das Ciências – Biologia*; *Ensino de Geologia e Biologia*; *Performance Artística – Dança*; *Química – Ensino*;
3. Dissertações em cursos de Mestrado com especialidade noutras áreas das Ciências da Educação, como por exemplo: *Análise e Organização do Ensino*; *Formação de Professores*; *Observação e Análise da Relação Educativa*; *Educação*.

É de referir que, no decorrer do trabalho conducente a esta caracterização, se constatou que, frequentemente, a designação utilizada nas dissertações não correspondia à designação oficial do curso.

Para agrupar as dissertações no Gráfico 3 de acordo com as categorias acima indicadas designou-se a primeira por *Dissertações em Mestrados em Avaliação*, a segunda por *Dissertações em Mestrados em Disciplinas Curriculares*; e a terceira por *Dissertações em Mestrados noutras áreas das Ciências da Educação*.



Este gráfico apresenta a distribuição das dissertações analisadas por áreas dos cursos de mestrado em que foram, pelo que se refere ao total das dissertações em estudo, 48.

Da análise do gráfico, constata-se que mais de metade das dissertações foi concluída no âmbito de cursos de mestrado com especialização em avaliação e que apenas 10 se realizaram em mestrados relativos a disciplinas do currículo do ensino básico e/ou secundário. Este aspecto é importante, na medida em que se verifica que a avaliação não existe como disciplina, ou componente substancial de disciplinas, no plano curricular da quase totalidade, senão da totalidade, das licenciaturas da via de ensino.

No sentido de podermos fazer uma leitura dos dados referidos no espaço temporal em análise, foi construído o Quadro 7, que apresenta a distribuição das dissertações concluídas em cada uma das três categorias utilizadas no Gráfico 3, de acordo com a sua data de conclusão.

Da análise do quadro verifica-se que, enquanto no caso das *Dissertações nas Disciplinas das Ciências da Educação* e das *Dissertações em Disciplinas Curriculares*, as datas de conclusão se encontram igualmente distribuídas ao longo do período de tempo em estudo, no caso das *Dissertações em Avaliação* encontramos dois momentos de maior número de conclusões de dissertações – o primeiro nos anos de 1996 e 1997 e

o segundo no ano de 2000. Estes dados vão ao encontro da análise feita no Gráfico 2, e sugerem que os cursos de mestrado com especialização em avaliação, abertos a partir do início da década de 90, terão deixado de funcionar anualmente ou de assumir o mesmo grau de importância relativamente à avaliação das aprendizagens de forma específica, nos restantes períodos. Para melhor compreender estes dados há que ter em conta que, em geral, os mestrados aqui referidos tinham uma componente curricular de dois anos após a qual se realizava uma dissertação cuja conclusão raramente acontecia antes de três anos após o início do respectivo Curso.

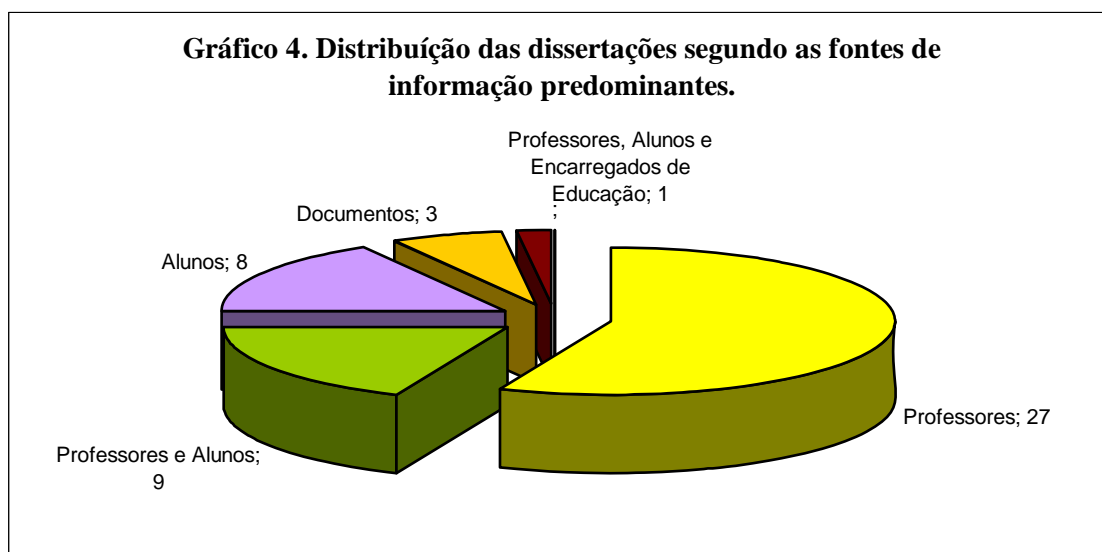
**Quadro 7. Número de dissertações concluídas em cada ano nas categorias *Dissertações em Mestrados em Avaliação, Dissertações em Mestrados em Disciplinas Curriculares e Dissertações em Mestrados noutras áreas das Ciências da Educação*.**

Ano \ Categoria	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	Total
<i>Dissertações em Mestrados em Avaliação</i>	0	1	4	9	3	2	4	0	2	1	26
<i>Dissertações em Mestrados noutras áreas das Ciências da Educação</i>	1	1	2	2	1	2	0	1	1	1	12
<i>Dissertações em Mestrados em Disciplinas Curriculares</i>	0	1	0	2	2	1	2	1	0	1	10
<b>Total</b>	1	3	6	13	6	5	6	2	3	3	48

### Caracterização das Fontes de Informação Utilizadas

No processo de análise das dissertações, pareceu importante referir e identificar as fontes de informação predominantemente utilizadas pelos investigadores. Para este efeito, só foram incluídas na categoria *Documentos* as dissertações em que a análise de documentos tenha sido a estratégia primordial ou mesmo exclusiva de recolha de dados de carácter basilar para o estudo em questão. Não foram aqui incluídas dissertações que recorrem à análise de documentos de forma a complementar a informação recolhida através de outras estratégias. Também é de referir que, nas situações em que se recorre à observação de aulas, sem que haja recurso a outras estratégias junto dos alunos ou referência explícita de que os alunos são alvo do estudo, estes não são contabilizados

em termos de análise numérica ou tipológica das fontes de informação (Anexo B). No Gráfico 4, apresenta-se a distribuição das dissertações em estudo pelas diferentes fontes de informação utilizadas.



Nota: Santo (2002) foi incluída na categoria *Professores* pelas razões já referidas no texto.

A análise do Gráfico 4 permite verificar que a grande maioria dos investigadores utilizou como fontes de informação primordiais os Professores (27) ou os Professores e Alunos simultaneamente (nove) e apenas oito dissertações se basearam nos alunos como fontes exclusivas de informação.

Esta tendência na escolha das fontes de informação mantém-se ao longo do período de tempo analisado, não se verificando qualquer alteração. Todavia, é de salientar que a última dissertação realizada com base em informação recolhida apenas junto de alunos foi concluída em 2000. Os dados sugerem que, na grande maioria das investigações analisadas, o professor é inequivocamente considerado como tendo o papel de actor principal da avaliação. De facto, é dada uma relevância à opinião dos professores mais de três vezes superior relativamente à opinião dos alunos. Com efeito, 27 dissertações têm como fonte exclusiva de informação os professores, ao passo que apenas oito recorrem exclusivamente aos alunos. No entanto, o facto dos autores das dissertações serem, eles próprios, professores, pode ser uma outra explicação para esta escolha, por se sentirem mais confortáveis num trabalho que envolve directamente os seus pares.

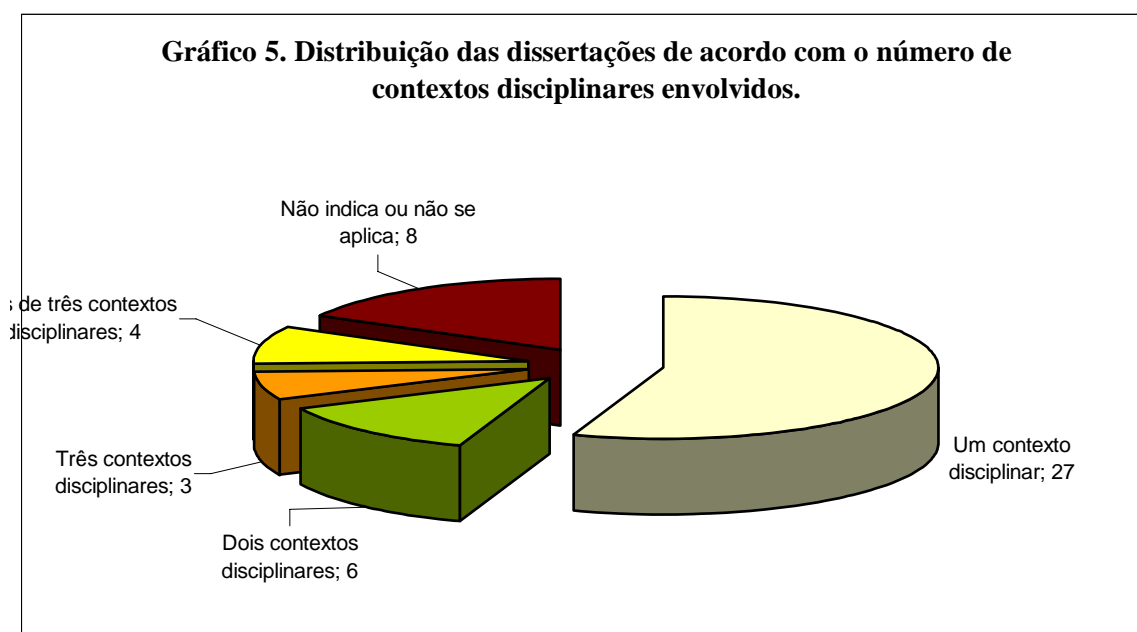
Por oposição à tendência descrita, de ter como fontes de informação professores e/ou alunos, evidenciam-se as dissertações de Lobo (1996) e a de Gouveia (2000), por apresentarem uma visão mais alargada dos actores no processo de ensino, aprendizagem

e avaliação, o que é, aliás, preconizado na legislação. De facto, em Lobo (1996) é incluída a opinião de um encarregado de educação, embora se tenha procurado que estes fossem em número idêntico ao de alunos e de professores – dois elementos – e, no caso de Gouveia (2000), embora o estudo recaia unicamente sobre alunos, estamos perante uma consciência clara da avaliação como forte condicionante e até conteúdo principal da comunicação entre a escola e as famílias, patente no próprio título – *Incidência da avaliação escolar nas relações família-escola em alunos do 4º ano de escolaridade*.

Apenas três dissertações utilizaram como fonte exclusiva de informação documentos de carácter legal (Ferreira, 1994), enunciados de tarefas de avaliação (Rodrigues, 2001), ou documentos contendo os resultados das avaliações dos alunos (Torres, 2000).

### Contexto Disciplinar das Investigações Analisadas

Outro aspecto interessante na caracterização geral das dissertações analisadas parece ser o que se relaciona com os contextos disciplinares em que se realizaram. O Gráfico 5 apresenta a distribuição das dissertações de acordo com o número de contextos disciplinares envolvidos, sendo contabilizadas para este efeito as 48 dissertações em estudo.



Nota: Para a elaboração deste gráfico inclui-se Gonçalves, 1997, na categoria *Um contexto disciplinar*, por se ocupar da disciplina de Educação Física, embora recolha informação junto de Professores de diversas disciplinas.

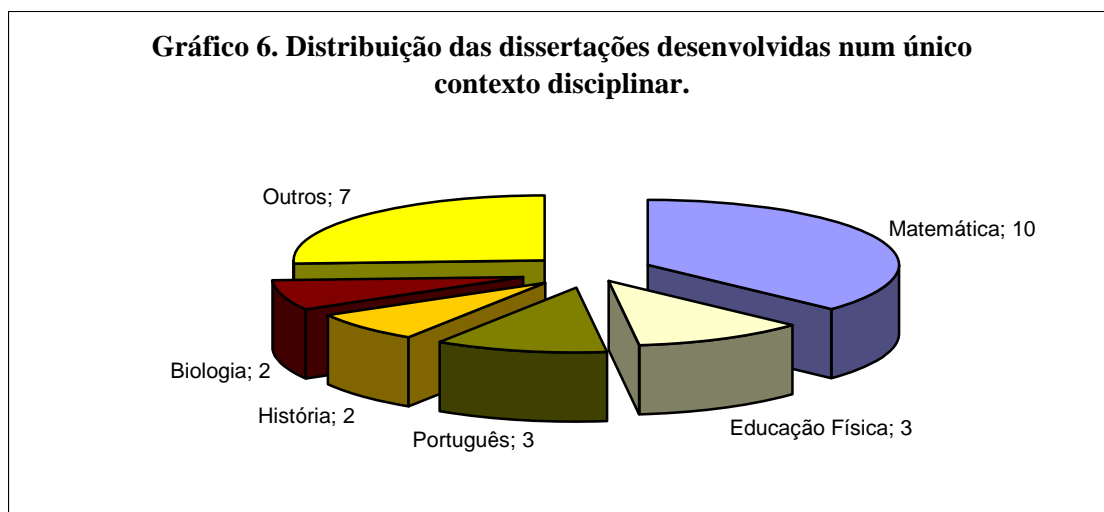
Da análise do Gráfico 5, conclui-se que a maioria das investigações foram levadas a cabo num único contexto disciplinar – 27 dissertações. Nesta categorização,

verifica-se que é reduzido o número de estudos que envolveu mais do que três contextos disciplinares, sendo de referir que, das quatro que se enquadram neste conjunto, uma ocupa-se de quatro contextos disciplinares, duas ocupam-se de seis contextos disciplinares, uma não refere o número exacto de contextos disciplinares que abrange e outra envolve um número superior a seis contextos disciplinares.

Analisando os contextos disciplinares em que ocorreram as 27 dissertações que se ocupam de um único contexto, obtemos o Gráfico 6.

Deste gráfico, destaca-se a *Matemática* como contexto disciplinar que predomina (10 dissertações). Neste grupo de estudos, existe uma distribuição equilibrada das datas de conclusão das dissertações ao longo de todo o período temporal considerado, a primeira tendo sido concluída em 1995 e a última em 2001. Quanto à instituição onde foram realizados, verificou-se que a maioria, quatro, foi concluída na UCP e outras quatro na UL, tendo as duas restantes sido realizadas nas universidades do Algarve e do Minho (UAL e UMI).

As dissertações no contexto disciplinar do *Português*, foram realizadas na UCP e concluídas em 1997. Duas das três dissertações no contexto disciplinar da *Educação Física* foram realizadas na FMH da UTL, sendo concluídas em 1997 e a terceira dissertação nesta área disciplinar foi concluída em 1999, na UCP.

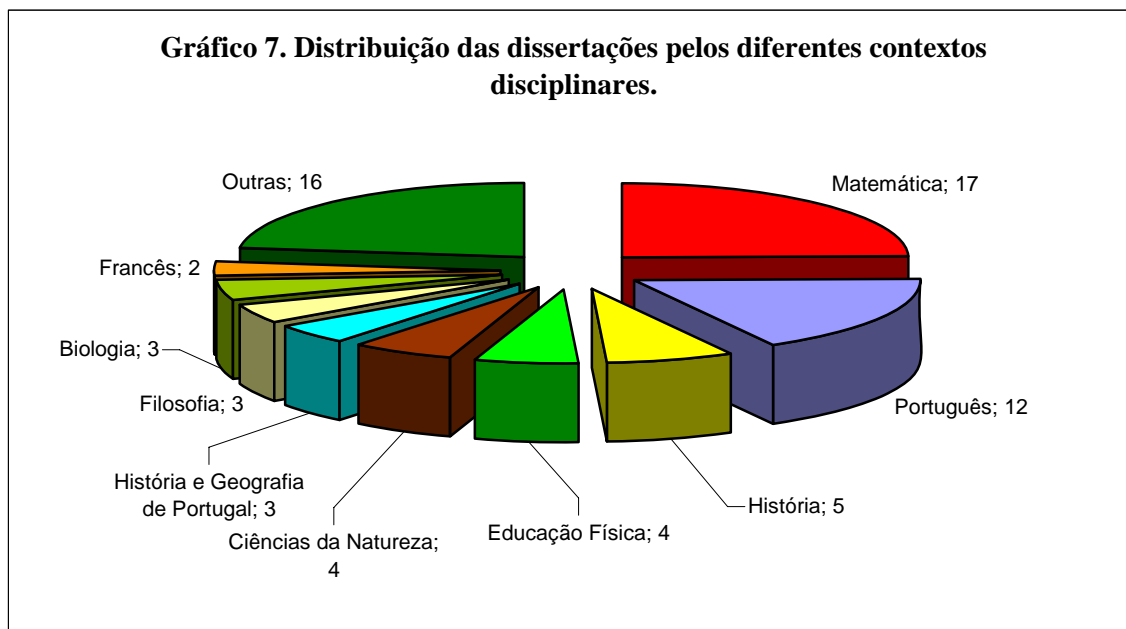


Nota: Para a elaboração deste gráfico inclui-se Gonçalves, 1997, na categoria *Um contexto disciplinar*, por se ocupar da disciplina de Educação Física, embora recolha informação junto de Professores de diversas disciplinas.

Na categoria *Outros* englobam-se contextos disciplinares vários, a saber: uma dissertação no contexto disciplinar de Alemão, de 2001, realizada na UL; uma no de Filosofia, de 1997, realizada na UCP; uma no de Dança, de 1998, realizada na FMH da

UTL; uma no de Ciências da Natureza, de 1998, realizada na UCP; uma no de Oficina de Artes, de 1999, realizada na UCP; uma no de Informática, de 2003, realizada na UAL; e uma no de Francês, de 1997, realizada na UCP.

O Gráfico 7 mostra a distribuição das dissertações de acordo com o contexto disciplinar em que foram realizadas, independentemente dos estudos se debruçarem sobre um ou mais destes contextos. Deste modo, a soma dos valores do gráfico é, necessariamente, superior ao total de dissertações analisadas, uma vez que, como vimos, 13 dissertações se basearam num trabalho empírico no âmbito de mais do que um contexto disciplinar. A análise do gráfico torna evidente que o contexto disciplinar mais utilizado é o da *Matemática*, seguido do *Português*. Verifica-se também que os contextos disciplinares de *História* e de *História e Geografia de Portugal* perfazem, no conjunto, um total de oito dissertações. Na categoria *Outros* foi incluída uma grande diversidade de contextos disciplinares referidos, desde a área das línguas estrangeiras, onde se incluíram o Inglês, o Francês e o Alemão, a área das ciências, onde se incluem as Técnicas Laboratoriais de Biologia, as Ciências Físico-Químicas, a Física e a Química, e ainda áreas artísticas como a Educação Visual, a Educação Visual e Tecnológica, a Dança, a Música, e a Oficina de Artes.



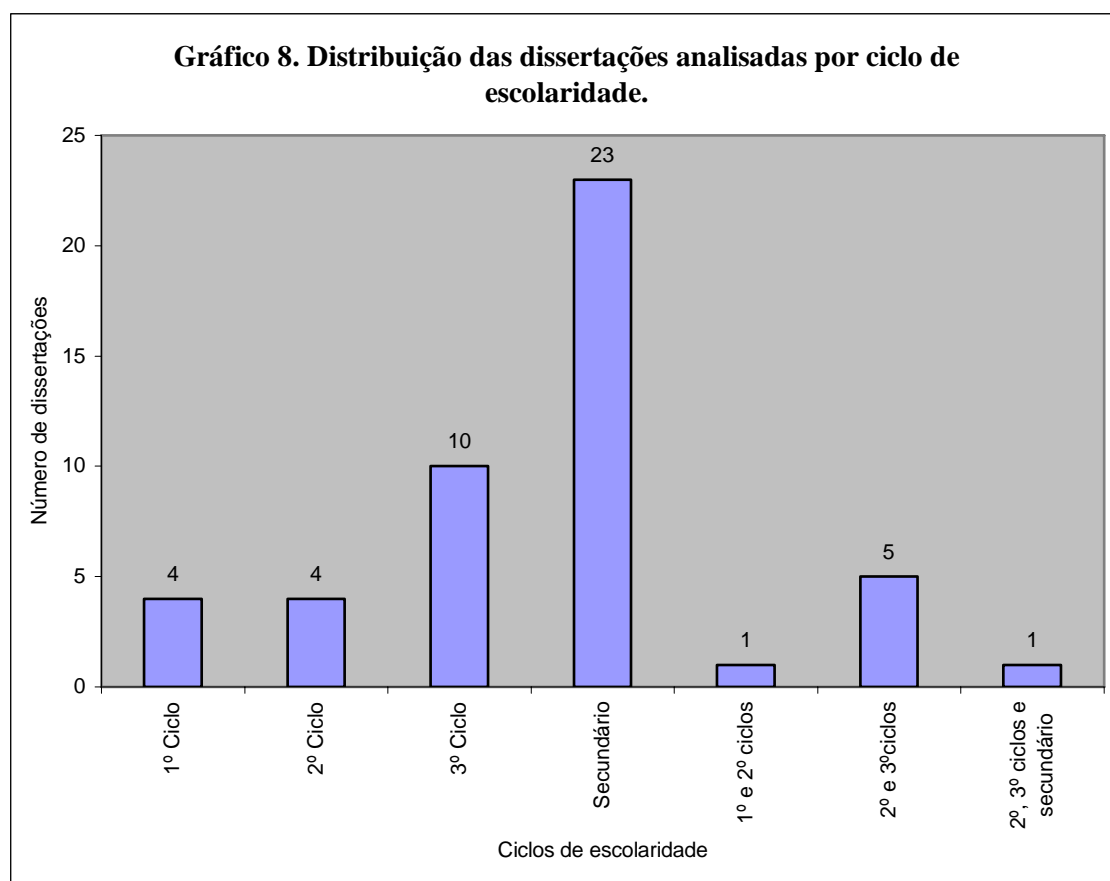
Nota: Para a elaboração deste gráfico inclui-se Gonçalves, 1997, no grupo de um *contexto disciplinar*, por se ocupar da disciplina de Educação Física, embora recolha informação junto de Professores de diversas disciplinas.

## **Níveis e Anos de Escolaridade**

Outro aspecto que mereceu atenção, no sentido de se proceder a uma caracterização geral das dissertações analisadas, foi o dos níveis de ensino e dos anos de escolaridade envolvidos nas investigações.

No Gráfico 8 apresenta-se a distribuição das dissertações analisadas por ciclo de escolaridade no contexto do qual decorreu a investigação. A análise do gráfico mostra que praticamente metade das dissertações analisadas se realizou exclusivamente no âmbito do ensino secundário (23 das 48 dissertações em estudo), seguindo-se-lhe o 3º ciclo (dez das 48 dissertações analisadas). Este aspecto pode prender-se com diversos factores, desde o ciclo de leccionação dos investigadores, nos casos em que estes eram Professores, até ao facto de o nível secundário ser aquele que marca a transição para o ensino superior e aquele onde se realizam avaliações externas, com alguma tradição. No entanto, esta última hipótese explicativa não se traduz de forma tão clara na distribuição das dissertações por problemáticas abordadas, como veremos a seguir. De facto, são poucas as dissertações que se debruçaram especificamente sobre a avaliação externa. É ainda de referir a existência de sete dissertações que abrangem mais do que um ciclo de ensino. No caso das dissertações que se realizaram no âmbito dos 2º e 3º ciclos simultaneamente, constata-se que há professores que leccionam estes dois ciclos de ensino, factor que pode ter condicionado os agrupamentos dos ciclos sobre os quais recaíram as dissertações, como são os casos de Campos (1996), Neves (1996), Alves (1997) e Araújo (2002). No caso de Martins, I. (1997), uma vez que este estudo foi feito no âmbito disciplinar de História e Geografia de Portugal e História, o facto de se trabalhar com o 2º e 3º ciclos em simultâneo pode prender-se também com alguma continuidade entre as duas disciplinas de um ciclo para o seguinte. No caso da dissertação que se debruça sobre o 2º e 3º ciclos e secundário, a disciplina que é alvo de estudo é leccionada desde o 2º ciclo até ao secundário, inclusive, e o objectivo do estudo prendia-se com a auscultação da opinião dos professores dos vários ciclos. Verifica-se ainda que o número de dissertações que se ocupam do 1º ciclo é reduzido, representando cerca de 11% do número total de dissertações analisadas.

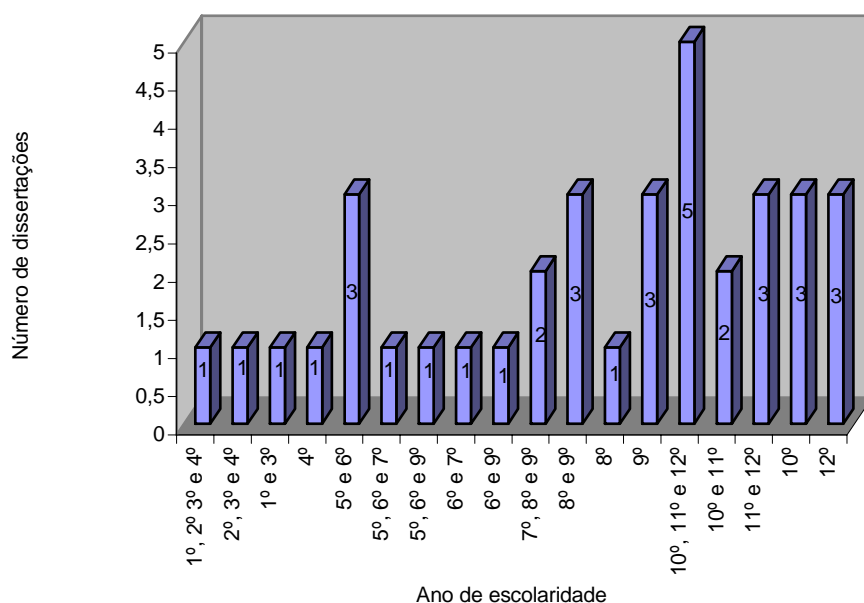




O Gráfico 9 permite analisar a distribuição das investigações realizadas por ano de escolaridade, sendo de salientar que, de acordo com a explicação dada na legenda, o número total de dissertações onde foi possível encontrar esta informação foi de 36.

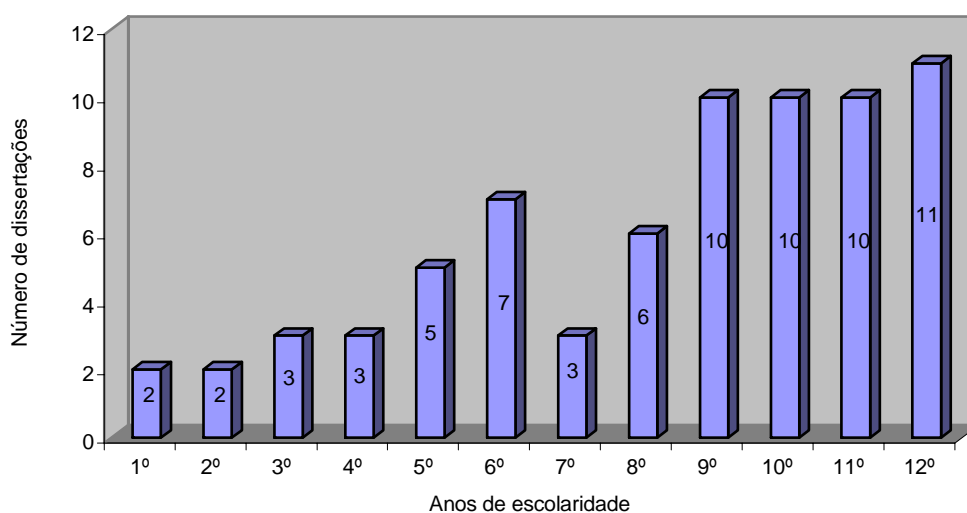
A partir da análise do gráfico, pode observar-se um pico de dissertações sobre os três anos que compõem o nível secundário (cinco dissertações). Igualmente se verifica que a grande maioria das dissertações se desenvolveu ao nível dos anos terminais de ciclo. Por outro lado, apenas quatro dissertações contemplam mais do que um ciclo de escolaridade. Este aspecto torna-se mais visível no Gráfico 10, onde se faz uma análise agregada do número de dissertações que se debruça sobre cada ano de escolaridade, pelo que o número de dissertações ultrapassa as apresentadas no gráfico anterior. De facto, a maioria das dissertações estudadas teve como contexto mais do que um ano de escolaridade.

**Gráfico 9. Distribuição desagregada das dissertações analisadas por ano de escolaridade.**



Nota: Não foram incluídas, por não referirem os anos de escolaridade, Afonso (1999), Alves (1997), Baptista (1997), Dias (1999), Ferreira (1994), Ferreira (2003), Fonseca (1998), Leite (1998), Martins, F. (1997), Ribeiro (1997), Santo (2002) e Santos (2003).

**Gráfico 10. Distribuição agregada das dissertações analisadas por ano de escolaridade.**



Nota: Não foram incluídas, por não referirem os anos de escolaridade, Afonso (1999), Alves (1997), Baptista (1997), Dias (1999), Ferreira (1994), Ferreira (2003), Leite (1998), Martins, F. (1997), Ribeiro (1997), Santo (2002) e Santos (2003).

Como se pode observar no Gráfico 10, os anos terminais de ciclo são os mais estudados, ou seja, o 4º, 6º, 9º e 12º anos, e são-no de forma crescente do primeiro ciclo até ao ensino secundário. Dentro de cada ciclo, nem sempre existe uma progressão

crescente de ano para ano. Com efeito, só no 2º e 3º ciclos é que existe esta progressão crescente, sendo que, no 1º ciclo, os 1º e 2º anos constituíram, cada um, contexto de estudo para duas dissertações, acontecendo o mesmo nos 3º e 4º anos (três dissertações). Dentro do ensino secundário, os 10º e 11º anos também serviram de contexto para o mesmo número de investigações (dez).

Constata-se que o 1º ciclo é o contexto menos estudado (dez dissertações) e que o secundário é o mais estudado (31 dissertações). Por outro lado, de entre os quatro ciclos de escolaridade, aquele que apresenta um aumento mais acentuado de dissertações de ano para ano é o terceiro. Assim, o 7º ano de escolaridade serve de contexto apenas a três dissertações, enquanto que o 9º serve de contexto de estudo a dez, ou seja, mais do triplo. Se, por um lado, é compreensível o elevado número de dissertações que se ocupam do 9º ano, por ser ano terminal de ciclo e, além disso, conclusivo do ensino obrigatório, já será passível de questionamento o número reduzido de investigações sobre o 7º ano, que marca o início de um novo ciclo, apresentando um substancial acréscimo de disciplinas, a par das transformações ligadas ao desenvolvimento dos jovens na idade prevista para a sua entrada neste ano.

Como já foi referido, observa-se uma maior incidência de estudos em anos terminais de ciclo, com maior incidência, porém, nos 12º e 9º anos. De facto, é no final do 12º ano que os alunos são sujeitos às avaliações externas cujos resultados têm fortes consequências a nível dos alunos, mas também das escolas, havendo uma divulgação pública destes resultados, de uma forma que tem vindo a captar cada vez mais a atenção da sociedade em geral e impele a tomadas de decisão por parte dos órgãos administrativos e políticos. No 9º ano, as avaliações são internas excepto nos casos da Matemática e Língua Portuguesa que têm uma componente externa através dos exames. Este é o ano de escolaridade que corresponde à conclusão do ensino obrigatório, pelo que a avaliação tem um carácter globalizante em termos das aprendizagens de todo o terceiro ciclo. O destaque do número de dissertações sobre o 6º ano, relativamente aos anteriores, poderá ser explicado por se tratar igualmente de um ano terminal de ciclo, sobre o qual recaem provas de aferição, embora de forma não sistemática.

Uma razão plausível para a escolha do nível e ano lectivo analisados nas dissertações poderá ser a ligação dos seus autores a esse nível de ensino, no caso da investigação ser conduzida por professores. A confirmação ou não desta hipótese não cabe no âmbito deste trabalho, embora seja considerada de interesse para um trabalho posterior.

## **Zonas do País em que se Realizaram as Investigações**

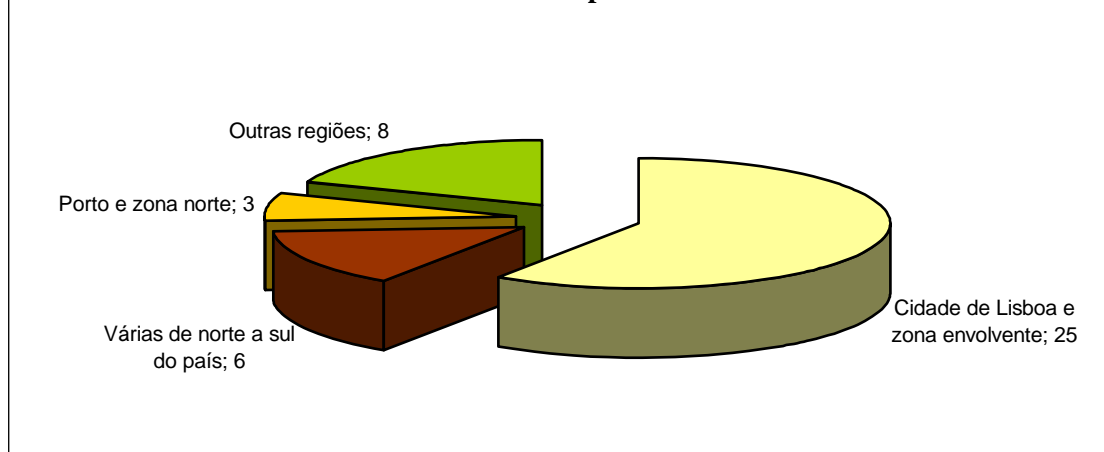
Analizadas as 48 dissertações seleccionadas para integrarem este estudo, foi possível identificar 46 que se basearam na recolha de dados junto de elementos da comunidade educativa de escolas dos 2º, 3º ciclos e secundárias. As duas restantes basearam-se na análise de documentos de carácter legal (Ferreira, 1994) e enunciados de provas de avaliação externa (Rodrigues, 2001), portanto não associados a nenhuma escola em particular.

Com o objectivo de proceder à análise da distribuição geográfica das escolas onde foram recolhidos os dados dos 46 estudos em análise, constituíram-se as quatro categorias seguintes de distribuição geográfica:

1. *Cidade de Lisboa e zona envolvente*, onde se incluiu todo o distrito de Lisboa, o concelho de Almada e as regiões designadas de “arredores de Lisboa”, “próxima de Lisboa”, “periferia da Grande Lisboa” e “região de Lisboa”; “grande Lisboa e Margem Sul do Tejo”;
2. *Porto e zona norte*, onde se incluiu o distrito de Porto e regiões referidas nas dissertações como “zona norte do país”, “grande Porto”, “distrito de Braga”, e “região norte do país”;
3. *Várias de norte a sul do país*, que inclui duas ou mais localizações integradas nas três grandes zonas genericamente designadas de norte, centro e sul;
4. *Outras regiões*, que inclui locais que claramente não se integram nas zonas incluídas em 1) e 2), mas também outras que podem eventualmente ser consideradas de fronteira, e que, por essa razão, são aqui referidas: “Setúbal”, “zona litoral suburbana”, “distrito de Aveiro”, zona norte do distrito de Santarém”, “concelho de Beja”, “centro do país” e “Algarve”.

O Gráfico 11 apresenta a distribuição relativa às zonas do país em que se localizavam as escolas dos ensinos básico e secundário onde os dados das referidas dissertações foram recolhidos – apenas foi possível recolher esta informação em 42 dissertações.

**Gráfico 11. Distribuição das escolas onde se recolheram os dados por zona do país.**



Nota: Não se incluíram por não referir estes dados ou por não se adequarem a esta análise – Ferreira (1994), Graça (1995), Boavida (1996), Pereira (1997), Amado (1998), e Rodrigues (2001).

Este gráfico evidencia claramente que os dados das investigações analisadas foram, na sua maioria, recolhidos em escolas situadas na *Cidade de Lisboa e zona envolvente*. De facto, existem 25 dissertações cujos dados foram aí recolhidos, o que representa mais de metade das dissertações que envolveram dados recolhidos em escolas. Deste grupo, 15 autores recolheram informação em escolas do distrito de Lisboa, cingindo-se, na maioria, a uma ou duas escolas, destacando-se Dias (1999), por envolver no estudo trinta escolas, e Gonçalves (1997), que envolveu todas as escolas secundárias dos concelhos de Amadora e Sintra, ou seja, 15 escolas. Dos restantes 10 estudos realizados em escolas situadas na *Cidade de Lisboa e zona envolvente*, duas recolheram dados em escolas apenas do concelho de Almada, Peres (1995) e Martins, F. (1997), envolvendo quatro e oito escolas, respectivamente, e as restantes não especificaram de forma clara a sua localização na região considerada, envolvendo entre uma a duas escolas na sua pesquisa.

Das três dissertações que recolheram dados em escolas do *Porto e zona norte*, apenas duas se restringem ao “distrito do Porto” e “grande Porto”, existindo grande disparidade no número de escolas que envolveram – apenas duas escolas, na dissertação de Pais (1998) e 29 na de Fonseca (1998), tendo, esta última, o intuito de recolher dados em todas as escolas secundárias com ensino da dança, localizadas no distrito do Porto. A terceira dissertação deste grupo realizou-se em Macieira, no distrito de Braga, envolvendo apenas uma escola.

No grupo de seis dissertações que foram incluídas na categoria *Várias de norte a sul do país*, Rocha (2003) foi a que envolveu um menor número de escolas – apenas

duas – uma localizada em Lisboa e outra no Algarve. Gil (1997) e Ribeiro (1997) não referem o número exacto de estabelecimentos de ensino que envolveram no estudo empírico mas, pelo número de diferentes localizações geográficas referidas, depreende-se que não foi inferior a cinco em qualquer dos dois casos. Em Gil (1997) são referidas as localidades: Lisboa, Alverca, Caldas da Rainha, Seixal e Castelo Branco. Ribeiro (1997) refere o centro e os arredores de Lisboa, a zona urbana do Minho, e duas zonas urbanas e uma zona rural no Alentejo. Finalmente, Gouveia (2000) envolveu 19 escolas dos concelhos de Cascais, Oeiras, Bragança, Alfândega da Fé, Mogadouro e Macedo de Cavaleiros. Afonso (2000) envolveu 36 escolas de “todo o país”, sem especificar as suas localizações, e Santo (2002) recolheu informação junto de 49 escolas e cinco centros de formação, sendo as escolas pertencentes aos distritos de Vila Real, Viseu e Coimbra.

Este é o grupo que, por abarcar diferentes regiões do país, poderia levar a crer que englobaria os estudos cujo trabalho empírico se teria realizado num maior número de escolas. No entanto, Rocha (2003) envolveu apenas quatro escolas, e tanto Gil (1997) como Ribeiro (1997) podem ter realizado recolha de dados apenas em cinco escolas, pois, de facto, não indicam o seu número exacto, indicando cinco localidades geograficamente diferentes.

Finalmente, na categoria *Outras regiões*, apenas duas dissertações envolveram recolha de dados em mais de três escolas: Ferreira (2003), que envolveu 31 escolas, localizadas no distrito de Aveiro, e Santos (2003), que envolveu todas as escolas secundárias do Algarve, num total de 17. Das restantes seis dissertações deste conjunto, quatro incluíram apenas uma escola – Henriques (1995) numa “zona litoral suburbana”, Martins (1999) numa “cidade ribatejana”, Mourão (1997) no “distrito de Santarém” e Torres (2000) na cidade de Setúbal – uma realizou-se junto de duas escolas do “concelho de Beja” e outra em três escolas da Cidade de Setúbal.

### **Síntese da Resposta à Primeira Questão**

Mais de metade das dissertações foram realizadas na UCP, seguindo-se-lhe a UL e a UTL. Curiosamente, na UP não foi encontrada qualquer dissertação que se inclui nesta investigação embora tenha uma FPCE.

A esmagadora maioria das dissertações seleccionadas para este estudo foi realizada em universidades localizadas em Lisboa.

Os anos de maior produção de dissertações aqui incluídas (36 em 48) foram de 1996 a 2000, inclusive, com um pico no ano de 1997 (em que foram concluídas 13).

Este aspecto pode estar relacionado com factores tais como: a implementação de cursos de mestrado em avaliação na UCP, nos primeiros anos da década de noventa; o destaque dado à avaliação das aprendizagens após o Despacho Normativo 98-A/92, nomeadamente ao nível de acções de formação promovidas quer pelo Ministério da Educação, através do IIE, quer através dos Centros de Formação das Associações de Escolas, quer ainda através das Associações de Professores (Fernandes, 2005) e ainda a criação de legislação sobre a avaliação das aprendizagens, especialmente a partir do Despacho Normativo 98-A/92.

A categoria em que se insere o maior número de estudos aqui analisados (26 dissertações) é a que foi designada por *dissertações em mestrados em avaliação*, sendo a categoria *dissertações em mestrados em disciplinas curriculares* aquela com menor número de estudos (dez) centrados na problemática da avaliação das aprendizagens.

A maior parte das dissertações (27), teve os professores como fonte exclusiva de informação, o que poderá indiciar, da parte dos autores dos trabalhos analisados, uma concepção de avaliação centrada nesta categoria de agentes educativos. Apenas uma dissertação incluiu informação recolhida junto de um encarregado de educação, além de professores e alunos.

A maioria das dissertações realizou-se no âmbito de um único contexto disciplinar (27), destacando-se a matemática, seguida da língua portuguesa. É de referir que também foi objecto de estudo a avaliação de aspectos particulares de algumas disciplinas: uma dissertação debruçou-se sobre a componente laboratorial de técnicas laboratoriais de biologia (TLB) e de ciências físico-químicas (CFQ); uma outra sobre a avaliação da resolução de problemas no âmbito da disciplina de matemática e uma terceira dissertação estudou o processo de avaliação da dança, uma unidade que faz parte do currículo da disciplina de educação física. Há que salientar a existência de dissertações no âmbito de áreas artísticas como a dança, a oficina de artes, a música, a educação visual e a educação visual e tecnológica. Curiosamente, também se encontrou uma dissertação no âmbito do ensino do alemão que existe como disciplina opcional num número relativamente restrito de escolas.

O ciclo de ensino onde se realizou maior número de dissertações (16) foi o secundário, seguido do 3º ciclo. Os anos terminais de ciclo são aqueles onde se verificou um maior número de estudos, número que aumenta à medida que se avança nos níveis de ensino. Nota-se uma escassez de estudos ao nível do 1º ciclo (quatro) que, afinal, constitui uma base primordial no que respeita à criação de hábitos e estratégias

de estudo bem como de desempenho dos alunos, à introdução ao cálculo e às noções básicas do pensamento geométrico e espacial, à estrutura e forma de funcionamento da língua materna, além das noções de cidadania e de relação e integração do aluno na comunidade que o rodeia e na sociedade em geral.

Em termos de distribuição geográfica das dissertações que realizaram trabalho empírico junto de escolas (46 investigações), 25 fizeram-no na *cidade de Lisboa e zona envolvente*, contrastando com o *Porto e zona norte* onde se realizaram apenas três dissertações. Seis dissertações debruçaram-se sobre escolas situadas em mais do que uma região do país, envolvendo entre quatro e 49 escolas, distribuídas por diferentes concelhos de diferentes distritos do país.

Das oito dissertações que elaboraram estudos em apenas uma localidade não incluída no *Porto e zona norte* nem *cidade de Lisboa e arredores*, distribuíram-se de Aveiro ao Algarve e envolveram entre uma e 31 escolas.

### **Caracterização dos Problemas e Questões de Investigação**

No sentido de dar resposta à segunda questão deste estudo – *Como se poderão descrever a natureza dos problemas e das questões de investigação das dissertações analisadas?* – começou-se por proceder ao levantamento do propósito de cada dissertação. Como foi referido no capítulo anterior, optou-se por considerar que o enunciado do problema deveria preferencialmente obedecer, cumulativamente, aos seguintes requisitos: a) ser claro em termos de linguagem/formulação; b) ser constituído por um número reduzido de frases; c) identificar o(s) objecto(s) de estudo; e d) ser identificável.

No entanto, uma vez que não se pretende fazer qualquer tipo de avaliação das dissertações em estudo, estes requisitos foram considerados da forma mais abrangente possível, de modo a identificar as problemáticas em análise.

Verificou-se que a maioria dos autores – 47 – enunciou um problema de investigação. Esta apresentação fez-se maioritariamente na introdução do documento (em 44 dissertações), geralmente logo no primeiro capítulo. Apenas três autores (Fonseca, 1998; Afonso, 1999; e Tavares, 1999) enunciaram o problema em capítulos intermédios ou finais da dissertação. Numa dissertação (Braga, 1999) não foi enunciado o problema, de modo que nos fosse possível identificá-lo. No entanto, foram elaborados objectivos e questões de investigação.



No que respeita à forma de especificar o problema de investigação, 33 autores fizeram-no através de “questões”, cinco através de “objectivos”, uma através de “objectivos e questões”, uma através de “questões problema” e quatro não utilizaram qualquer destas designações para as especificações feitas. Apenas em três dissertações não foi possível identificar qualquer forma de especificação do problema de investigação. Quanto ao seu número, constatou-se que mais de metade (30) dos autores não elaborou mais do que cinco especificações/itens, embora seis elaborassem de nove a 12. Uma análise de conteúdo dos problemas e/ou suas especificações, de todas as dissertações, fez emergir informação, ao nível do trabalho empírico, que se sintetiza no Quadro 8.

**Quadro 8. Informação acerca do estudo empírico contida nos problema e/ou suas especificações.**

Informação acerca do estudo empírico a realizar no que respeita a:	Número de dissertações
Tipo de indivíduos alvo de estudo (professores e/ou alunos)	46
Âmbito disciplinar	27
Nível de ensino	27
Ano (s) de escolaridade	5
Período temporal abrangido	3

Para uma leitura adequada do Quadro 8, define-se “indivíduos alvo de estudo” como tratando-se dos indivíduos que estão a ser estudados. Pretende-se, deste modo, fazer uma distinção clara entre estes elementos e aqueles que constituem fontes de informação. Por exemplo, Gonçalves (1997) propôs-se estudar as práticas de avaliação dos professores (indivíduos alvo de estudo), a partir de informação recolhida junto de professores e de alunos (fontes de informação). Ou seja, nesta dissertação, os professores foram, simultaneamente, alvo de estudo e fontes de informação, enquanto que os alunos foram apenas mais uma fonte de informação.

Do conjunto das 48 dissertações analisadas, verifica-se que apenas três não identificaram, no problema e/ou sua especificações, nenhum tipo de sujeito em estudo. Destas, duas têm um carácter histórico, baseando-se na análise de documentos (Ferreira, 1994 e Rodrigues, 2001) e uma baseia-se num trabalho empírico sobre professores (Alaíz, 1996).

Das 21 dissertações que não deram informação acerca do contexto disciplinar em que o estudo se iria realizar, é de referir que três se centraram no 1º ciclo sem se debruçarem sobre nenhuma área disciplinar em particular. Das restantes 18, o estudo empírico pretendeu abordar temas transversais a diferentes disciplinas, como Afonso

(1999), temas envolvendo diferentes contextos disciplinares, como Campos (1996) e Araújo (2002), ou ainda temas dentro de um único contexto disciplinar, como Queirós (1997) e Santos (2003).

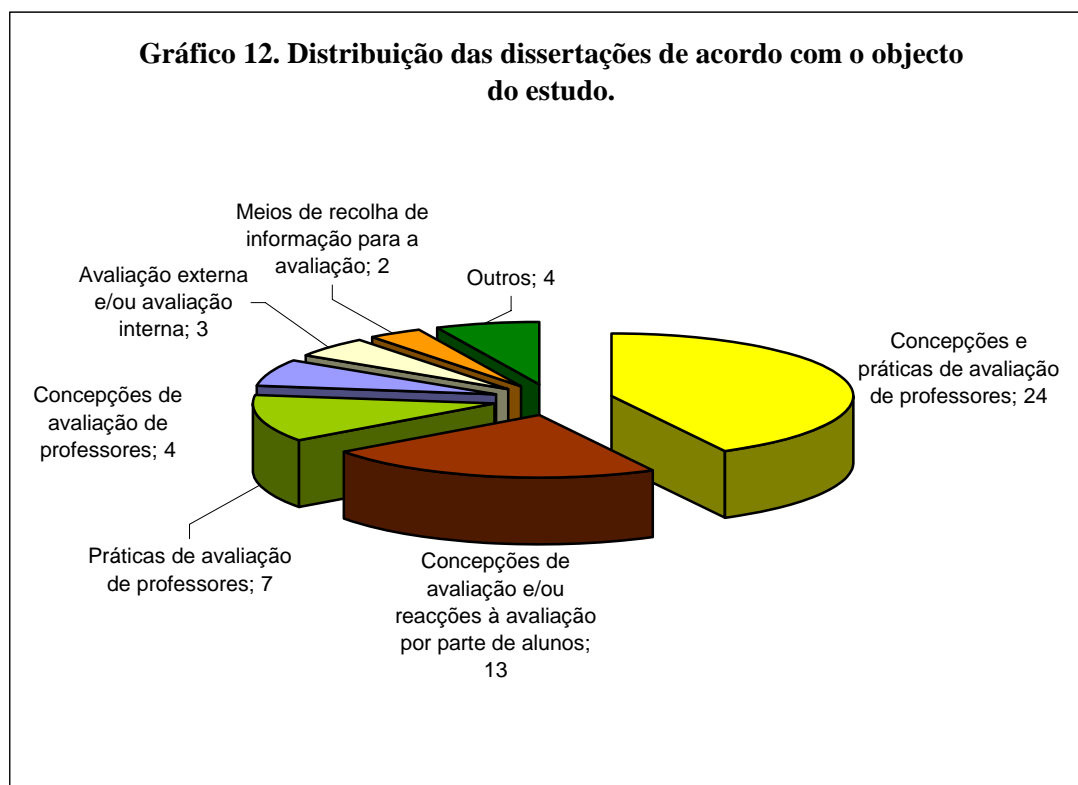
É ainda de salientar que três autores indicaram o período temporal que o estudo empírico visou abranger. Deste conjunto, uma dissertação faz uma abordagem histórica da avaliação baseando-se essencialmente em diplomas legais (Ferreira, 1994), outra ocupa-se da análise de provas de avaliação externa em matemática, também numa perspectiva histórica, cujo horizonte temporal se situa ao longo do último quartel do século XX (Rodrigues, 2001). A terceira dissertação deste grupo tem por propósito fazer uma análise comparativa entre os resultados das avaliações contínua, das provas globais e dos exames nacionais, ao longo de três anos lectivos consecutivos, numa abordagem *à posteriori* (Torres, 2000).

Conclui-se que, na apresentação e especificação do problema em estudo, a maioria das dissertações dá indicação de qual a fonte prioritária de informação, bem como do âmbito disciplinar em que decorrerá o estudo. Mais de metade destas dissertações referiram o nível de ensino que constituiu o objecto de estudo e cinco delas especificaram mesmo o(s) ano(s) de escolaridade. A informação acerca do período temporal sobre o qual decorreria a investigação verificou-se apenas em três dissertações que se propuseram efectuar uma abordagem histórica da avaliação ao nível das determinações legais, ou ao nível das provas de avaliação externa, existindo uma dissertação que se ocupou da comparação dos resultados das avaliações interna e externa ao longo de três anos lectivos.

A partir da análise do problema, das questões e/ou dos objectivos constantes em cada dissertação analisada, foi possível definir as seguintes categorias relativas ao objecto de estudo das dissertações analisadas: *Concepções e práticas de avaliação de professores*, *Concepções de avaliação e/ou reacções à avaliação por parte de alunos*, *Práticas de avaliação de professores*, *Concepções de avaliação de professores*, *Avaliação externa e/ou avaliação interna*, *Meios de recolha de informação para a avaliação* e *Outros*.

O Gráfico 12 apresenta a distribuição das dissertações analisadas pelas diferentes categorias de objecto de estudo. É de referir que, uma vez que existem dissertações que foram incluídas em mais do que uma categoria, a soma dos valores do gráfico é superior às 48 analisadas. A análise do Gráfico 12 permite verificar que

praticamente metade das dissertações – 24 – definiu um objecto de estudo relacionado com as *Concepções e práticas de avaliação de professores*. Neste grupo, duas dissertações foram incluídas em mais do que uma categoria (Lobo, 1996 e Rafael, 1998); em ambos os casos, existe uma preocupação em analisar as *Concepções e práticas de professores*, bem como *Concepções de avaliação e/ou reacções à avaliação por parte de alunos*.



Fazendo uma análise mais fina deste conjunto de dissertações, constata-se o seguinte: quatro dissertações recorrem à opinião de professores e de alunos para compreender as concepções e as práticas de avaliação dos professores (Gonçalves, 1997; Leite, 1998; Lobo, 1996; e Rafael, 1998); e que uma dissertação se ocupa exclusivamente das concepções e das práticas de classificação dos professores (Pais, 1998). A maioria (15 em 24 estudos) centrou-se num número não superior a quatro professores. Mais concretamente, oito dissertações tiveram como amostra três professores, cinco dissertações tiveram como amostra dois professores e duas dissertações contaram com quatro professores na sua amostra. No entanto, quatro estudos ultrapassam os 80 indivíduos como fonte de informação para o seu estudo, indiciando uma metodologia do tipo de sondagem, a saber: Santos (2003), que auscultou 85 professores; Dias, (1998), que trabalhou junto de 94 professores; Rocha

(2003), cuja amostra incluiu 104 professores; e Gonçalves (1997), que recolheu informação para o seu estudo junto de 205 professores e 776 alunos. No entanto, destes quatro estudos, apenas um abrangeu mais do que uma zona geográfica do país (Rocha, 2003) tratando-se, curiosamente, do mesmo estudo que abrangeu o menor número de escolas. De facto, o estudo de Rocha (2003) realizou-se com base em apenas quatro escolas de duas regiões bem distintas do país, Lisboa e Algarve. As três dissertações restantes, do conjunto de estudos que envolveram mais de 50 indivíduos na amostra, centraram-se numa única região, sendo a de Santos (2003) a que é geograficamente mais abrangente, na medida em que incluiu todas as escolas secundárias do Algarve (17). Gonçalves (1997) também fez um estudo exaustivo, ao nível de todas as escolas secundárias dos concelhos de Amadora e Sintra, num total de 15, e Dias (1999) trabalhou junto de 30 escolas do distrito de Lisboa.

No que respeita ao número de escolas envolvidas no trabalho de campo, no âmbito das 24 dissertações incluídas na categoria *Concepções e práticas de avaliação de professores*, predominam aquelas que se realizaram junto de uma ou duas escolas, num total de 16 dissertações. Deste subgrupo, quatro não indicam a localização das escolas junto das quais trabalharam, sabendo-se que em duas o trabalho de campo envolveu duas escolas, numa envolveu três escolas e noutra envolveu quatro escolas.

As restantes 14 dissertações incluídas nesta categoria realizaram-se junto de não mais do que quatro escolas, situadas num único distrito, existindo três dissertações que, embora centradas num único distrito (Dias 1999 e Gonçalves, 1997) ou região – Algarve, no caso de Santos (2003) – recolheram informação junto de um elevado número de escolas – entre 15 e 30. Das duas dissertações cujo trabalho de campo decorreu em várias regiões (mais de três) do país, nenhuma dá indicação do número exacto de escolas cujo trabalho envolveu, depreendendo-se, quer pelo número de regiões referidas, quer pelo de participantes envolvidos, que não terá sido inferior a cinco, nem superior a 20. Assim, pode concluir-se que, no conjunto das 24 dissertações inseridas na categoria *Concepções e práticas de avaliação de professores*, predominam os estudos baseados num número muito reduzido de escolas, localizadas num único distrito (14 dissertações), embora três dissertações abranjam um número elevado de escolas num só distrito ou região do país. Apenas duas dissertações se ocuparam com mais de duas regiões, ou seja, envolveram estudos de maior abrangência geográfica embora incluindo um número de indivíduos relativamente pequeno (não superior a 20).

No conjunto de 13 dissertações sobre as *Concepções de avaliação e/ou reacções à avaliação de alunos*, cinco estão incluídas em mais do que uma categoria (Lobo, 1996; Couto, 1997; Rafael, 1998 e Silva, 2000; Araújo, 2002). Os casos de Lobo (1996) e Rafael (1998) foram já explicitados anteriormente. Couto (1997) ocupa-se das *Concepções de avaliação e/ou reacções à avaliação de alunos*, de *Meios de recolha de informação para avaliação* e inclui-se ainda na categoria *outros*, uma vez que procura estabelecer a relação entre a implementação das provas globais e a forma como estas interferem na organização e funcionamento de um grupo disciplinar. Silva (2002) debruça-se sobre *Concepções de avaliação e/ou reacções à avaliação de alunos* e *Práticas de avaliação de professores*. Finalmente, Araújo (2002), além de investigar as *Concepções de avaliação e/ou reacções à avaliação de alunos*, analisa as *Concepções de avaliação de professores*.

Relativamente às fontes de informação, cinco dissertações centraram-se na opinião exclusiva de alunos (Henriques, 1994; Peres, 1995; Mourão, 1997; Silva, 1997; e Braga, 1999); três recolheram informação junto de professores e de alunos num número quase equivalente de ambos os tipos de informantes referidos (Alves, 1998; Leite, 1998; e Araújo, 2002); quatro recolheram informação junto de professores e alunos, embora o número de alunos fosse bastante superior ao de professores (Couto, 1997; Martins, I., 1997; Rafael, 1998; e Silva, 2000); e uma única dissertação recolheu informação junto de dois professores, dois alunos e um encarregado de educação (Lobo, 1996). Este conjunto de dissertações fundamentou-se na informação obtida junto de um número muito diverso de participantes, predominando aquelas em que a amostra é superior a uma centena – oito das 13 dissertações (Henriques, 1994; Peres, 1995; Couto, 1997; Martins, I., 1997; Mourão, 1997; Silva, 1997; Braga, 1999; e Araújo, 2002). No entanto, estes estudos em larga escala basearam-se, na sua maioria (seis dissertações), na recolha de informação junto de uma única escola (Henriques, 1994; Couto, 1997; Martins, I., 1997; Mourão, 1997; Silva, 1997; e Braga, 1999), uma dissertação realizou o seu estudo junto de duas escolas da Grande Lisboa (Araújo, 2002) e outra, em quatro escolas do concelho de Almada (Peres, 1995). Destes oito estudos em larga escala numérica, em termos de amostra, cinco tiveram lugar no distrito de Lisboa. Assim, no conjunto das treze dissertações incluídas na categoria de *Concepções de avaliação e/ou de acções à avaliação de alunos* é possível afirmar que: a sua maioria – oito de 13 – abrange uma amostra superior a 100 indivíduos; a maior parte recolheu informação

apenas junto de alunos (cinco dissertações), ou de alunos e de professores (sete dissertações), existindo ainda uma que também auscultou um encarregado de educação; os estudos realizaram-se maioritariamente numa única escola (10 de 13 dissertações), ou seja, num único local do país, havendo uma que recolheu informação em quatro escolas, todas elas situadas em duas localidades do concelho de Almada; a maioria dos estudos (nove das 13 dissertações) foi realizada no distrito de Lisboa e nos “arredores de Lisboa”. Verifica-se que nesta categoria predominam os estudos em larga escala, em termos de amostra, embora com uma abrangência geográfica muito restrita, essencialmente incluídos na região de Lisboa.

Do conjunto de sete dissertações incluídas na categoria *Práticas de avaliação de professores*, duas foram também incluídas noutras categorias (Diogo, 2001 e Silva, 2000). O caso de Silva (2000) já foi referido anteriormente; quanto a Diogo (2001), ocupa-se também do efeito da avaliação externa sobre o processo de ensino e aprendizagem. Na categoria agora em análise, apenas uma dissertação não ultrapassou os dez indivíduos como fontes de informação (Alaíz, 1996). De facto, seis dissertações trabalharam com amostras de número superior aos 35 sujeitos, tendo, uma delas, auscultado a opinião de 133 professores. É ainda de referir que só duas das sete dissertações se basearam na informação recolhida junto de alunos (Silva, 2000; e de Diogo, 2001). As quatro dissertações restantes só se basearam na informação recolhida junto de professores (Fonseca, 1998 e Martins, F., 1997).

Por outro lado, ao limitarem a sua amostra a professores, o que acontece em quatro das sete dissertações, estão na verdade a questioná-los acerca da ideia que eles têm das suas próprias práticas, não confrontando a sua concepção com a dos alunos, que são testemunhas presenciais do que se pratica dentro das salas de aula. Das sete dissertações inseridas nesta categoria, observa-se que uma se ocupou essencialmente das práticas avaliativas no contexto da avaliação formativa (Martins, F., 1997); duas se ocuparam do efeito da avaliação externa nas práticas de ensino, aprendizagem e avaliação; uma ocupou-se das práticas avaliativas de uma componente do programa da disciplina de Educação Física, a dança (Fonseca, 1998); e outra ocupou-se da avaliação da componente laboratorial, também ela uma parte integrante do currículo das duas disciplinas sobre as quais recaiu o estudo – Ciências Físicas e Químicas (CFQ) e Técnicas Laboratoriais de Biologia (TLB) (Afonso, 2000).

No conjunto das quatro dissertações inseridas na categoria *Concepções de avaliação de professores*, duas estão incluídas em mais do que uma categoria (Afonso, 1999 e Araújo, 2002). O segundo foi já mencionado; Afonso (1999) foi incluído também na categoria *Outros* por fazer uma reflexão acerca da dimensão ética e pedagógica da avaliação. Quanto à amostra, duas basearam-se em informação recolhida apenas junto de professores (Alves, 1997 e Afonso, 1999), uma recolheu informação junto de professores e de alunos (Araújo, 2002) e outra junto de professores e de formadores da formação em alternância (Santo, 2002). Destes quatro estudos, apenas um abrange menos de uma centena de indivíduos como fontes de informação – Afonso (1999) – que trabalhou junto de 24 professores de escolas localizadas numa “zona suburbana de Lisboa”. Das três restantes, uma utilizou uma amostra de 148 professores de três escolas de Setúbal (Alves, 1997), outra auscultou 155 professores e 113 alunos (Araújo, 2002), e a terceira recolheu informação junto de 58 formadores e de 333 professores. Destes estudos, dois realizaram-se no distrito de Lisboa, um na cidade de Setúbal e uma única – Santo (2002) – envolveu a pesquisa junto de 49 escolas distribuídas por três distritos do centro e norte do país (Coimbra, Viseu e Vila Real). Assim, nesta categoria, os estudos são essencialmente abrangentes em número de participantes, mas com um grau restrito de abrangência geográfica.

Prosseguindo com a análise do Gráfico 12, verifica-se a existência de três dissertações incluídas na categoria designada por *Avaliação externa e/ou avaliação interna*, uma delas concluída em 1999, que procurou fazer uma comparação entre a avaliação externa (AE) e a avaliação interna (AI) na disciplina de Biologia, no final do 12º ano. Este estudo decorreu em duas escolas secundárias do distrito de Beja, incluindo um total de 15 professores e 218 alunos, e utilizou os seguintes instrumentos de recolha de dados: entrevistas e inquéritos por questionário aos professores; análise das tabelas de classificação interna final e dos exames de dois anos lectivos consecutivos, relativamente aos alunos (Tavares, 1999). Outra dissertação pertencente a esta categoria foi defendida em 2000, tendo sido o trabalho de campo realizado numa única escola secundária de Setúbal, incluindo uma amostra de oito professores e 937 alunos, com o intuito de investigar a relação entre os resultados da avaliação contínua, os da avaliação sumativa final interna e os do exame, ao longo de um triénio, na disciplina de Matemática (Torres, 2000). Finalmente, a terceira dissertação relacionada com este tema data de 2001, tendo por objectivo ver qual o eventual efeito da AE sobre as

práticas dos professores, recorrendo a uma amostra de 10 professores e de 25 alunos do 12º ano, de várias escolas do distrito de Lisboa. Destas três dissertações apenas uma é incluída em mais do que uma categoria (Diogo, 2001) e esta é aquela que envolve a amostra mais reduzida. Todos estes estudos foram realizados numa região geograficamente restrita, uma só localidade, embora todas elas diferentes – Beja, Setúbal e Lisboa.

Na categoria *Instrumentos* foram incluídas duas dissertações – Couto (1997) e Rodrigues (2001). A primeira teve como intenção investigar até que ponto as provas globais influenciam o trabalho de um grupo disciplinar e, uma vez que se preocupava com questões relacionadas com a construção e conteúdo da prova, foi incluída nesta categoria. A segunda dissertação teve como objectivo analisar a evolução dos exames de Matemática ao longo do último “quarto do século XX”, recorrendo à análise de mais de duas centenas de enunciados de provas (Rodrigues, 2001).

Foi necessário criar a categoria *Outros* por se considerar que existiam aspectos específicos, em cada uma das quatro dissertações aí incluídas, que não se integravam nas restantes categorias. Ou seja, neste grupo de dissertações, metade foi também incluída noutras categorias (Couto, 1997 e Afonso, 1999) como foi salientado acima. Couto (1997) teve a particularidade de procurar descrever até que ponto a imposição de um suporte avaliativo (a prova global) comum a um grupo disciplinar, constituído por quatro professores, afectou a organização e a planificação do trabalho destes, quer em grupo disciplinar, quer individualmente, ao nível do currículo ensinado por cada um. Afonso (1999) fez uma investigação onde procurou, pela auscultação de 24 professores, analisar as reflexões acerca da dimensão ética e pedagógica da avaliação, tendo sido o único estudo a referir a dimensão ética da avaliação como objecto de estudo. No universo das dissertações analisadas, Ferreira (1994) foi a única que intentou, através da análise de documentos legais, fazer uma história da avaliação ao longo de aproximadamente um século, ao nível dos liceus. Finalmente, Gouveia (2000) realizou um estudo junto de 364 alunos do 1º ciclo com a intenção de compreender como a avaliação interfere nas relações entre a escola e a família. Trata-se de um aspecto também único, no conjunto de dissertações analisadas, que importa salientar na medida em que, por vezes, a avaliação sumativa constitui o principal meio de comunicação entre estas duas instituições (escola e família), provocando preconceitos e formas de interacção altamente condicionadas pela interpretação dos resultados dessa avaliação.



Como já foi referido atrás, na maior parte das dissertações investigaram-se *Concepções e práticas de avaliação* (24). Estas dissertações foram concluídas sobretudo entre 1996 e 2003. É de salientar que, embora não se trate do objecto central do estudo das dissertações, os seguintes temas são abordados por um número significativo destas 24 dissertações:

1. Estratégias e meios de recolha de informação para avaliação (Campos, 1996; Neves, 1996; Gil, 1997; Amado, 1998; Leite, 1998; Dias, 1999; Martins, 1999; Ferreira, 2000; Peres, 2000; Silva, 2000; Santos 2002; Santos, 2003);
2. Domínios avaliados (Figueiredo, 1997; Gil, 1997; Leite, 1998; Dias, 1999);  
e
3. Função formativa da avaliação (Alaíz, 1996; Lobo, 1996; Martins, F., 1997; Pereira, 1997; Queirós, 1997; Braga, 1999; e Dias, 1999).

Todos estes aspectos estão, na verdade, interligados, tratando-se de factores centrais numa avaliação para a aprendizagem, o que parece relacionar-se com as inovações do ponto de vista legal, realizadas em 1991 e 1992, em que se preconizava claramente a preponderância da avaliação formativa. O impacto do Decreto-Lei 98-A/92, onde se explicita claramente esta ideia, terá sido o motor de seis dissertações analisadas, que se ocuparam especificamente com a implementação de inovações ao nível legal (Peres, 1995; Araújo, 1996; Lobo, 1996; Alves, 1997; Afonso, 1999; e Torres, 2000), e cuja maioria foi concluída entre 1996 e 1997.

### **Síntese da Resposta à Segunda Questão**

Apenas numa das 48 dissertações não foi possível identificar o problema, embora tenham sido definidos objectivos e questões de investigação. Nas restantes 47, a maioria (44) apresenta o problema de investigação logo no primeiro capítulo e três apresentaram-no em capítulos intermédios. No que respeita à especificação do problema de estudo em questões ou objectivos, apenas três autores não o fizeram de forma a ser possível identificá-los. A maioria dos autores (30) não definiu mais do que cinco questões ou objectivos de investigação, embora seis tenham elaborado entre nove a 12.

Da informação contida no problema e/ou suas especificações, a quase totalidade dos autores (46) indicou os indivíduos alvo de estudo e mais de metade indicou o âmbito disciplinar, bem como o nível de ensino em que o estudo se iria realizar. Há que referir ainda que três autores fornecem indicação acerca do período temporal que o

estudo se propõe abranger, dois deles porque fizeram uma abordagem histórica da avaliação e outro porque compara os resultados das avaliações internas e os da avaliação externa dos alunos de uma escola do ensino secundário, ao longo de três anos lectivos.

Procurando categorizar os tipos de problemáticas em estudo, nas 48 dissertações, constatámos que pouco mais de metade (24) se ocupou das *Concepções e práticas de avaliação de professores*, temática seguida, em número de dissertações que o abordaram, pelas *Concepções de avaliação e/ou reacções à avaliação por parte de alunos* (13). Constatase que a maior parte das dissertações se ocupou dos professores, quer ao nível das concepções e práticas, quer ao nível de cada uma delas individualmente – quatro debruçaram-se sobre as *Concepções de avaliação de professores* e sete sobre as *Práticas de avaliação de professores*.

Treze autores investigaram as *Concepções de avaliação e/ou reacções à avaliação por parte de alunos*, nomeadamente as consequências da avaliação sobre as suas estratégias de estudo ou sobre a sua atitude em sala de aula.

Apenas três dissertações se ocuparam expressamente da *Avaliação externa e/ou avaliação interna*. Seria de esperar um número mais elevado de estudos sobre a problemática da *Avaliação externa e/ou avaliação interna*, já que a avaliação externa tem consequências na transição para níveis superiores de ensino, e ainda pelo seu impacto ao nível dos média e sociedade em geral.

O assunto menos abordado de forma central e exclusiva foi o dos *Meios de recolha de informação para efeitos de avaliação*, do qual apenas dois autores se ocuparam. Este também é um tema que seria de esperar que merecesse maior atenção, já que condiciona fortemente os resultados dos alunos e os domínios e níveis de complexidade avaliados.

No conjunto das 24 dissertações que estudaram as *Concepções e práticas de avaliação de professores* constata-se que metade abordou as estratégias e meios de recolha de informação, sete abordaram a dimensão formativa da avaliação e quatro deram especial atenção aos domínios alvo de avaliação (cognitivo, psico-motor e comportamental).

A referência explícita à implementação de inovações ao nível legal foi feita por seis autores, a maioria dos quais concluiu a sua dissertação nos anos 1996 e 1997. Este aspecto está provavelmente associado à criação e implementação do Despacho normativo nº 98-A/92 e a toda uma série de acções por parte do Ministério da Educação e organismos a ele ligados, como o Instituto de Inovação Educacional, no sentido de promover uma avaliação com um carácter mais formativo, mais partilhado ao nível dos seus agentes e mais transparente no que respeita à partilha de critérios.

### **Caracterização dos Procedimentos Metodológicos**

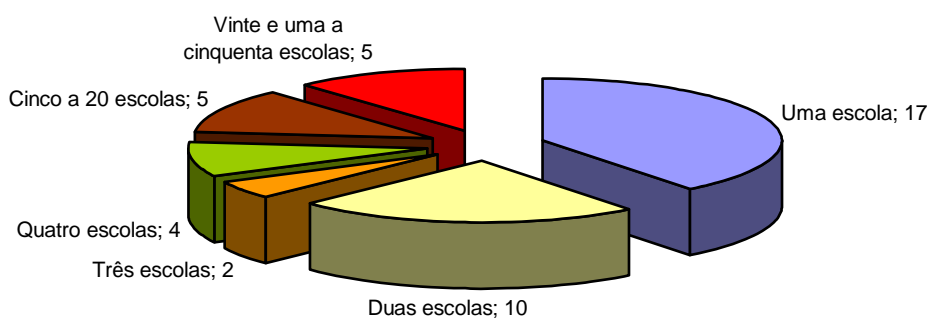
#### **Número de Participantes e de Escolas**

Para dar resposta à terceira questão de investigação – *Quais foram as principais estratégias de investigação e os procedimentos metodológicos utilizados pelos investigadores?* – começou-se por analisar a amostra de cada dissertação. Como já foi referido, aquando da análise do Gráfico 4, três das 48 dissertações em estudo usaram uma amostra constituída essencialmente por documentos. No entanto, no caso de Torres (2000), o trabalho centrou-se ao nível dos alunos de uma escola secundária de Setúbal, enquanto que as duas outras dissertações, Ferreira (1994) e Rodrigues (2001), fizeram um trabalho de análise documental que, devido à sua natureza, não envolveu nenhum tipo de indivíduos ou de instituição de ensino na sua amostra. Deste modo, nesta secção dedicada ao número de participantes e de escolas, estas duas dissertações não serão alvo de análise. Por outro lado, outros três investigadores (Gil, 1997; Pais, 1998; e Diogo, 2001) não indicaram um número exacto de escolas, ou tal não está suficientemente perceptível, pelo que o Gráfico 13 se refere apenas a 43 dissertações das 48 em estudo.

No Gráfico 13 está representada a distribuição das dissertações, tendo em contra o número de escolas em que cada estudo se realizou.

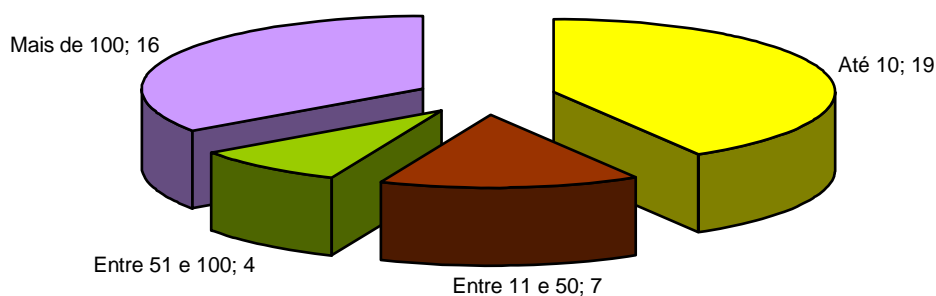
Pode observar-se que, dos 43 estudos em análise nesta secção, 33 ocorreram junto de uma a quatro escolas, sendo que mais de metade (27) destas investigações se concretizou em apenas uma ou duas escolas (17 no primeiro caso e 10 no segundo). O número de dissertações que se desenvolveram em mais de 20 escolas é relativamente inferior (cinco).

**Gráfico 13. Distribuição das dissertações de acordo com o número de escolas envolvidas no estudo.**



O Gráfico 14 mostra a distribuição das dissertações tendo em conta o número de participantes envolvidos em cada uma das investigações. Nesta análise não se incluem Ferreira (1994) nem Rodrigues (2001), pelas razões acima indicadas, referindo-se o Gráfico 14 a um total de 46 dissertações.

**Gráfico 14. Distribuição das dissertações de acordo com o número de participantes nas investigações.**



Da análise do Gráfico 14, é possível constatar que o número de dissertações (19) com uma amostra bastante reduzida, é relativamente próximo do das investigações que envolveram um número relativamente elevado de participantes (16). Ou seja, os dados sugerem que há um certo equilíbrio entre estudos realizados em larga escala, mais abrangentes, mais extensos, e os estudos realizados numa escala mais reduzida, mais focados e, provavelmente, mais intensos.

Refira-se ainda que seis investigações envolveram mais de 51 participantes e menos de 100 e que outras seis envolveram mais de dez e menos de 50 participantes.

De forma a relacionar os dados constantes nos Gráficos 13 e 14, foi elaborado o Quadro 9, onde se apresenta a constituição das amostras dos estudos que envolveram entre uma e quatro escolas. Uma vez que, como já foi referido, duas das dissertações em análise não realizaram estudo empírico junto de escolas, e que Pais (1998) e Diogo (2001) não indicam o número exacto de escolas junto das quais ocorreu o estudo empírico, e Gil (1997), Gonçalves (1997), Martins, F. (1997), Ribeiro (1997), Fonseca (1998), Dias (1999), Afonso (2000), Gouveia (2000), Santo (2002), Ferreira (2003) e Santos (2003) realizaram estudos empíricos abrangendo cinco ou mais escolas, o Quadro 9 refere-se a um total de 33 dissertações.

**Quadro 9. Participantes (professores – P - alunos – A- e encarregados de educação – E) e número de dissertações que envolveram uma, duas, três e quatro escolas.**

Nº de escolas	Nº de participantes	Nº de dissertações por tipo de participantes na amostra	Total de dissertações
1	1 a 10	5 de P; 2 de P e A; 1 de P, A e um E	8
	11 a 50	2 de P e A	2
	51 a 100	0	0
	Mais de 100	2 de P e A; 5 de A	7
2	1 a 10	7 de P	7
	11 a 50	1 de P e A	1
	51 a 100	0	0
	Mais de 100	2 de P e A	2
3	1 a 10	1 de P	1
	11 a 50	0	0
	51 a 100	0	0
	Mais de 100	1 de P	1
4	1 a 10	1 de P	1
	11 a 50	1 de P	1
	51 a 100	0	0
	Mais de 100	1 de P e 1 de A	2

O Quadro 9 mostra que, da totalidade de dissertações que se realizaram junto de estabelecimentos de ensino (43), 72% (33) recorreram a um número inferior a cinco escolas. Destas 33 dissertações, 17, ou seja, cerca de metade, realizou-se junto de uma única escola, constatando-se, neste grupo, um certo equilíbrio entre estudos empíricos que envolveram um número não superior a 10 indivíduos (oito) e aqueles cuja amostra ultrapassou uma centena de indivíduos (sete).

Em termos de constituição da amostra, pode concluir-se que, enquanto que nos estudos que envolvem uma amostra de não mais de 10 elementos (oito), predominam os

que se debruçam apenas sobre professores (cinco), os que envolveram amostras com mais de uma centena de indivíduos, predominam os que inquiriram apenas alunos (5).

No conjunto de 10 dissertações cujo trabalho empírico envolveu duas escolas, sete (70%) utilizaram uma amostra não superior a 10 indivíduos, todas elas recorrendo apenas a professores. Uma única dissertação inquiriu entre 11 a 50 sujeitos e duas dissertações recolheram informação junto de mais de 100 pessoas.

Das duas dissertações cujo trabalho empírico envolveu três escolas, uma trabalhou com uma amostra de três professores e outra com 148 professores. Finalmente, no conjunto de quatro dissertações cujo estudo abrangeu quatro escolas, uma teve como amostra quatro professores, outra, 24 professores e duas tiveram amostras superiores a 100 pessoas.

É de referir que, neste conjunto de 33 dissertações, existe uma concentração de estudos (27) sobre apenas uma ou duas escolas.

O Quadro 10 mostra a constituição das amostras nos estudos que envolveram um máximo de quatro escolas. Neste caso já se inclui Pais (1998) que, embora não especifique o número exacto de escolas sobre as quais incidiu o seu trabalho empírico, o total de participantes que seleccionou (quatro professores) permite integrá-lo neste quadro. Assim, o Quadro 10 diz respeito a um total de 34 dissertações.

**Quadro 10. Constituição das amostras das dissertações que se realizaram junto de não mais de quatro escolas.**

Nº de escolas	Nº de participantes	Nº de dissertações por tipo de participantes na amostra	Total de participantes
Uma a quatro	1 a 10	15 de P, 2 de P e de A, 1 de P, de A e de E	18
	11 a 50	1 de P e 3 de P e A	4
	51 a 100	0	0
	Mais de 100	2 de P, 6 de A e 4 de P e de A	12

Nota: Professores (P), alunos (A) e encarregados de educação (E)

Constata-se, pois, que perto de três quartos das dissertações que fizeram estudos empíricos junto de estabelecimentos de ensino envolveram um número não superior a quatro instituições, verificando-se, mesmo, que, cerca de metade envolveu apenas uma escola.

No que respeita à constituição da amostra, predominam as dissertações onde foram inquiridos apenas professores, sobretudo em amostras não superiores a 10 indivíduos. De facto, a maioria das dissertações que não ultrapassou os quatro estabelecimentos de ensino, teve apenas professores na sua amostra (18). Seguiram-se os estudos que envolveram professores e alunos (9) e, finalmente, aquelas cujo trabalho empírico decorreu unicamente junto de alunos (seis). Enquanto que as amostras constituídas apenas por professores ocorreram, essencialmente, junto de uma única escola (15), as que recorreram exclusivamente a alunos foram estudos com amostras superiores a 100 elementos. Esta disparidade é compreensível se for tido em conta que um único professor interage com um grupo de alunos – normalmente a turma – ao passo que a experiência escolar dos alunos se processa, essencialmente, dentro do grupo turma. Assim sendo, o professor pode constituir, em si mesmo, uma unidade de estudo pela sua singularidade de forma de interagir com grupos de alunos no contexto da tríade ensino, aprendizagem e avaliação.

O Quadro 11 apresenta a constituição das amostras das dissertações que se realizaram junto de mais de cinco escolas, tendo por universo apenas 11 dissertações, uma vez que, como já foi referido, Ferreira (1994) e Rodrigues (2001) não realizaram trabalho de campo junto de escolas, Diogo (2001) não indicou o número de escolas junto das quais fez o seu trabalho empírico e 34 dissertações estudaram um universo não superior a quatro escolas.

**Quadro 11. Constituição das amostras das dissertações que se realizaram junto de mais de cinco escolas**

Número de escolas	Número de participantes	Número de dissertações por tipo de sujeitos na amostra	Número de dissertações
Cinco a 20 escolas	De um a 10	0	0
	11 a 50	3 de P	3
	51 a 100	1 de P	1
	Mais de 100	1 de P e A e 1 de A	2
Mais de 21 escolas	De um a 10	0	0
	11 a 50	0	0
	51 a 100	2 de P e 1 de A	3
	Mais de 100	1 de P e 1 de P e F	2

Nota: Professores (P), alunos (A) e formadores de centros de formação (F)

A análise do Quadro 11 mostra um equilíbrio entre o número de dissertações que se realizou junto de 5 a 20 escolas (seis), e o número daquelas que se realizaram junto de mais de 21 escolas (cinco). No que respeita ao número de escolas, não há, no

conjunto de dissertações em análise, trabalhos que se tenham realizado junto de mais de cinco escolas e cuja amostra não ultrapasse os 10 elementos. Se tal acontecesse, estaríamos, provavelmente, perante um número de cinco ou mais estudos de caso acerca de um ou dois indivíduos, uma vez que o número de escolas teria de ser superior a cinco e o número de elementos da amostra inferior a 11. A condução de um estudo deste tipo – de cinco a 10 estudos de caso – seria certamente muito difícil de levar a cabo, em simultâneo, por uma só pessoa, acrescentando a dificuldade do trabalho ter de ser desenvolvido no contexto de um curso de mestrado, que tem restrições temporais para a sua realização.

Neste conjunto de 11 dissertações, continuam a predominar aquelas cuja amostra é constituída exclusivamente por professores (sete em 11), apenas duas tendo recorrido a professores e a alunos, e uma unicamente a alunos. Paralelamente, constata-se que a maioria de estudos realizados em mais de cinco escolas são constituídos por amostras de 51 a 100 elementos, seguindo-se-lhes, em número equivalente de elementos da amostra, aquelas que inquiriram entre 11 e 50 elementos e as que recorreram a mais de 100 elementos (três dissertações cada).

O Quadro 12 apresenta a constituição das amostras das dissertações em análise, de acordo com o número de elementos que as constituem. Neste Quadro não se incluem Ferreira (1994), nem Rodrigues (2001), uma vez que não envolveram instituições de ensino para a sua elaboração. Assim, o universo em análise no Quadro 12 é de 46 dissertações.

**Quadro 12. Constituição das amostras das dissertações, de acordo com o seu número de elementos.**

Número de participantes	Número de dissertações por tipo de sujeitos na amostra	Número de dissertações
De um a dez	15 de P, 2 de P e A, 1 de P, A e E	18
11 a 50	4 de P e 4 de P e A	8
51 a 100	3 de P e 1 de A	4
Mais de 100	2 de P, 6 de P e A, 7 de A e 1 de P e F	16

Nota: Professores (P), alunos (A), encarregados de educação (E) e formadores de centros de formação (F)

Os dados apontam para a existência de estudos em larga escala realizados num reduzido número de escolas, uma a quatro, ou seja, estudos que pretenderam auscultar uma vasta maioria ou mesmo a totalidade do universo em análise. Já quando o número



de escolas ultrapassa as cinco, quatro dissertações ocupam-se de 11 a 50 participantes, outras quatro de 51 a 100 participantes e mais quatro dissertações de mais de uma centena de participantes. Assim, de uma forma geral, pode-se concluir que à medida que se analisam estudos com um maior número de escolas cresce também o número de indivíduos que, nos casos dos cinco estudos que incluíram mais de 20 escolas, três autores recolheram informação junto de uma amostra que variou entre meia centena e uma centena de indivíduos e dois autores auscultaram mesmo mais de uma centena.

No que respeita à relação entre o tipo de sujeitos incluídos que constituíram a amostra e o número de escolas envolvidas, constatou-se que, nos estudos que se ocuparam de menos de uma dezena de participantes, predominaram as amostras constituídas apenas por professores (15 dissertações em 18) não havendo nenhum estudo que se ocupasse apenas de alunos. Já nos 16 estudos que envolveram mais de uma centena de indivíduos predominaram as amostras constituídas por professores e alunos (sete das 16), seguidas das constituídas apenas por alunos (sete das 16) e pelas constituídas apenas por professores (duas de 16). Assim, verifica-se que, nas dissertações com um elevado número de participantes (mais de 100), o número de dissertações, de acordo com a constituição das amostras, não apresenta valores tão díspares como no conjunto de dissertações que se ocupou de amostras não superiores a dez indivíduos, cuja quase totalidade (14 em 18) se ocupou exclusivamente do estudo de professores.

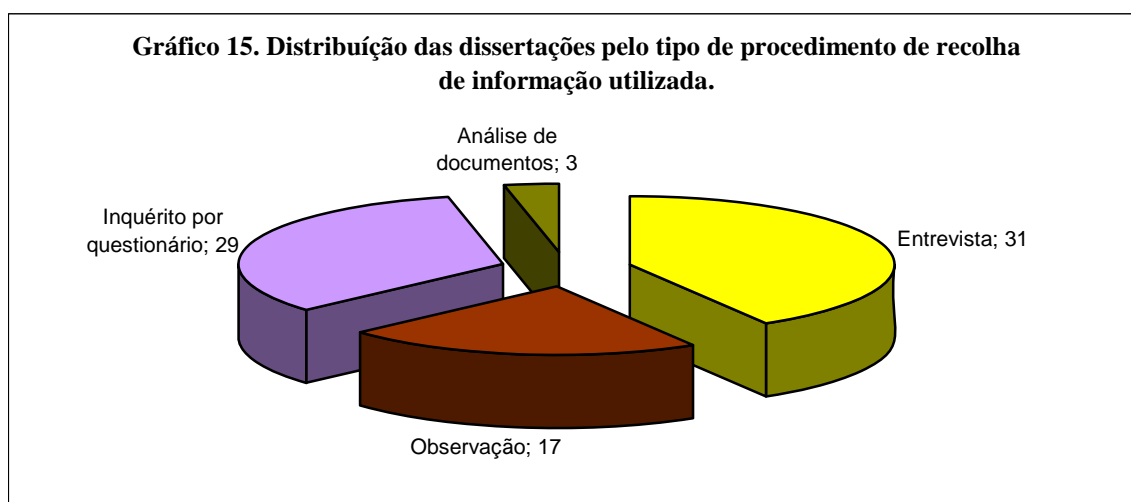
Será interessante relacionar estes dados com as estratégias de recolha de informação, a cuja análise se procederá em seguida.

### **Procedimentos de Recolha de Informação**

A análise das 48 dissertações que foram consideradas nesta investigação permitiu identificar quatro procedimentos primordiais de recolha de informação por parte dos investigadores: o inquérito por entrevista, a observação, o inquérito por questionário e a análise documental. Verificou-se ainda que, na grande maioria dos casos, os investigadores utilizaram mais do que um procedimento. Os Gráficos 15 a 21 apresentam as diferentes estratégias de recolha de informação utilizadas, quer de forma isolada, quer agrupada, e ainda a sua relação com a dimensão da amostra. Não foram aí incluídas as dissertações cuja análise incidiu sobre documentos de origem legal, ou produzidos pelos professores, alunos ou escola (planificações de aulas, documentos de

avaliação escrita, e outros) e cujo objectivo tenha sido apenas o de complementar a informação recolhida através de outros procedimentos de recolha de informação, inequivocamente primordiais.

No Gráfico 15, está indicada a distribuição das 48 dissertações analisadas, de acordo com o tipo de procedimento de recolha de informação utilizado. Uma vez que diversos estudos recorreram a mais do que um procedimento de investigação, a soma dos valores constantes no gráfico é superior ao das dissertações em análise (48). É ainda de referir que na categoria *Análise de documentos* foram incluídas apenas dissertações em que a análise de documentos escritos tenha sido basilar para a investigação, e não uma forma de complementar a informação recolhida através de outros procedimentos.



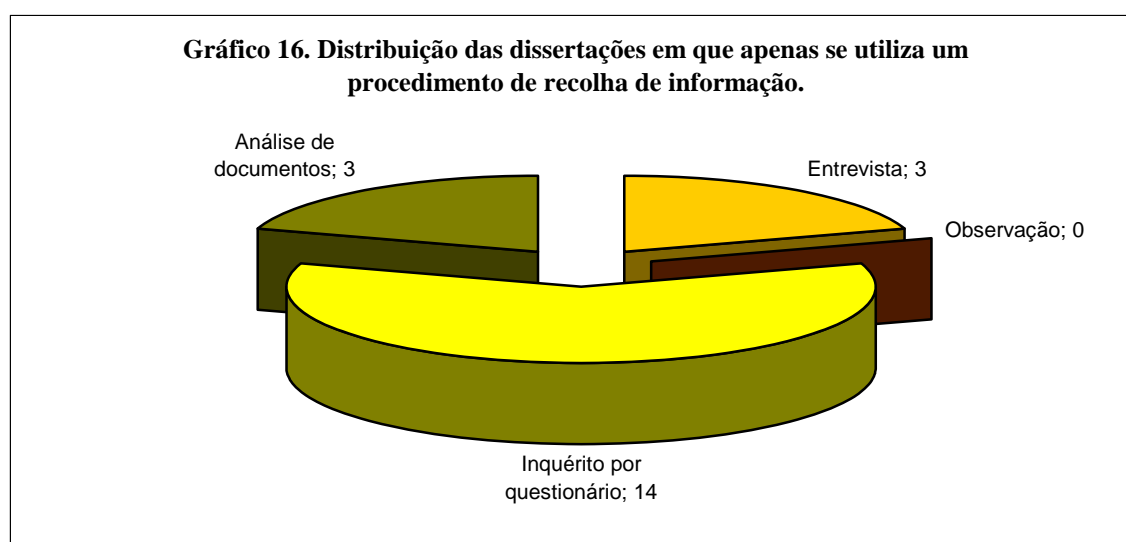
Os dados constantes no Gráfico 15 mostram que o procedimento de recolha de informação mais utilizada foi a entrevista (31), seguida do inquérito por questionário (29) e finalmente da observação (17). Só num número reduzido de investigações o procedimento seguido foi a análise documental (3). É curioso que apenas 17 autores refiram o recurso à observação quando 31 dissertações tiveram como objecto de estudo as *Concepções e práticas de avaliação de professores* e as *Práticas de avaliação de professores* (ver Gráfico 12). Na verdade, para investigar práticas de sala de aula a observação é das estratégias mais recomendadas.

No entanto, neste conjunto de 46 dissertações, que trabalharam junto de estabelecimentos de ensino, apenas 18 têm um objecto de estudo que não envolve mais de 10 indivíduos e dentro destas, 16 foram realizadas com recurso à observação. Assim, constata-se que no caso de estudos que envolvem a recolha de informação junto de um

número reduzido de sujeitos, na quase totalidade dos casos (16 em 18), recorreu-se à observação como procedimento de investigação.

Esta relação entre a dimensão da amostra de cada estudo e o recurso à observação sugere a possibilidade de existência de relações entre a amostra e os diferentes procedimentos de recolha de informação utilizadas. Nesse sentido, começou por se fazer o levantamento da utilização das diferentes procedimentos individualmente e em diferentes conjugações.

No Gráfico 16 apresenta-se a distribuição das dissertações em que se utiliza apenas um procedimento de recolha de informação.



A análise do Gráfico 16 mostra que foram identificadas 20 dissertações em que foi utilizada apenas um procedimento. É de referir, mais uma vez, que situações em que a análise de alguns documentos como forma complementar de recolha de informação, bem como a realização de entrevistas especificamente para ajudar à construção de um outro meio de recolha de informação, não foram incluídas nas categorias *Análise de documentos* e *Entrevista*, respectivamente.

Este gráfico evidencia a distribuição destas dissertações, mostrando que em mais de metade delas foi utilizado o inquérito por questionário. Nestas catorze dissertações, verifica-se que a amostra variou entre os 94 e os 900 indivíduos. Assim, conclui-se que o inquérito por questionário foi utilizado como procedimento único de recolha de informação em situações em que a dimensão da amostra era muito grande. A maioria destas dissertações baseou-se em estudos empíricos que decorreram num único distrito

(duas no de Setúbal, uma no de Santarém, uma no do Porto e duas no de Lisboa), e três destes desenvolveram-se em uma a três escolas (Peres, 1995; Alves, 1997; e Mourão, 1997). Ainda neste grupo de nove dissertações que recorreu apenas ao inquérito por questionário para recolha de informação, apenas três abrangeram mais de dois concelhos (Gouveia, 2000; Santo, 2002; e Rocha, 2003).

Constata-se que o inquérito por questionário foi utilizado como procedimento único de recolha de informação em estudos empíricos que se centraram em amostras numericamente extensas, mas geograficamente restritas a um único distrito (Peres, 1995; Alves, 1997; Dias, 1999; Fonseca, 1998; Gonçalves, 1997; e Mourão, 1997).

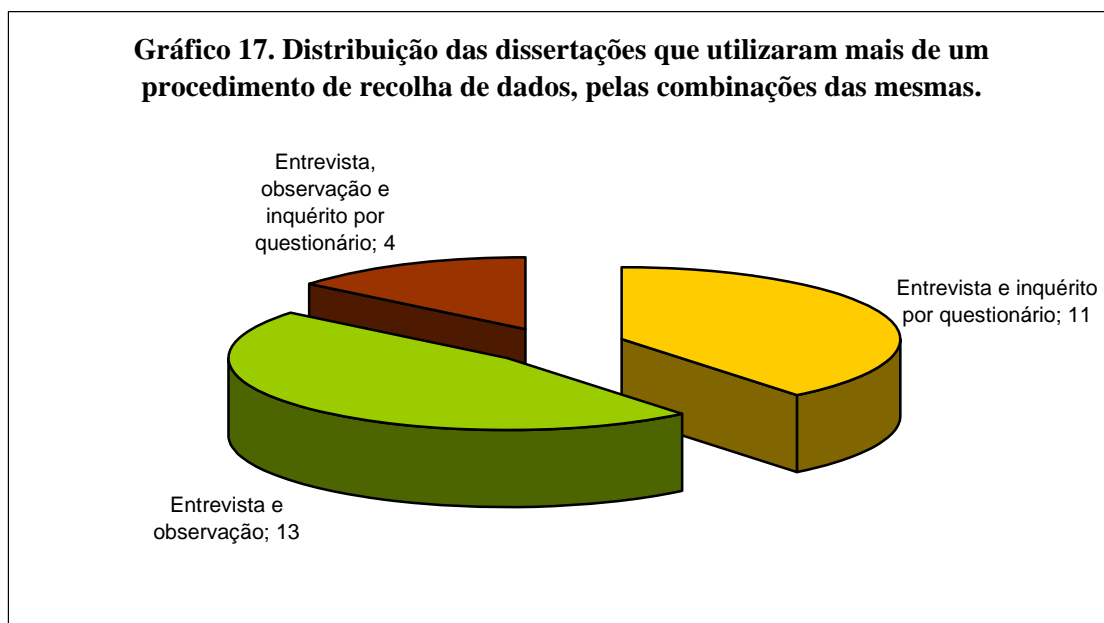
A entrevista foi utilizada como procedimento único de recolha de informação em apenas três das 46 dissertações que se basearam em estudos realizados junto de instituições de ensino. Estes três estudos tiveram amostras de dimensão relativamente reduzida, embora bastante diferentes. No caso de Santos (2002) o estudo centrou-se em três professores de duas escolas do 1º ciclo localizadas no distrito de Lisboa. Já Leite (1998) recorreu a seis professores e a seis alunos de uma mesma escola secundária do distrito de Lisboa. Finalmente, Afonso (1999) seleccionou a amostra com maior dimensão, 24 professores, leccionando em quatro escolas no 1º e no 2º ciclo, no distrito de Lisboa.

Conclui-se então que a entrevista foi pouco usada como procedimento exclusivo de recolha de informação (três dissertações em 48), tendo ocorrido em situações em que a amostra era de reduzida dimensão (três a 24 indivíduos) e geograficamente restritas – todas decorreram no distrito de Lisboa.

A análise de documentos foi utilizada como meio exclusivo de recolha de informação em três das 48 dissertações analisadas (Ferreira, 1994; Rodrigues, 2001; Torres, 2000). Uma delas traça um percurso histórico da avaliação (Ferreira, 1994), outra analisa os enunciados de avaliações externas realizadas durante um determinado período de tempo (Rodrigues, 2001) e uma terceira realizou um estudo comparativo dos resultados dos alunos em avaliações internas e externas de uma única escola secundária de Setúbal, exclusivamente através da análise dos documentos que continham esta informação (Torres, 2000). A análise de documentos foi, pois, usada como procedimento de recolha de informação único num número muito reduzido de dissertações (três em 48), todas elas com propósitos bastante distintos.

Finalmente, é de referir que a observação não foi utilizada como procedimento isolado para a recolha de informação, o que não surpreende, uma vez que esta exige um contacto prévio com elementos da (s) escola (s), como conselho executivo, professores e/ou coordenadores. O recurso a este procedimento surge, no conjunto das 48 dissertações em estudo, a par com a entrevista (em 13 dissertações) ou com a entrevista e o inquérito por questionário (em seis dissertações).

O Gráfico 17 apresenta a distribuição das dissertações de acordo com os agrupamentos de procedimentos de recolha de informação utilizados. Uma vez que 15 dissertações utilizaram apenas um procedimento, o Gráfico 17 apresenta a distribuição das restantes 33.



O gráfico mostra que 24 das 33 dissertações que utilizaram mais do que um procedimento para a recolha dos dados utilizaram dois procedimentos em conjunto, havendo um certo equilíbrio entre as que usaram a entrevista e a observação (13) e as que utilizaram a entrevista e o questionário (11). A conjugação da entrevista com a observação de aulas foi utilizada em situações em que a amostra era pequena. Assim, das 13 dissertações que recorreram a estes procedimentos em conjunto, oito tiveram uma amostra constituída por dois a três indivíduos, 12 tiveram uma amostra inferior a 10 indivíduos, sendo a amostra maior constituída por 18 elementos. Já no caso da conjugação da entrevista com o inquérito por questionário, as amostras incluíram números elevados de indivíduos, sendo a mais pequena constituída por quatro

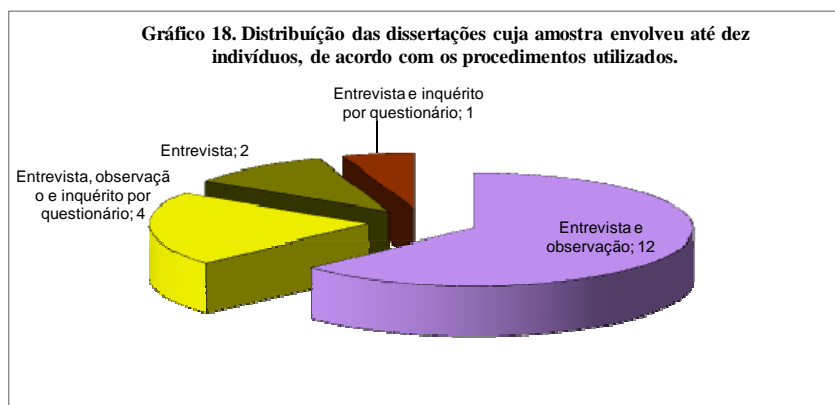
professores (Pais, 1998) e a maior por quatro professores e 386 alunos (Couto, 1997). Neste grupo de 11 dissertações verifica-se que mais de metade, 8, têm amostras com mais de 30 indivíduos e destas, cinco dissertações têm amostras com mais de uma centena de indivíduos.

A conjugação menos utilizada foi a da entrevista juntamente com a observação de aulas e o inquérito por questionário, a que recorreram quatro autores. Deste grupo, duas tiveram uma amostra constituída por três indivíduos, outra recorreu à informação recolhida junto de quatro Professores e outra estudou os sete professores constituintes de um grupo disciplinar. Estes números sugerem que o recurso à observação de aulas foi feito maioritariamente em situações em que a amostra era relativamente reduzida, não ultrapassando os quatro elementos – 10 dissertações em 13 que recorrem à entrevista em conjunto com a observação de aulas e duas em seis que recorreram à entrevista, à observação de aulas e ao inquérito por questionário. Este último meio de recolha de dados, o inquérito por questionário, foi utilizado essencialmente em estudos cujas amostras continham um número relativamente elevado de indivíduos, mais de 30 indivíduos – 10 dissertações em 14 que recorreram à entrevista em conjunto com o inquérito por questionário e duas dissertações em seis que utilizaram a entrevista, a observação de aulas e o inquérito por questionário.

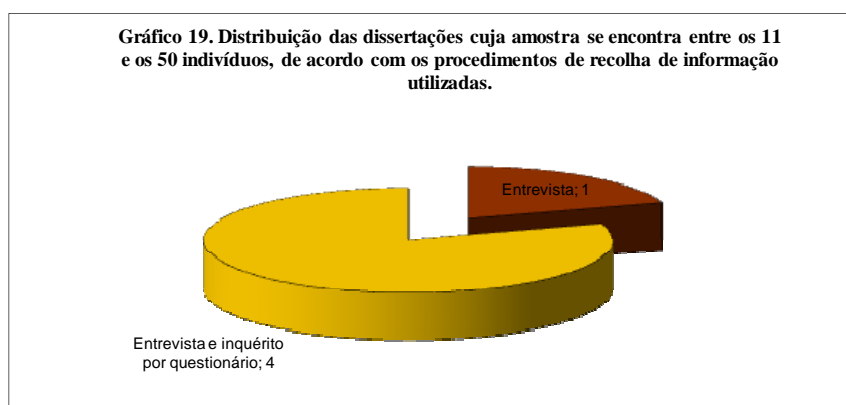
É também interessante sublinhar o facto de em três investigações se ter utilizado a estratégia da entrevista de forma isolada, envolvendo entre 3 e 24 indivíduos na amostra.

Fazendo um cruzamento do número de indivíduos da população estudada com as estratégias de recolha de informação, obtêm-se os Gráficos 18, 19, 20 e 21. O Gráfico 18 apresenta a distribuição das dissertações cuja amostra envolveu até dez indivíduos, de acordo com os procedimentos utilizados.

Da análise do gráfico acima verifica-se que, no conjunto de 18 dissertações que não envolveu mais de 10 indivíduos, em todas foi utilizada a entrevista como procedimento de recolha de dados. Apenas em duas dissertações esta constituiu o único procedimento utilizada e cerca de 67% dos autores recorreram à entrevista conjugada com a observação. Por outro lado, o inquérito por questionário, procedimento normalmente utilizado para inquirir um número elevado de indivíduos, foi usado em quatro dissertações.

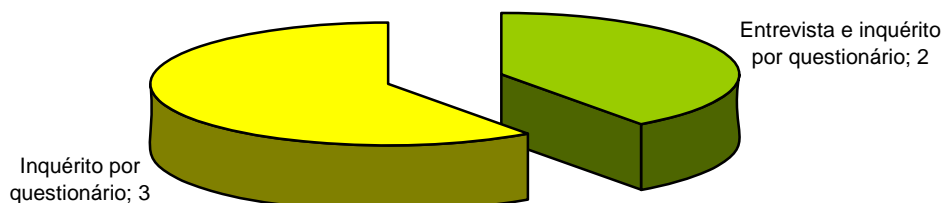


O Gráfico 19 apresenta a distribuição das dissertações cuja amostra se encontra entre os 11 e os 50 indivíduos, de acordo com os procedimentos de recolha de informação utilizadas.



O Gráfico 19, que se reporta a um universo de 10 dissertações, mostra que exactamente metade teve como procedimentos de recolha de informação conjugadas a entrevista e o inquérito por questionário. Também neste grupo de dissertações se verifica que todas recorreram à entrevista, sendo de destacar que duas usaram exclusivamente este procedimento (Leite, 1998 e Afonso, 1999). O Gráfico 20 apresenta a distribuição das dissertações, cuja amostra envolve entre 51 e 100 indivíduos, de acordo com os procedimentos de recolha de informação utilizadas.

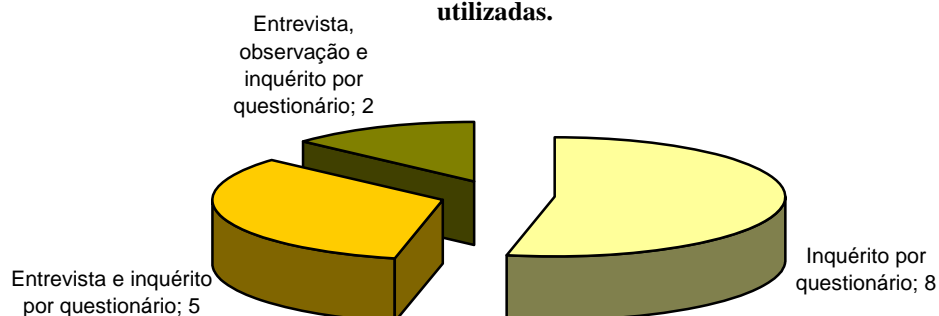
**Gráfico 20. Distribuição das dissertações, cuja amostra envolve entre 51 e 100 indivíduos, de acordo com os procedimentos de recolha de informação utilizadas.**



O Gráfico 20 reporta-se a um universo de apenas quatro dissertações, tendo três recorrido ao inquérito por questionário para recolha de dados e apenas uma (Dias, 1999) utilizou o inquérito por questionário conjugado com a entrevista.

O Gráfico 21 apresenta a distribuição das dissertações, cujas amostras ultrapassam os 100 sujeitos, de acordo com os procedimentos de recolha de informação utilizados.

**Gráfico 21. Distribuição das dissertações, cujas amostras ultrapassam os 100 sujeitos, de acordo com os procedimentos de recolha de informação utilizadas.**



Das 15 dissertações que constituem o universo do Gráfico 21, pouco mais de metade utilizou unicamente o inquérito por questionário, embora este procedimento tenha sido utilizado em todas elas, sendo que em cinco foi conjugado com a entrevista e em duas com entrevista e observação (Henriques, 1995 e Braga, 1998). De salientar, no entanto que o número de entrevistas foi reduzido nestas duas investigações (sete e oito, respectivamente).



A análise dos Gráficos 18 a 21 permite concluir que, à medida que aumenta o número de indivíduos da população estudada, diminui o recurso à entrevista conjugada com o inquérito por questionário. Esta conjugação de procedimentos foi utilizada em mais de metade das dissertações cuja amostra não ultrapassou os dez indivíduos. Paralelamente ao aumento do número de indivíduos da população estudada, aumenta o recurso ao questionário como procedimento de recolha de dados utilizado isoladamente. No grupo de dissertações em que o número de indivíduos da população estudada varia entre os 10 e os 100, em mais de metade é utilizada a entrevista em conjunto com o questionário.

### **Estratégias de Recolha de Informação**

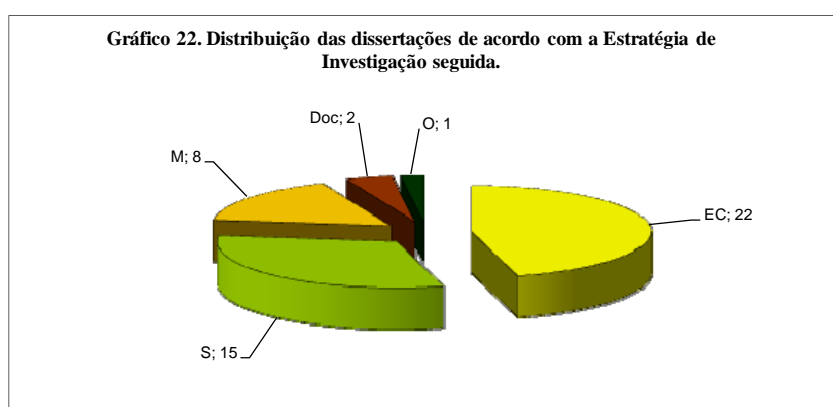
Com o objectivo de classificar as dissertações analisadas em termos da sua estratégia de investigação, tomou-se por base, em primeiro lugar, a forma como cada autor os denominou e caracterizou. Em seguida, associaram-se a esses dados a dimensão da amostra de cada dissertação, bem como os procedimentos de recolha de informação utilizados. Foram, então, definidas as estratégias que a seguir se caracterizam:

- *Investigação Documental (Doc)* – inclui dissertações cuja fonte de informação é única ou quase exclusivamente constituída por documentos escritos e que tem por objectivo caracterizar uma situação enquadrada num determinado período temporal;
- *Estudo de Caso ou Tipo Estudo de Caso (EC)* – inclui dissertações que indicam expressamente tratar-se de estudos de caso ou que, não o fazendo, se debruçam sobre uma amostra restrita, do ponto de vista numérico e da dispersão a nível geográfico, recorrendo essencialmente a formas de recolha de informação que exigem uma relação mais ou menos próxima entre o observador e o observado, como a entrevista e a observação de aulas;
- *Investigação por Sondagem (S)* – inclui dissertações que se debruçam sobre uma amostra de grande dimensão, eventualmente dispersa em termos geográficos, recorrendo a formas de recolha de informação possíveis de implementar com pouco ou nenhum contacto entre o observador e o observado, como o questionário;
- *Investigação Mista (M)* – inclui dissertações onde se utilizaram estratégias de recolha de informação que exigem uma relação entre observador e observado e

conduzem a dados qualitativos, por um lado, e estratégias que não exigem qualquer tipo de relação entre observador e observado e conduzem a dados passíveis de quantificação, por outro. Também se incluíram nesta categoria dissertações em que foi realizada uma análise quantitativa de dados de entrevistas ou de documentos escritos de uma amostra superior a vinte indivíduos;

- *Outras (O)* – inclui dissertações que não se enquadram nas estratégias anteriores.

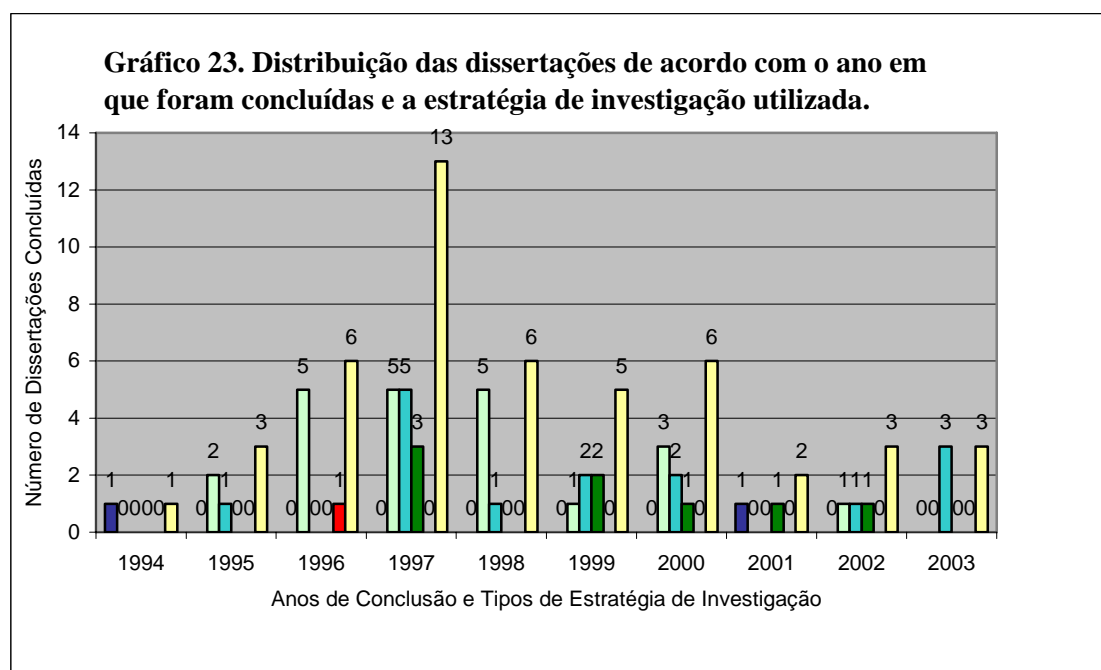
No Anexo C, encontram-se os dados de cada dissertação que serviram de base para a sua inclusão em cada uma das categorias de estratégias. A partir dele, foi construído o Gráfico18, que evidencia o predomínio de *EC* (22 dissertações), seguido dos *S*.



O *EC* foi o mais utilizado, o que está de acordo com a existência de uma maioria de dissertações a investigar uma população não superior a 10 indivíduos, ao elevado recurso à entrevista e observação em conjunto, bem como à existência de três dissertações que recorrem unicamente à entrevista para recolha de dados. Da mesma forma, verifica-se uma maioria de estudos que incidem sobre uma população superior a 100 indivíduos, o que está em consonância com o elevado recurso ao inquérito por questionário, que, em procedimento utilizado isoladamente, constitui mesmo a maioria, bem como à predominância do recurso à entrevista e ao inquérito por questionário em conjunto. Estes dados estão na base do número elevado de *S*, presente neste estudo.

Os dados traduzem um quase equilíbrio entre a tendência para os estudos em profundidade de situações particulares, abrangendo um pequeno número de indivíduos, através de uma metodologia de observador mais ou menos participante – *EC* – e a tendência para estudos com uma grande abrangência numérica e mesmo geográfica, em termos da população alvo, recorrendo a procedimentos com menor participação do investigador no terreno, como o inquérito por questionário utilizado de forma isolada e a entrevista associada ao inquérito por questionário.

O Gráfico 22 apresenta a distribuição das dissertações de acordo com o ano em que foram concluídas e a estratégia de investigação utilizado. Neste gráfico a categoria T refere-se ao número total de dissertações concluídas em cada ano.



Da análise do gráfico constata-se que os *EC* não foram utilizados em apenas três dos dez anos em estudo, 1994, 2001 e 2003. Esta foi ainda a estratégia de investigação mais utilizado em quatro dos dez anos em que foi utilizada e foi utilizada em número idêntico ao misto em dois destes oito anos, 1999 e 2002. Por outro lado os *S* são aqueles que seguem em número de anos de utilização (sete) e em número de dissertações que a ele recorreram. Este tipo de estratégia foi utilizado em dissertações defendidas entre 1995 e 2003. Já as *M* só foram utilizadas em quatro dos dez anos em estudo, fazendo delas recurso um menor número de dissertações do que aquelas que recorreram às estratégias de investigação anteriores. Os anos que apresentam uma maior equivalência entre a utilização das estratégias *EC*, *S* e *M* foram os de 1999, 2000 e 2002.

### **Síntese da Resposta à Terceira Questão**

Das 43 dissertações que indicaram o número de escolas junto das quais se realizou o trabalho empírico, perto de metade (17) fê-lo junto de uma única escola e 33 não ultrapassaram as quatro escolas.

Quanto ao número de indivíduos junto dos quais se realizou a recolha de dados empíricos, 19 dissertações não ultrapassaram os dez e 16 ultrapassaram uma centena. Daqui se observa um certo equilíbrio entre estudos em profundidade junto de um número reduzido de participantes e estudos em larga escala. No entanto, ao analisar simultaneamente o número de escolas e o número de participantes, constatou-se que, das 46 dissertações que realizaram estudos junto de instituições de ensino, naquelas que se realizaram num número não superior a quatro escolas, é muito baixa a discrepância entre as que não ultrapassaram os dez indivíduos na amostra (18 em 33 que realizaram estudos em menos de cinco escolas), e as que auscultaram mais de uma centena de indivíduos (12 em 33).

Das quatro dissertações que seguiram uma temática diferente das restantes, salienta-se uma investigação documental – abordagem histórica da avaliação com base em documentos de carácter legislativo; uma análise dos efeitos da implementação da prova global na dinâmica de trabalho de um grupo disciplinar, com reflexos quer ao nível individual dos professores, quer ao nível do currículo ensinado; uma reflexão sobre a dimensão ética e pedagógica da avaliação e uma análise das repercussões da avaliação sobre a comunicação escola/família.

Sobretudo os dois últimos trabalhos merecem uma especial atenção, na medida em que abordam temáticas pouco investigadas no nosso país e que merecem especial relevância ao nível da literatura internacional. Na verdade, estudiosos como Black e Wiliam (1998) demonstraram que a avaliação pode dar lugar a efeitos perversos, tais como punição e incidência nefasta sobre o auto-conceito do aluno, factores que interferem negativamente na relação aluno/disciplina, aluno/professor e mesmo aluno/escola em geral.

Relativamente à comunicação escola/família, tal como alerta Perrenoud (1999), a avaliação constitui, por vezes, o único veículo de comunicação entre estas duas entidades, daí a necessidade de cautela relativamente à probabilidade de indução de preconceitos, por parte dos encarregados de educação, para com a escola, nem sempre realistas, mas capazes de criar equívocos comunicacionais muito nefastos.

No que diz respeito à análise comparativa das problemáticas em estudo e das metodologias adoptadas, constata-se que dos 24 autores que se propuseram estudar as concepções e práticas de avaliação de professores, nove não fizeram qualquer tipo de observação em sala de aula. Deste grupo, quatro basearam-se somente em entrevistas, dois, em inquéritos por questionário e três, em entrevistas complementadas por inquéritos por questionário. Por outro lado, dos sete autores que se propuseram estudar as práticas de avaliação dos professores, seis também não utilizaram a observação em sala de aula, tendo quatro recorrido à entrevista complementada com inquérito por questionário e dois apenas a questionários realizados junto de professores. Estes factos parecem indiciar uma certa incongruência entre os objectivos propostos e a metodologia utilizada.

## **Capítulo V**

### **CONCLUSÕES E REFLEXÕES FINAIS**

Neste último capítulo, apresenta-se uma panorâmica geral da investigação, relembrando as três principais questões que a orientaram, justifica-se a metodologia seguida e, depois de sumariados os principais resultados, procede-se a uma reflexão em torno de alguns aspectos considerados relevantes, que emergem do estudo e que conduzem a algumas recomendações. Finalmente, são apontados caminhos de investigação futura, com vista ao aprofundamento da problemática da investigação no domínio da avaliação das aprendizagens.

#### **Introdução**

Nas duas últimas décadas, tem sido produzido um número relativamente elevado de trabalhos de reflexão e de investigação centrados na avaliação das aprendizagens, para o que terá contribuído a tomada de consciência dos investigadores, fruto de trabalhos de investigação de uma série de autores, sobre os quais esta dissertação também se debruçou, das repercussões da avaliação sobre as práticas de ensino e estratégias de aprendizagem. Uma revisão da literatura sobre esta temática evidencia que a avaliação de carácter não certificativo tende a ser cada vez mais processual, partilhada e mesmo negociada com os alunos, com recurso a tarefas diversificadas e a contextos informais de avaliação.

Por outro lado, a partir da década de 1980, assiste-se à emergência, com maior consistência teórica, do construtivismo, corrente de pensamento que encara a aprendizagem como uma construção social, em que é dado relevo à atribuição de significados, integrada numa rede de valores e crenças e de numa estrutura mental pré-existent, embora em mutação contínua, fruto da integração de novos saberes. Procura-se, então, a integração dos processos de ensino e de aprendizagem em contextos que se aproximem da realidade envolvente, levando o aluno a empenhar-se em tarefas de natureza experimental. Neste contexto, a avaliação aparece como um processo fundamental das aprendizagens dos alunos, criando condições para o desenvolvimento das suas competências de auto-avaliação e metacognitivas. Desta forma, valoriza-se, de modo particular, a avaliação formativa.

Relativamente ao contexto português, segundo Fernandes (2004, 2005), sobretudo a partir dos anos 1990, assiste-se a um desenvolvimento assinalável do estudo

da avaliação das aprendizagens, sendo de salientar três domínios de análise: políticas educativas; produção de materiais; e investigação. Por outro lado, as orientações emanadas da tutela reforçam a ideia de que a avaliação se deve ocupar de saberes em acção ou em utilização, ou seja, das chamadas competências, importadas do mundo da economia empresarial. A avaliação assume, deste modo, novas dimensões, como a de diagnóstico, com vista a possíveis adaptações curriculares específicas para cada turma ou mesmo para cada aluno. Adicionalmente, é instituída a obrigatoriedade da auto-avaliação por parte dos alunos cuja co-responsabilização e consciencialização do processo avaliativo se pretende, ao mesmo tempo que se preconiza uma maior diversidade de estratégias avaliativas.

Também a partir dos anos 1990, num ambiente marcado por reformas, inovações, reflexões e estudos neste domínio, são criados os primeiros cursos de mestrado em avaliação e, ao longo desta mesma década, em alguns cursos de mestrado com especialização no ensino de disciplinas curriculares, foram produzidas dissertações que se centraram na avaliação das aprendizagens. Assim, surgem estudos sobre as concepções de avaliação de professores; a relação entre as concepções e as práticas de avaliação dos professores; as concepções avaliativas dos alunos; história da avaliação; estratégias avaliativas e/ou domínios objecto de avaliação e ainda sobre avaliação externa.

Pareceu assim pertinente, relevante e útil produzir um trabalho de sistematização da investigação realizada no âmbito de programas de mestrado que podem contribuir para a sua caracterização. Esta foi a principal motivação deste estudo, em que se levou a cabo a análise e síntese de alguns aspectos das dissertações realizadas no âmbito de programas de mestrado sob a responsabilidade de universidades com assento no Conselho de Reitores das Universidades Portuguesas (CRUP), de acordo com a composição deste organismo no ano lectivo de 2002/2003. No presente trabalho, foram analisadas as dissertações de mestrado sobre a problemática da avaliação das aprendizagens, concluídas entre 1994 e 2003, inclusive, de acordo com os critérios de selecção que foram apresentados no Capítulo III. O objectivo foi descrever, tão fielmente quanto possível, os contextos gerais de investigação, as problemáticas em análise e as metodologias utilizadas, num corpo significativo de investigações realizadas nas universidades portuguesas, nos dez anos acima referidos.

Esta dissertação procurou dar resposta a três grandes questões de investigação:

1. Como é que se poderão caracterizar as investigações realizadas?
2. Como se poderão descrever a natureza dos problemas e das questões de investigação das dissertações analisadas?
3. Quais foram as principais estratégias de investigação e os procedimentos metodológicos utilizados pelos investigadores?

A metodologia seguida, sintetizada no Quadro 2 do Capítulo III, é de carácter descritivo e seguiu os passos que se consideram integrantes de uma síntese de investigações. Tal como defendem Suri e Clarke (1999), não houve lugar a qualquer juízo sobre a qualidade das dissertações em análise, tendo-se, antes, optado pela utilização de critérios de inclusão/exclusão de carácter conceptual e não valorativo. Assim, procedeu-se a uma leitura cuidada e subsequente análise de conteúdo dos capítulos referentes à problemática e à metodologia, frequentemente complementadas com a leitura de outras secções das dissertações, sempre que isso se revelou necessário para a recolha de informação indispensável para dar resposta às questões de partida.

A investigação distribuiu-se por cinco grandes etapas. Numa primeira fase, procedeu-se à selecção do conjunto de dissertações a serem alvo da análise. Esta selecção foi realizada com base em três tipos de pesquisa complementares – por meio da Internet, do contacto escrito com as instituições e do contacto directo e presencial com bibliotecas universitárias e a Biblioteca Nacional. O objectivo foi garantir a maior exaustividade de pesquisa e a abrangência da totalidade, ou quase totalidade, das dissertações que iam de encontro aos requisitos do presente estudo. O segundo momento consistiu na elaboração de uma base de dados, assente nas categorias que foram emergindo. Em seguida, os dados já organizados foram analisados e interpretados, à luz do quadro teórico de referência. Finalmente, foram elaboradas sínteses globais da informação analisada, que permitiram dar resposta às questões de partida.

### **Principais conclusões e reflexões**

Mais de metade dos estudos analisados tiveram lugar numa única Universidade – a Universidade Católica Portuguesa – seguindo-se-lhe outras duas, em Lisboa. Nas Faculdades de Ciências da Educação do Porto e de Coimbra não foram concluídas dissertações passíveis de integrar este estudo. Verifica-se, pois, que apenas três



universidades se destacam na produção de dissertações em avaliação. Este aspecto reflecte a relevância que a avaliação assume dentro dos interesses investigativos das Universidades referidas, nomeadamente de Departamentos e Unidades de Investigação.

O pico verificado em 1997 na produção deste tipo de investigação explica-se por uma multiplicidade de factores de natureza diversa – institucional, legislativa e estrutural – de entre os quais terá um peso significativo a abertura, na UCP, no início dos anos 1990, de cursos de mestrado em avaliação.

De salientar ainda que o estudo revelou que a avaliação é investigada indirectamente, através das concepções dos intervenientes e, ainda assim, a relevância recai essencialmente nos professores, ou seja, o indivíduo que aprende raramente está no centro destes processos investigativos. Com efeito, contrariamente àquilo para que a literatura tem vindo a alertar, no sentido da necessidade dos investigadores se debruçarem sobre as práticas avaliativas efectivas, os estudos empíricos analisados neste trabalho que se debruçam sobre essas práticas fazem-no, essencialmente, com base em concepções dos sujeitos que as implementam (os professores), em detrimento da observação de aulas, como estratégia prioritária de recolha de dados.

No que respeita ao contexto disciplinar em cujo âmbito decorreu o trabalho empírico, distinguiu-se a matemática. Este resultado suscita algumas interrogações: se, por um lado, o interesse pelas questões avaliativas na disciplina de matemática poderá ter origem nos fracos resultados que os alunos nela obtêm, por outro lado, será lícito interrogarmo-nos sobre a divulgação dos resultados destas investigações, já que parecem não ter qualquer efeito positivo sobre o desempenho dos alunos. Uma questão que não foi possível analisar no presente trabalho prende-se com a área científica de origem dos autores das dissertações – serão, na sua maioria, professores de matemática?

O ciclo de escolaridade mais estudado foi, de longe, o secundário. Diferentes razões poderão ser apontadas para este facto, entre as quais se salientam as seguintes: contexto profissional dos autores das dissertações; o secundário como ciclo de conclusão dos estudos pré-universitários; o 12º ano como aquele em que os alunos são sujeitos a uma avaliação externa e consequente pressão institucional, política e social aliados.

Em contrapartida, o 1º ciclo é o menos estudado. Se tivermos em conta as frequentes recomendações de uma boa formação de base, necessária às aprendizagens posteriores, frequentemente enfatizada, tanto por estudiosos das questões educativas

como pelos decisores políticos – quer a nível nacional, quer a nível europeu – este resultado torna-se altamente dissonante.

Os resultados da presente investigação revelam também o pouco interesse que a *avaliação externa* tem suscitado nos investigadores, contrariamente ao preconizado pela literatura internacional, quer de carácter teórico, quer de carácter empírico. Este aspecto parece estar relacionado com o facto de, embora este tipo de avaliação existir no nosso país desde há alguns anos, não ter tido o carácter decisivo que tem noutros países com uma forte tradição na implementação de exames standardizados. Deste modo, enquanto na literatura mais recente se debate a existência de um certo paralelismo entre avaliação formativa *versus* avaliação sumativa e avaliação externa *versus* avaliação interna (Shepard, 2001), no nosso país ainda se estão a dar os primeiros passos nos estudos empíricos sobre os eventuais efeitos da avaliação externa sobre a avaliação interna. Constatase, assim, que a avaliação externa, bem como as suas repercussões na avaliação interna, já com um *corpus* sustentado na literatura a nível internacional, ainda está por estudar no nosso país. Do mesmo modo, a avaliação aferida, que conta já com cerca de sete anos de existência, não despertou ainda o interesse da investigação educativa.

Embora o contexto de avaliação mais estudado tenha sido o de sala de aula, ainda assim, a complexa relação entre avaliação formativa e avaliação sumativa, que tem sido objecto de tanta reflexão a nível internacional, carece, no nosso país, de sustentação empírica, pela falta de consistência do estudo das práticas avaliativas.

Ao nível dos procedimentos e estratégias metodológicos, a investigação demonstrou a existência de um quase equilíbrio entre estudos em profundidade, junto de um número reduzido de participantes – estudos de caso – e estudos em larga escala, do tipo sondagem, sendo significativamente menor o número das investigações mistas.

No entanto, é de salientar que a maioria dos estudos de caso carecem de sustentabilidade metodológica, na medida em que a observação – fundamental para se compreender a realidade em análise de forma directa – ou simplesmente não é utilizada, ou é realizada de forma insuficiente. De facto, a realidade de “sala de aula” torna-se fictícia na presença de observadores, até que estes se integrem no grupo, o que só é conseguido após um determinado número de observações (Estrela, 1990 e Sierra, 2006).

Um outro aspecto que constitui um risco para integridade dos dados obtidos e até mesmo conclusões a que a sua análise conduz, numa dissertação, é a relação existente

entre o investigador e os sujeitos que se disponibilizam para constituir fontes de observação, num estudo de caso baseado num número reduzido de sujeitos observados. Efectivamente, trata-se de uma questão de ética investigativa que, embora de difícil controlo, não pode ser descurada, a bem do rigor científico.

Por outro lado, tendo em conta que, ao nível das problemáticas mais abordadas nos estudos analisados, emergiu a preocupação com as práticas avaliativas, constata-se que esta preocupação não tem um paralelo ao nível metodológico, na medida em que os procedimentos de recolha de informação *in loco*, como a observação, são relativamente pouco utilizados, quer em número, quer em duração.

No que respeita aos estudos por sondagem, nem todos tiveram em conta os passos necessários à sua construção. Nos estudos analisados que recorrem ao questionário, nem sempre é clara a existência de entrevistas exploratórias, pré-testagem do questionário, ou a sua apresentação a um painel de especialistas.

É ainda de referir o reduzido número de dissertações que recorreram aos três procedimentos em conjunto – entrevista, observação e inquérito por questionário – que configura uma preocupação com a triangulação dos dados no sentido da garantia de validade interna.

Por fim, refira-se que, no geral, poder-se-á dizer que, sendo inegável o interesse que estes estudos representam para o avanço do conhecimento sobre a avaliação no nosso país, eles não diferem muito significativamente, ao nível dos resultados, talvez por terem abordagens muito semelhantes da temática em análise. Serão de salientar, pela sua novidade e pelo interesse e actualidade de que se revestem, uma análise dos efeitos da avaliação tanto sobre os professores como sobre o currículo ensinado, uma reflexão sobre a dimensão ética e pedagógica da avaliação e, finalmente, um estudo sobre as repercussões da avaliação na comunicação escola/família.

Os resultados deste estudo revelaram que a avaliação das aprendizagens tem vindo a ser, de um modo crescente, objecto de interrogações que têm conduzido a um número já significativo de dissertações de mestrado sobre esta problemática. Sendo a avaliação imprescindível para a qualidade da aprendizagem, podendo, pois, constituir-se como um meio privilegiado de combater o insucesso escolar, será desejável que cada vez mais investigadores se interessem por novas abordagens do tema, ou pelo aprofundamento das já existentes. Também neste campo foi possível observar a fraca

divulgação das dissertações realizadas e inferir da pouca influência dos trabalhos de investigação sobre a determinação das políticas educativas.

Este não é um trabalho que se possa dar por concluído. Limitações de tempo e outras deixaram fragilidades que só poderão ser colmatadas com o continuar da investigação que suscitou novas interrogações, abrindo outras tantas pistas motivadoras relativamente a investigação futura. Esta é uma primeira sistematização da investigação produzida em Portugal sobre avaliação no âmbito de dissertações de mestrado, com vista ao conhecimento do estado da arte nesta temática. No entanto, torna-se da máxima urgência uma análise e síntese das principais conclusões a que este conjunto de dissertações chegou, o que não coube no âmbito deste trabalho. Tendo em conta os resultados a que se chegou, apontam-se como prioritárias investigações sobre a avaliação no primeiro ciclo de escolaridade, a utilização de metodologias de recolha de informação *in loco* e um maior recurso ao aluno como fonte de informação.

## REFERÊNCIAS BIBLIOGRÁFICAS

- Abrantes, P. e Araújo, F. (Orgs.), (2002). *Avaliação das aprendizagens. Das concepções às práticas*. Lisboa: DEB.
- ALAIZ, V.(1996). *Práticas de avaliação formativa e de diferenciação pedagógica*. Tese de Mestrado em Ciências da Educação (Avaliação em Educação). Lisboa. Universidade Católica Portuguesa.
- Allal, L. (2001). La métacognition en perspective. In G. Figari e M. Achouche (Coords). *L'activité évaluative réinterrogée: Regards scolaires et socioprofessionnels*, pp. 104-106. Bruxelles: De Boeck Université.
- Alkin, M.C. e Christie, C.A. (2004). An evaluation theory tree. In M.C. Alkin (Ed.), *Evaluation Roots*. Thousand Oaks, CA: Sage.
- ALVES, I. (1998). *Concepções de alunos de 10 anos acerca da avaliação - quatro estudos de caso*. Tese de Mestrado em Ciências da Educação (Avaliação). Lisboa. Universidade Católica Portuguesa.
- ALVES, J. (1997). *Representações da avaliação por parte dos professores dos 2º e 3º ciclos em três escolas do distrito de Setúbal*. Tese de Mestrado em Ciências da Educação (Avaliação Educacional). Lisboa. Universidade Católica Portuguesa.
- Archbald, D. e Newmann, F. (1992). Approaches to assessing academic achievement. In H. Berlak, F. Newmann, E. Adams, D. Archbald, T. Burgess, J. Raven e T. Romberg (Eds.), *Toward a new science of educational testing and assessment*, pp. 139-180. Albany, NY: State University of New York Press.
- Bangert-Drowns, R. e Rudner, L. (1991). *Meta-analysis in educational research*. ERIC Digest.
- Benavente, A., Alai, V., Barbosa, J., Campos, C., Carvalho, A. e Neves, A. (1995). *Novo modelo de avaliação do ensino básico: formas de implementação local*. Lisboa: Instituto de Inovação Educacional.
- Biggs, J. (1998). Assessment and Classroom learning: A role for summative assessment? *Assessment in Education: Principles, Policy & Practice*, 5, 1, pp. 103-110.
- Birenbaum, M. (1996). Assessment 2000: towards a pluralistic approach to assessment. In M. Birenbaum e F. J. R. C. Dochy (Eds.), *Alternatives in assessment of*

- achievements, learning processes and prior knowledge. Evaluation in education and human services* (pp. 3-29). Boston, MA: Kluwer Academic Publishers.
- Black, P., e Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education: Principles, Policy & Practice*, 5, 1, 7-74.
- Bogdan, R. e Biklen, S. (1994). *Investigação qualitativa em educação: uma introdução à teoria e aos métodos*. Porto: Porto Editora.
- Campos, C. (1996). *Concepções e práticas de professores sobre avaliação das aprendizagens*. Tese de Mestrado em Ciências da Educação (Avaliação Educacional). Lisboa. Universidade Católica Portuguesa.
- Cardinet, J. e Laveault, D. (2001). L'activité évaluative en education: évolutions des préoccupations des deux côtés de l'Atlantique. In G. Figari e M. Achouche (Coords). *L'activité évaluative réinterrogée: Regards scolaires et socioprofessionnels*, pp. 15-29. Bruxelles: De Boeck Université.
- Cohen, L. e Manion, L. (1994). *Research methods in education*. London: Routeledge and Kegan Paul.
- Cole, D., Ryan, C., e Kick, F. (1995). *The road to authentic assessment and portfolios in portfolios across the curriculum and beyond*. Thousand Oaks: SAGE.
- Coll, C., Martín, E., e Onrubia, J. (2001). La evaluación del aprendizaje escolar: dimensiones psicológicas, pedagógicas y sociales. In C. Coll, J. Palacios e A. Marchesi (Coords.) *Desarrollo psicológico y educación 2. Psicología de la educación escolar* (pp. 549-572). Madrid: Alianza.
- Cooper, H. e Lindsay, J. (1998). Research Synthesis and Meta-analysis. In Bickman, L. e Rog, D. *Handbook of Social Research Methods*, pp315-337. London: SAGE Publications.
- Cortesão, L. e Torres, M. (1983). *Avaliação Pedagógica I: Insucesso escolar*. Porto: Porto Editora.
- Cortesão, L. e Torres, M. (1984). *Avaliação Pedagógica II: Perspectivas de Sucesso*. Porto: Porto Editora.
- Crooks, T. (1988). The Impact of Classroom Evaluation Practices on Students. *Review of Educational Research*, 58 (4), pp. 438-481.

- Cumming, J.J. e Maxwell, G.S. (1999). Contextualising authentic assessment. *Assessment in Education*, 6(2), 177-194
- De Ketele, J-M. (2001). Place de la notion de compétence dans l'évaluation des apprentissages. In Gérard Figari e Mohammed Achouche (Coord). *L'activité évaluative réinterrogée: Regards scolaires et socioprofessionnels*, pp. 102-108. Bruxelles: De Boeck Université.
- De Landsheere, G. (1976). *Avaliação Contínua e Exames*. Coimbra: Almedina. (pp. 21-55)
- Estrela, A. (1990). *Teoria e Prática de Observação de Classes: Uma Estratégia de Formação de Professores*. Lisboa: Instituto Nacional de Investigação Científica
- Estrela, A., Marmoz, L., Pires, J. e Pereira, O. (1998). *Investigação e reforma educativa*. Lisboa: Instituto de Inovação Educacional.
- Estrela, A. e Nóvoa, A. (Orgs.) (1999). *Avaliações em educação: Novas perspectivas*. Lisboa: EDUCA, Faculdade de Psicologia e de Ciências da Educação.
- Fernandes, D. (1992). *Prática e perspectivas de avaliação. (Dois anos de experiência no Instituto de Inovação Educacional)*. Documento policopiado.
- Fernandes, D. (1994). Avaliação das aprendizagens: das prioridades de investigação e de formação às práticas na sala de aula. *Revista Educação*, 8, pp. 15-20.
- Fernandes, D. (2004). *Avaliação das aprendizagens: Uma agenda, muitos desafios*. Lisboa: Texto Editores.
- Fernandes, D. (2005). *Avaliação das aprendizagens: Desafios às teorias, práticas e políticas*. Lisboa: Texto Editores.
- FERREIRA, M. (1994). *A classificação dos alunos na legislação escolar: notas para a história da avaliação no ensino liceal (1836-1919)*. Tese de Mestrado em Ciências da Educação (Psicologia da Educação). Coimbra: Universidade de Coimbra – faculdade de Psicologia e de Ciências da Educação.
- Figari, G. (2001). Revue de questions de recherche sur l'évaluation Ou: L'activité évaluative entre cognition et réponse sociale: nouveaux défis pour les évaluateurs. (Artigo não publicado).
- Figari, G. e Achouche, M. (2001). *L'activité évaluative réinterrogée: Regards scolaires et socioprofessionnels*. Bruxelles: De Boeck.

- FIGUEIREDO, M. (1997). *A formação integral do aluno no âmbito da disciplina de português: concepções e práticas de professores do ensino básico – estudo de caso*. Tese de Mestrado em Ciências da Educação (Avaliação Educacional). Lisboa. Universidade Católica Portuguesa.
- Fraenkel, J. e Wallen, N. (1990). *How to Design and Evaluate Research in Education*. EUA: McGraw-Hill.
- FONSECA, A. (1998). *A avaliação da dança no ensino secundário: caracterização da situação no distrito do Porto*. Tese de Mestrado em Performance Artística – dança. Lisboa. Universidade Técnica de Lisboa – Faculdade de Motricidade Humana.
- Gay, L. (1981). *Educational Research – Competencies for analysis & Application*. Columbus, Ohio: Charles E. Merrill Publishing Co.
- Gipps, C. (1994). *Beyond testing: Towards a theory of educational assessment*. Londres: Falmer.
- Gipps, C. (1999). Socio-cultural aspects of assessment. In A. Iran-Nejad e P. Pearson (Eds.), *Review of Research in Education*, 24, pp. 355-392. Washinton, DC: AERA.
- Gipps, C. e Cumming, G. (2003).  
<http://literacyconference.oise.utoronto.ca/papers/gipps.pdf>
- Gipps, C. e Stobart, G. (2003). Alternative assessment. In T. Kellaghan e D. Stufflebeam (Eds.), *International handbook of educational evaluation*, pp. 549-576. Dordrecht: Kluwer.
- Glass, G. (1976). Primary, secondary and meta-analysis research. *Educational Research*, 5, pp. 3-8.
- Glass, G., McGaw, B. e Smith, M. (1981). *Meta-analysis in social research*. Beverly Hills, CE: SAGE.
- Guba, E. e Lincoln, Y. (1989). *Fourth generation evaluation*. London: Sage.
- Guba, E. e Lincoln, Y. (1994). Competing paradigms in qualitative research. In N. Denzin e Y. Lincoln (Eds.), *Handbook of qualitative research*, pp. 105-117. Londres: Sage.
- Hadji, Ch. (1994). *A avaliação, regras do jogo. Das intenções aos instrumentos*. Porto: Porto Editora.



- Harlen, W. (2006). On the relationship between assessment for formative and summative purposes. In Gardner, J. (Ed.). *Assessment and Learning*. London: Sage
- Herman, J. e Golan, S. (1993). The effects of standardized testing on teaching and schools. *Educational Measurement: Issues and Practice*, 12 (4), 20-25, 41-42.
- Huba, M. e Freed, J. (2000). Using portfolios to promote, support, and evaluate learning. In *Learner-centered assessment on college campus*. Boston: Alyn and Bacon, pp. 233-268.
- Hunter, J., Schmidt, F. e Jackson, G. (1982). *Meta-Analysis – cumulating research findings across studies*. London: SAGE Publications.
- Kellaghan, T. e Madaus, G. (2003). External (Public) Examination. In T. Kellaghan e D. L. Stufflebeam (Eds.), *International Handboock of Educational Evaluation* (pp. 577-602). Dordrecht: Kluwer Academic Publishers.
- Kellaghan, T., Madaus, G. e Airasian, P. (1980). *Standardized Testing in Elementary Schools: Effects on Schools, Teachers, and Students*. New York: National Institute of Education.
- Lee, D. e Gavine, D. (2003). Goal-setting and self-assessment in Year 7 students. *Educational Research*, Vol. 45, No. 1, Spring 2003, 49-59.
- Leite, C. e Fernandes, P. (2002). A avaliação da aprendizagem. Novos contextos, novas práticas, Porto: Edições ASA.
- LOBO, A. (1996). *AAA: Aprendizagem Assistida pela Avaliação: um sorriso difícil para o novo sistema de avaliação: percepções de alunos interveniente*. Tese de Mestrado em Educação. Lisboa. Universidade Católica Portuguesa.
- Madaus, G., Haney, W. e Kreitzer, A., (2000). The role of testing in evaluations. In D. Stufflebeam, G. Madaus, e T. Kellaghan, (Eds.) *Evaluation models: Viewpoints on educational and human services evaluation* (2<sup>nd</sup> Edition), pp. 113-126. Dordrecht: Kluwer.
- McMillan, J., Myran, S. e Workman, D. (1999). *The impact of mandated statewide testing on teachers' classroom assessment and instructional practices*. Virginia: Virginia Commonwealth University.
- Marzano, R. (1998). *A Theory-Based Meta-Analysis of Research on Instruction*. ERIC (<http://www.eric.ed.gov>)

- Marsh, C. (1997). *Planning, management and ideology. Key concepts for understanding curriculum*. London: The Falmer Press. (pp.169-182)
- Nevo, D. (1986). The conceptualization of education: An analytical review of the literature. In House, E. (Ed.), *New directions in educational evaluation*, pp. 15-29. London: Falmer.
- Noizet, G. e Caverni, J-P. (1985). *Psicologia da Avaliação Escolar*, pp. 19-21. Coimbra: Coimbra Editores.
- RAFAEL, M. (1998). *Avaliação em matemática no ensino secundário: concepções e práticas dos professores e expectativas dos alunos*. Tese de Mestrado em Educação (Didáctica da Matemática). Lisboa. Universidade de Lisboa – Faculdade de Ciências.
- Revista *Inovação*: Vol. 3, nº 4, (1999). Instituto de Inovação Educacional.
- Revista *Noesis*, nº 23 de 1992.
- Revista *Noesis*, nº40 de 1996.
- RODRIGUES, A. (2001). *Exames Nacionais em Matemática: que espécie de avaliação?* Tese de Mestrado em Supervisão Pedagógica do Ensino da Matemática. Braga. Universidade do Minho.
- Patton, Q. (1990). *Qualitative Evaluation and Research Methods*. London: SAGE Publications
- Perrenoud, Ph. (1999). Não mexam na minha avaliação! Para uma abordagem sistémica da mudança pedagógica. In Estrela, A. E Nóvoa, A. (Orgs.). *Avaliações em educação: Novas perspectivas*, pp. 155-173. Lisboa: EDUCA, Faculdade de Psicologia e de Ciências da Educação.
- Plomp, T., Howie, S. e McGaw, B. (2003). International studies of educational achievement. In Kellaghan e Stufflebeam, D. (Eds.), *International handbook of educational evaluation*, pp. 951-978. Dordrecht: Kluwer.
- Ramos, J. (1989). *Bases pedagógicas de la evaluación. Guía práctica para educadores* (pp. 51-105). Madrid: Síntesis.
- Ribeiro, L. (1997). *Avaliação da aprendizagem*. Lisboa: Texto editora

- Ribeiro, A. e Ribeiro, L. (1998). *Planificação e avaliação do ensino-aprendizagem*. Lisboa: Universidade Aberta.
- Rosenthal, R. (1984). *Meta-Analytic Procedures for Social Research*. London: SAGE Publications.
- Rudner, L., Glass, G., Evartt, D. e Emery, P. (2002). *A User's Guide to the Meta-Analysis of Research Findings*. ERIC (<http://www.eric.ed.gov>)
- Sadler, D. (1998). Formative assessment: Revisiting the territory. *Assessment in Education: Principles, Policy e Practice*, 5, 1, pp. 77-84.
- SANTO, G. (2002). *Concepções implícitas de professores sobre aprendizagem, ensino e avaliação*. Tese de Mestrado em Psicologia. Coimbra. Universidade de Coimbra - Coimbra: Universidade de Coimbra – faculdade de Psicologia e de Ciências da Educação.
- SANTOS, M. (2002). *Os professores do 1º ciclo do Ensino Básico – suas práticas e concepções sobre a avaliação das aprendizagens (estudo de casos)*. Tese de Mestrado em Ciências da Educação (Avaliação em Educação). Lisboa. Universidade Católica Portuguesa.
- Scriven, M. (2003). Evaluation theory and metatheory. In In Kellaghan e Stufflebeam, D. (Eds.), *International handboock of educational evaluation*, pp. 15-30. Boston: Kluwer.
- Shepard, L. (2001). The role of classroom assessment in teaching and learning. In V. Richardson (Ed.). *Handbook of research on teaching* (4<sup>th</sup>. Edition), pp. 1066-1101. American Educational Research Association. New York: Macmillian.
- Sierra, F. (2006). *Métodos de investigación etnográfica en Ciencias Sociales*. Zragoza: Mira Editores
- Smith, M. e Glass, G. (1977). Meta-analysis of psychotherapy outcomes studies. *American Psychologist*, 32 (9), 752-760.
- Stiggings e Conklin (1992). *In teacher's hands: Investigating the practices of classroom assessment*. Albany, NY: State University of New York Press.
- Sprinthal, N. e Sprinthal, R. (1990). *Psicologia Educacional*. Lisboa: McGraw-Hill.
- Sturman, L. (2003). Teaching to the test: science or intuition? *Educational Research*, Vol. 45, No. 3, Winter 2003, 261-273.

- Suri, H. e Clarke, D. (1999). *Revisiting Methods of Literature Synthesis*. ERIC (<http://www.eric.ed.gov>)
- Torrance, H. (1993). Formative Assessment: some theoretical problems and empirical questions. *Cambridge Journal of Education*, vol. 23, No 3, (pp 333-343).
- Tuckman B W (1978). *Conducting Educational Research*. New York: Harcourt Brace Jovanovich.
- Valadares, J. e Graça, M. (1998). *Avaliando para melhorar a Aprendizagem*. Amadora: Plátano Edições Técnicas.
- Van Dalen, D. (1979). *Understanding educational research: an introduction*. New York: McGraw-Hill.
- Wiggins, G. P. (1993). *Assessing Student Performance: Exploring the Purpose and Limits of Testing*. San Francisco: Jossey-Bass Publishers.
- William, D., Lee, C., Harrison, C. e Black, P. (2004). Teachers developing assessment for learning: impact on student achievement. *Assessment in Education*, vol.11, nº 1. USA: Carfax Publishing.

## **REFERÊNCIAS LEGISLATIVAS**

- Lei nº 46/86 de 14 de Outubro. *Diário da República*, I Série, 10598-10601.
- Despacho Normativo nº 162/1991. *Diário da República*, 244, I Série B, 10598-10601.
- Despacho Normativo nº 98-A/1992. *Diário da República*, 140, I Série B, 2908(2)-2908(4).
- Decreto-Lei nº 6/2001. *Diário da República*, 15, I Série A, 258-265.
- Despacho Normativo nº 338/1993. *Diário da República*, 247, I Série B, 5934-5937.
- Despacho Normativo nº 30/2001. *Diário da República*, 166, I Série B, 4438-4441.
- Despacho Normativo nº 1/2005. *Diário da República*, 3, I Série B, 71-76.

# **ANEXOS**

# **ANEXO A**

**Dissertações analisadas no estudo: data, autor e título.**

Ano	Autor	Título
1994	FERREIRA, Maria	A classificação dos alunos na legislação escolar: notas para a história da avaliação no ensino liceal (1836-1919)
1995	GRAÇA, Maria	Avaliação de resolução de problemas - contributo para o estudo das relações entre as concepções e as práticas pedagógicas dos professores
1995	HENRIQUES, Maria	Avaliação – uma estratégia formativa no processo de ensino e de aprendizagem da História
1995	PERES, Maria	Representações sociais da avaliação – o ponto de vista dos alunos de quatro escolas da grande Lisboa
1996	ALAIZ, Vitor	Práticas de avaliação formativa e de diferenciação pedagógica
1996	BOAVIDA, Joaquim	Concepções e práticas de avaliação das aprendizagens de professores do 1º ciclo do ensino básico: três estudos de caso
1996	CAMPOS, Cristina	Concepções e práticas de professores sobre avaliação das aprendizagens – dois estudos de caso
1996	LOBO, Aldina	AAA: Aprendizagem Assistida pela Avaliação: um sorriso difícil para o novo sistema de avaliação: percepções de alunos interveniente
1996	MARTINS, Maria	A avaliação das aprendizagens em matemática: concepções dos professores
1996	NEVES, Anabela	Observação das concepções e práticas de dois professores do ensino básico
1997	ALVES, José	Representações da avaliação por parte dos professores dos 2º e 3º ciclos em três escolas do distrito de Setúbal
1997	BAPTISTA, José	Avaliação de atitudes na escola: concepções e práticas de professores reflectidas em dois estudos de caso
1997	COUTO, Maria	As provas globais e a sua influência no funcionamento de um grupo disciplinar e na uniformização do currículo: estudo de caso
1997	FIGUEIREDO, Maria	A formação integral do aluno no âmbito da disciplina de português: concepções e práticas de professores do ensino básico – estudo de caso
1997	GIL, Dulcineia	Reflexões de professores de língua portuguesa do 2º ciclo do Ensino Básico sobre avaliação de aprendizagens
1997	GONÇALVES, Acácio	Avaliação em Educação Física: o que pensam professores e alunos
1997	MARTINS, Filomena	A influência das práticas de avaliação formativa de professores no processo de ensino-aprendizagem
1997	MARTINS, Irene	Percepções dos alunos sobre avaliação: a sua relação com as estratégias de estudo
1997	MOURÃO, Paulo	O pensamento do aluno: percepções pessoais e crenças sobre o sucesso e insucesso em educação física
1997	PEREIRA, António	Avaliação formativa: uma utopia pedagógica?: modelos teóricos e práticas de professores do 3º ciclo do ensino básico: um estudo de caso
1997	QUEIRÓS, Maria	Avaliar para ensinar e aprender – concepções e práticas avaliativas de professores do 3º ciclo
1997	RIBEIRO, Custódia	Representações do sistema de avaliação de alunos e das práticas avaliativas nos professores de História (a nível do 3º Ciclo de Ensino Básico): um estudo de caso
1997	SILVA, Catarina	Os alunos e a avaliação – concepções dos alunos do ensino básico – 9º ano, relativamente à avaliação das aprendizagens
1998	ALVES, Isabel	Concepções de alunos de 10 anos acerca da avaliação - quatro estudos de caso
1998	AMADO, Nélia	Concepções e práticas de professores de matemática do ensino secundário sobre avaliação: três estudos de caso
1998	FONSECA, Albina	A avaliação da dança no ensino secundário: caracterização da situação no distrito do Porto

1998	LEITE, Maria	Contribuição para uma análise do significado da avaliação do rendimento escolar em ciências naturais: um estudo com professores e alunos
1998	PAIS, Paulo	Práticas classificativas de professores do ensino secundário - significados e valores
1998	RAFAEL, Maria	Avaliação em matemática no ensino secundário: concepções e práticas dos professores e expectativas dos alunos
1999	AFONSO, Maria	Avaliação e ética - um contributo
1999	BRAGA, David	Avaliação na voz dos alunos - representações da prática da avaliação formativa numa escola
1999	DIAS, Luis	Concepções e práticas dos professores de educação física acerca da avaliação formativa
1999	MARTINS, Maria	Práticas avaliativas, momentos formativos: análise das concepções e das práticas avaliativas dos professores de Oficina de Artes.
1999	TAVARES, Teresa	Estudo comparativo entre as classificações das validações interna final e de exame, da disciplina de biologia, do 12º ano de escolaridade: um estudo de caso
2000	AFONSO, Maria	A componente laboratorial e a avaliação das aprendizagens dos alunos: um estudo com professores de Ciências Físico-Químicas e Técnicas laboratoriais de Química
2000	FERREIRA, Odette	O português no ensino secundário: conteúdos e avaliação
2000	GOUVEIA, Manuel	Incidência da avaliação escolar nas relações família-escola em alunos do 4º ano de escolaridade
2000	PERES, Ana	Concepções e práticas avaliativas de professores de matemática do 2º ciclo do ensino básico
2000	SILVA, José	Avaliação de investigações matemáticas: uma experiência
2000	TORRES, Maria	Avaliação de resultados em matemática no contexto da reforma curricular do ensino secundário
2001	DIOGO, Ana	Efeitos da avaliação externa nas práticas de ensino/aprendizagem – contributo para o seu estudo no Ensino Secundário em Portugal, a exemplo da disciplina de Alemão no 12º ano.
2001	RODRIGUES, Ana	Exames Nacionais em Matemática: que espécie de avaliação?
2002	ARAÚJO, Filomena	A avaliação no ensino básico – o que pensam alunos e professores
2002	SANTO, Gisela	Concepções implícitas de professores sobre aprendizagem, ensino e avaliação
2002	SANTOS, Maria	Os professores do 1º ciclo do Ensino Básico – suas práticas e concepções sobre a avaliação das aprendizagens (estudo de casos)
2003	FERREIRA, Cristina	A avaliação das aprendizagens no trabalho laboratorial em biologia: uma proposta para o ensino secundário
2003	ROCHA, Carla	A avaliação das aprendizagens no 1º ciclo do ensino básico: das crenças aos instrumentos
2003	SANTOS, Maria	Concepções implícitas dos professores e práticas de avaliação



## **ANEXO B**

**Dissertações analisadas no estudo: data, autor e informação acerca das fontes de informação.**

ANO	AUTOR	REGIÃO	ESCOLAS	CICLOS	ANOS	PROFESSORES	ALUNOS	DISCIPLINAS	OUTROS
1994	FERREIRA, M. Edite	Não menciona ou não se aplica (NM/NA)	NM/NA	Secundário (liceus)	NM/NA	NM/NA	NM/NA	Português	Legislação: 66 documentos da Colecção Oficial de Legislação Portuguesa; 3 documentos do Diário do Governo.
1995	GRAÇA, M. Margarida .	NM/NA	4	3º	8º e 9º	4	NM/NA	Matemática	NM/NA
1995	HENRIQUE S, M. da Encarnação	Zona Litoral suburbana	1	3º	7º, 8º e 9º	NM/NA	Cerca de 190	História	NM/NA
1995	PERES, M. João	Concelho de Almada (2 localidades)	4	3º	9º (todos os alunos)	NM/NA	427	Matemática	Documentos oficiais: legislação:4.
1996	ALAIZ, Vitor	Periferia Grande Lisboa	1	2º	5º e 6º	4	(5 turmas)	Matemática	NM/NA
1996	BOAVIDA, Joaquim .	NM/NA	2	1º	2º, 3º e 4º	3	NM/NA	NM/NA	NM/NA
1996	CAMPOS, Cristina	Próxima de Lisboa	1	2º e 3º	5º, 6º e 7º	2	NM/NA	Português, Inglês e HGP	NM/NA
1996	LOBO, Aldina	Grande Lisboa	1	3º	8º	2	2	Português e CN	1 Encarregado de Educação
1996	MARTINS, M.da Paz	Grande Lisboa	1	3º	7º, 9º, 10º e 11º	3	NM/NA	Matemática	NM/NA
1996	NEVES, Anabela	Grande Lisboa e margem sul do Tejo	2	2º e 3º	5º, 6º e 9º	2	86	HGP, Português e Francês	NM/NA
1997	ALVES, José	Setúbal	3	2º e 3º	NM/NA	148	NM/NA	NM/NA	NM/NA
1997	BAPTISTA, José	Lisboa	2	Secundário	NM/NA	2	NM/NA	Economia, sociologia e várias ligadas	NM/NA

								à química	
1997	COUTO, Maria	Distrito de Lisboa	1	Secundário	10º e 11º	4	386	Filosofia	NM/NA
1997	FIGUEIREDO, M. Cecília	Arredores de Lisboa	1	3º	8º e 9º	3	2 turmas do 8º e 1 do 9º	Português	Todos os instrumentos de avaliação utilizados com os alunos.
1997	GIL, Dulcineia	-Lisboa, Alverca, Caldas da Rainha, Seixal e Castelo Branco	Pelo menos 5	2º	5º e 6º	20	NM/NA	Português	Instrumentos de avaliação
1997	GONÇALVES, Acácio	Concelhos da Amadora e de Sintra	15	Secundário	10º, 11º e 12º	205 universo);	766	Educação Física (a amostra inclui professores de todas as outras disciplinas do secundário)	NM/NA
1997	MARTINS, Filomena	Concelho Almada	8	2º	NM/NA	71	NM/NA	Matemática e CN	NM/NA
1997	MARTINS, Irene	Periferia Lisboa	1	2º e 3º	6º e 9º	2	359	HGP História	1 professor responsável pela sala de estudo e pertencente ao observatório da qualidade; 1 professor de Ensino. Especial.
1997	MOURÃO, P.	Zona norte do distrito de Santarém	1	Secundário	10º, 11º e 12º	NM/NA	315	Educação Física	NM/NA
1997	PEREIRA, António	NM/NA	3	3º	8º e 9º	3	NM/NA	Francês	NM/NA
1997	QUEIRÓS,	Concelho da área da	2	3º	7º, 8º e 9º	7	NM/NA	Português,	NM/NA

	M. Liseta	Grande Lisboa						Matemática e CN	
1997	RIBEIRO, Custódia	Várias nos Arredores Lisboa, uma ligada ao Ministério da Defesa no centro de Lisboa, uma urbana no Minho e duas urbanas e uma em meio rural, no Alentejo	Entre 5 a 15	3º	NM/NA	15	NM/NA	História	NM/NA
1997	SILVA, Catarina	Concelho limítrofe de Loures	1	3º	9º	NM/NA	147 (todas turmas 9º ano)	NM/NA	NM/NA
1998	ALVES, Isabel	Freguesia do distrito de Lisboa	1	Secundário	10º	2	4	Português e Matemática	NM/NA
1998	AMADO, Nélia	N	2	Secundário	11º 12º	3	NM/NA	Mat	Doc oficiais e dos professores
1998	FONSECA, Albina	Distrito Porto	29	Secundário	NM/NA	133 de Dança	NM/NA	Dança	NM/NA
1998	LEITE, M. Luísa	Lisboa	1	Secundário	NM/NA	6	6	CN	NM/NA
1998	PAIS, Paulo	Grande porto	Pelo menos duas	Secundário	11º e 12º	4	NM/NA	Matemática, Português, Inglês e Filosofia	NM/NA
1998	RAFAEL, M. Amélia	Arredores de Lisboa	1	Secundário	10º	3	15	Matemática	NM/NA
1999	AFONSO, M. Rosa	Suburbana de Lisboa	4	1º e 2º	NM/NA	24	NM/NA	Português, HGP, CN, Matemática, Música, EVT	NM/NA
1999	BRAGA, David	Macieira (distrito de Braga)	1	3º	9º	NM/NA	261	NM/NA	NM/NA
1999	DIAS, Luís	Distrito de Lisboa	30	2º, 3º e Secundário	NM/NA	94	NM/NA	Educação Física	NM/NA

1999	MARTINS, M. da Graça	Centro do país (cidade ribatejana)	1	Secundário	11° e 12°	3	4	Oficina de. Artes	NM/NA
1999	TAVARES, Teresa	Concelho Beja	2	Secundário	12°	15	218	Biologia	NM/NA
2000	AFONSO, M. Deolinda	Todo o país	36	Secundário	10°, 11° e 12°	77	NM/NA	CFQ TLB	NM/NA
2000	FERREIRA, Odette	Arredores Lisboa	1	Secundário	10°, 11° e 12°	2	NM/NA	Português e Latim	NM/NA
2000	GOUVEIA, Manuel	Concelho de Cascais, Oeiras, Bragança, Alfândega da Fé, Mogadouro e Macedo de Cavaleiros	19	1°	4°	NM/NA	364	NM/NA	NM/NA
2000	PERES, Ana	Arredores Lisboa	2	2°	5° e 6°	3	NM/NA	Matemática	NM/NA
2000	SILVA, José	Arredores Lisboa	2	Secundário	10°	2	(2turmas)	Matemática	NM/NA
2000	TORRES, M. Júlia	Setúbal	1	Secundário	10°, 11° e 12°	8	937.	Matemática	Fichas biográficas alunos; pautas; Dossiers de DT de cada Turma; Dossier do grupo de Matemática; Livros de ponto; Registo biográfico dos professores.
2001	DIOGO, Ana	Distrito de Lisboa	várias	Secundário	12°	10	25	Alemão	NM/NA
2001	PEREIRA, M. da Graça	NM/NA	NM/NA	NM/NA	NM/NA	28	NM/NA	Matemática	NM/NA
2001	RODRIGUE S, Ana	NM/NA	NM/NA	Secundário	12° e propedêutico	NM/NA	NM/NA	Matemática	Provas de Av de ano

									propedêutico:1; Provas Específicas: 5; Provas Escritas:3; Provas de Aferição:2; Provas de Exame Nacional:7; Provas 235 e equivalentes:11 ; total =29
2002	ARAÚJO, Filomena	Grande Lisboa	2	2º e 3º	6º e 7º	155	113	Português, Matemática, Língua Estrangeira, Educação Visual, EVT, Educação Física	NM/NA
2002	SANTO, Gisela	Região Norte: Distrito Vila Real, Viseu, Coimbra	49 e 5 Centros de Formação de Professores	Secundário	NM/NA	333 + 58 formadores	NM/NA	Português, Matemática, Biologia, História, FQ e Filosofia	5 centros formação e 58 formadores
2002	SANTOS, M. de Fátima	Região de Lisboa	2	1º	1º e 3º	3	NM/NA	NM/NA	NM/NA
2003	FERREIRA, Cristina Leite	Distrito de Aveiro	31	Secundário	N	61	NM/NA	Biologia	NM/NA
2003	ROCHA, Carla	Lisboa e Algarve	4	1º	1º, 2º, 3º e 4º	104	NM/NA	NM/NA	NM/NA

2003	SANTOS, M.	Algarve	17	Secundár io	NM/NA	85	NM/NA	Informática	NM/NA

## **ANEXO C**

**Dissertações analisadas no estudo: data, autor e informações acerca da forma de recolha de dados.**



ANO	AUTOR	ENTREVISTAS	OBSERVAÇÕES	QUESTIONÁRIO	DOCUMENTOS	OUTROS	NOTAS DE CAMPO
1994	FERREIRA, M. Edite	NM/NA	NM/NA	NM/NA	66 documentos da Coleção Oficial de Legislação Portuguesa e 3 documentos do Diário do Governo	NM/NA	NM/NA
1995	GRAÇA, M. Margarida .	- 16: 4 a cada prof, semi-estruturadas, áudio-gravadas.	-16 aulas: 4 de cada prof, com investigadora não interveniente, registadas.	NM/NA	NM/NA	- Conversas informais com registos.	SIM
1995	HENRIQUES, M. da Encarnação	7 estruturadas (já no 9º ano); e registo escrito audi-gravadas.	participante, ao grupo 1 e à turma já no 8º ano	1º ano do estudo os dois grupos responderam a um inquérito no início do ano lectivo final do 1º ano do estudo o mesmo inquérito com algumas alterações ao 1 grupo 2º ano do estudo só a uma turma do 8º ano com alunos que continuaram no estudo (19), resposta orientada+escolha múltipla	Teste referente a critérios no final do ano aos dois grupos, só com perguntas de escolha múltipla	- 296 referências escritas de alunos sobre o que o que era para eles a avaliação.	NM/NA
1995	PERES, M. João	NM/NA	NM/NA	- pré-teste a 40 alunos.	- oficiais (4)	- Grelhas de categorização.	NM/NA
1996	ALAIZ, Vitor	- informais várias, regulares; - 4 semi-abertas aos	- 10 não focadas,, não participantes de 4 turmas e de todos os	NM/NA	- fichas de trabalho, fichas de avaliação, sumários das aulas, pautas de av	- Construção de grelha de análise de ambientes de salas de aula;	- Diários de bordo

		alunos	professores; - 50 aulas com gravação em vídeo. - observações naturalistas, em ambientes informaisd, dentro do espaço escola.		sumativa; - docs da escola; - cadernos dos alunos.	- Seminários quinzenais entre Dezembro e Abril com as profs; - Diário das aulas feito pelas professoras.	
1996	BOAVIDA, Joaquim .	- 6: 2 a cada prof., semi-estruturadas, cerca de 60', áudio-gravadas, transcritas e lidas; - entrevistas semi-estruturadas às directoras das duas escolas.	- Cc 17 horas, audio-gravadas e com registo descritivo.	NM/NA	NM/NA	- Esboço das planificações.	SIM
1996	CAMPOS, Cristina	2 a um professor de 90'; 3 a um professor entre 45' e 60'; 1 aos dois professores de 3h; semi-estruturadas, áudio-gravadas, transcritas, protocolos lidos pelos entrevistados TOTAL: 6	Naturalista; 33h: 4 a 5 aulas de cada turma.	NM/NA	2 textos de cada prof acerca da avaliação em geral e sobre instrumentos de avaliação em particular; vários instrumentos de avaliação (testes, registos de observação, inventários, outros); portfólios de alguns alunos; pautas de classificação das turmas observadas dos 2º e 3º períodos.	Reacção a cartões com frases sobre práticas avaliativas	SIM
1996	LOBO, Aldina	- 4: duas a cada prof; - 1 de conjunto alunas; - 1 EE;	- 8: 4 de LP e 4 de CN, praticamente não participante, abertas.	NM/NA	- planificações, grelhas de registo, fichas de controlo da ap.; testes; redacções,	- Conversas informais, observações ocasionais, conversas telefónicas.	NM/NA

		- todas gravadas.			outros; - documentos da escola.		
1996	MARTINS, M. da Paz	- 6: duas a cada prof, gravadas; 2 a 2,5 horas cada, semi-estruturadas, transcritas e dadas a ler; - as 2ºs entrevistas entre 1 a 1,5 horas.	- participante, com registo escrito, gravadas em vídeo - 12 a 16 em cada turma do 11º (56); - 1 de cada turma de outros anos – 6.	NM/NA	- Fichas de trabalho usadas nas aulas; - Fichas de avaliação.	NM/NA	NM/NA
1996	NEVES, Anabela	Áudio-gravadas, 30 a 50', transcritas; semi-estruturadas 6 a um professor, de 30', com protocolo revisto pelo professor; 4 a outro professor, 45 a 50', com protocolo revisto pelo professor; 1 a cada presidente do CE de cada escola.=2	*num sistema narrativo aberto, sem categorias pré-definidas, não participante, com registos num diário de campo um professor - 11 na turma de HGP e 10 na turma de LP outro professor – 14 de uma turma total= 35	NM/NA	*3 documentos sobre instrumentos de registo de observação (do IIE) sobre os quais os professores se pronunciaram textos produzidos pelos dois professores acerca da avaliação, instrumentos de avaliação e instrumentos de registo de observação; fichas de avaliação (testes) nas turmas observadas; curriculum vitae dos professores; pautas classificativas das turmas observadas.	comentário a cartões (3) com frases acerca das práticas avaliativas	SIM – uma parte destinada às descrições dos fenómenos observados e outra reservada a comentários e dúvidas do observador.
1997	ALVES, José	- 8 informais.	NM/NA	- 148: predominância de perguntas fechadas de tipo Likert de 4 níveis e algumas questões abertas; - houve um pré-questionário.	- Doc oficiais.	NM/NA	NM/NA
1997	BAPTISTA,	Semi-estruturadas,	Não participante, com	NM/NA	Escalas de atitudes usadas	NM/NA	NM/NA

	José	baseada num guião, audio-gravadas e transcritas com protocolos revistos, intercaladas de observação de aulas 3 a uma professora e 4 a outra TOTAL 7	grelha construída, inspirada em dois dos EUA 8 aulas de uma professora e 10 da outra TOTAL 18		por uma prof, fichas de avaliação (testes) das turmas observadas, registos de observação realizados pelas professoras Documentos elaborados pela escola, pautas das notas e curriculum das duas professoras.		
1997	COUTO, Maria	semi-estruturadas, com guiões, áudio-gravadas 2 ao delegado: 1ª exploratória de 75', 2ª no final de 15'; Profs das PG do 10º: 1 a cada, individuais, 45' e 50'; Profs da PG do 11º: 1 a um Prof 1 a grupo de 4 alunos; TOTAL:6	NM/NA	Objectivo de entender como os alunos tinham sentido as PG; Submetido à opinião de 4 profs dos alunos alvo, revisto pelo orientador, prétestados a alunos que iam realizar outras PG(95 alunos); Elaboração de questionário respondido logo a seguir à realização das PG; 11º responderam 169 dos 197 alunos; 10º responderam 212 dos 252 alunos Constituído por 26 questões em escala de Likert de 5 níveis; TOTAL: 381= 212(10º)+169(11º)	NM/NA	NM/NA	SIM
1997	FIGUEIREDO, M. Cecília	Semi-estruturada, guião flexível com questões adaptadas e	observação sistemática, apoiada em grelha de observação e redacção	1 após a primeira entrevista	todos os documentos de avaliação usados pelas professoras	NM/NA	NM/NA

		outras novas 2 entrevistas diferentes, individuais ?	de relatórios após a observação não participante TOTAL DE 27				
1997	GIL, Dulcineia	- 8 (2 entrevistas a 4 profs) semi- estruturadas, áudio- gravadas e transcritas.	NM/NA	- 20 abertos pré- testados; - 20 fechados.	- 40 instrumentos de av; - 20 testes de avaliação	NM/NA	NM/NA
1997	GONÇALVES , Acácio	NM/NA	NM/NA	1ª versão revista por grupo de profs “com muita exp nestas questões”; parte foi adaptada de pré-existente; foram feitos 3 tipos adaptados aos 3 tipos de amostra, mas com os mesmos objectivos, ou seja, muito semelhantes; foi feito pré-teste junto de 9 profs de EF e 14 alunos TOTAL 130+75+766=9 71	NM/NA	NM/NA	NM/NA
1997	MARTINS, Filomena	Entrevistas exploratórias, semi- estruturadas, com guião construído; A 2 professores do 4º grupo, um a leccionar só Mat e outro só CN; 30', áudio-gravadas, transcritas e protocolos revistos;	NM/NA	Construção de pré- questionário, essencialmente só questões fechadas, validado por 3 profs sendo dois deles mestres em CE, depois por 6 profs com características da amostra e depois por	NM/NA	Encontros com CE e Delegados disciplinares das diferentes escolas para explicar o trabalho que se pretendia fazer	NM/NA

		análise de conteúdo das mesmas TOTAL: 2		dois doutores em CE; Versão final com 26 questões sendo apenas uma de resposta aberta dado a todos os profs do 4º grupo das escolas; TOTAL: 70			
1997	MARTINS, Irene	- Entrevistas formais às profs, gravadas, transcritas e dadas a ler, semi-estruturadas; - 1 à prof do ens especial; - 1 a prof da sala de estudo e observ. Da qualidade; - entrevistas informais a pares de alunos para selecção dops mesmos; - 1 entrevista semi-estruturada a cada al = 8, gravadas e transcritas; - 3 entrevistas a cada aluno não gravadas = 24; - 1 entrevista a 3 alunos extra do 9º ano, semi-estruturada e transcrita.	NM/NA	- 348 a alunos.	NM/NA	- Diários feitos pelos alunos (4 do 6º e 1 do 9º)	NM/NA
1997	MOURÃO, PAULO	NM/NA	NM/NA	Foram inquiridos 341 e só 315 foram aceites	NM/NA	NM/NA	NM/NA
1997	PEREIRA, António	Uma entrevista a cada professor, gravada,	Elaboração de grelha para a observação e	3 a cada professor para completar informação	NM/NA	Documentos adaptados ou produzidos pelos	NM/NA

		mas um dos profs não queria gravação e optou-se por conversas TOTAL 3 Conversas: 5 a cada professor, Conversas com guião construído a partir de dados da 1º entrevista e as observações Perguntas abertas que eram registadas telegraficamente e mais fechadas, com registos mais precisos TOTAL 15	fazia-se registo escrito anexo para algo mais Directa, não participante e distanciada (na sala de aula) “na terminologia de Estrela (1990)” 6 aulas de cada prof TOTAL 18	das conversas e observações TOTAL 9		professores: de apoio às aulas assistidas, planificações e aulas assistidas	
1997	QUEIRÓS, M. Liseta	*com pilotagem a 2 profs, com questões abertas, individuais, áudio-gravadas, cerca de 1h, transcritas, semi-estruturadas, com guião; uma a cada professor=7	*construção de grelha e sua pilotagem a duas aulas de outros professores, não participante; 3 de LP, 3 de M = 12 aulas. (de 4 professores)	um a todos os professores=7*	*Artefactos: projecto educativo das escolas, actas do conselho pedagógico, actas dos grupos disciplinares envolvidos no estudo, planificações, instrumentos de avaliação, apontamentos dos professores, pautas, critérios de avaliação, planos de apoio educativo dos alunos, fichas de trabalho, relatórios dos alunos.	- Encontros informais	NM/NA
1997	RIBEIRO, Custódia	Semi-directivas; guião construído a partir da análise dos questionários; a 5	NM/NA	Com 21 questões abertas mas com um pequeno quadrado para a resposta, responderam	NM/NA	NM/NA	NM/NA

		professores que responderam ao questionário, cerca de 50 a 60', individuais e áudio-gravadas,		15 professores de 14 escolas públicas e uma de ensino militar que lecciona desde o 2º ciclo até ao superior. 15			
1997	SILVA, Catarina	Guião construído, semi-estruturada, gravadas, individuais, entre 15 a 25'. 12 a alunos escolhidos pelos seus professores	NM/NA	11 questões tipo Likert de 5 categorias, testado previamente a uma população semelhante, com 3 questões abertas e 3 fechadas, cerca de 20' para aplicação 147 alunos de todas as 6 turmas do 9º ano	NM/NA	NM/NA	NM/NA
1998	ALVES, Isabel	- 8 – duas a cada aluno- áudio-gravadas; - 15 a 40'; transcritas; -semi-estruturadas, individuais, com guião, mas não rígido.	- em sistema narrativo aberto, sem categorias pré-definidas. - não participante; - com anotações num caderno; -17 de Port=8+4+4. E 8 de Mat.=4+4 - vigilância das Provas Globais de P e de M	- 8 – 2 a cada aluno	- trabalhos dos alunos caso: trabalhos realizados nas disciplinas, testes, apontamentos diversos e cadernos diários.	- conversas informais às duas professoras, sem guião, sem registo áudio, praticamente diárias	SIM
1998	AMADO, Nélia	- 6 semi-estruturadas, áudio-gravadas e transcritas e revistas	- Cc 27 aulas	NM/NA	- oficiais; -fichas de trabalho, de av, grelhas de av e planificações	- Conversas informais; - Observação da correcção e classificação pelas profs, de alguns testes.	SIM
1998	FONSECA, Albina	NM/NA	NM/NA	- 133 construído, - perguntas abertas e escolha múltipla.	NM/NA	NM/NA	NM/NA
1998	LEITE, M. Luísa	- 1 semi-estruturada a cada participante com	NM/NA	NM/NA	NM/NA	NM/NA	NM/NA



		guião, gravadas e transcritas; - 1 exploratória a 1 prof; - 1 exploratória a 1 aluno - ambas para a construção do guião					
1998	PAIS, Paulo	- livres, de tipo exploratório; - formais, semi-estruturadas, gravadas e transcritas, de aprofundamento.	NM/NA	- 40: 1 fechado e outro aberto, tecnicamente validados por painel de especialistas e pré-testados. (o fechado inclui uma resposta aberta).	- grelhas de registo de dados usados pelos profs na audição do pensamento em voz alta.	- Audição e registo de pensamento em voz alta.	NM/NA
1998	RAFAEL, M. Amélia	- 6 ( 2*3) aos profs, - 15 aos alunos (a 5 alunos de cada professor) - semi-estruturadas, de longa duração, áudio-gravadas, dadas a ler ( no caso dos professores).	- 9 aulas (3+6).	NM/NA	NM/NA	- De apoio às aulas; - registo resultados alunos. - notas de conversas informais e das reuniões de grupo; produzidos pelos profs (fichas de trabalho, testes sumativos, fichas de registo das informações dos alunos) fornecidos pelo CE (pautas, elet <sup>os</sup> de caracterização da turmas)	NM/NA
1999	AFONSO, M. Rosa	- aos professores, uma a cada, 24 - semi-estruturadas, com protocolo, transcritas e com análise de conteúdo.	NM/NA	NM/NA	NM/NA	NM/NA	NM/NA
1999	BRAGA,	4 alunos (2 a cada).	houve, participante,	validado, pré-testado,	Sim	NM/NA	NM/NA

	David	Semi-estruturadas, com guião construído a partir de entrevistas informais, gravadas; a 4 alunos tipo	mas não refere o número	aplicado a 117 alunos em 97/98 e a 144 em 98/99; dados tratados em SPSS e respostas abertas com análise de conteúdo.			
1999	DIAS, Luís	NM/NA	NM/NA	- 20 pré-questionário; - 94 construído, maioritariamente fechado.	NM/NA	NM/NA	NM/NA
1999	MARTINS, M. da Graça	semi-estruturada, semi-directiva, aos Professores e aos alunos, áudio gravadas, transcritas protocolo analisado; 60' a 90'; TOTAL: 7= 3Profs +4 alunos	1ª fase não estruturada; 2ª fase: Estruturada, não participante, com guião de observação; TOTAL: 2= uma de cada professor	NM/NA	NM/NA	Instrumentos de avaliação utilizados e os privilegiados pelos actores, notas e apontamentos considerados importantes, relatórios de aproveitamento escolar dos alunos, testes executados pelos professores (incluindo PG), matrizes, grelhas de observação e planificações.	SIM
1999	TAVARES, Teresa	3 profes; semi-estruturadas, gravadas, análise de conteúdo	NM/NA	12 (1 foi eliminado sendo inicialmente 13); construído com base nas entrevistas; pré-testado a 4 profes; fechado; aplicado a todos os profes de Biologia do 12º do distrito de Beja.	Tabelas de classificação interna final e de exame de 95/96 e 96/97	NM/NA	NM/NA
2000	AFONSO, M. Deolinda	- estudo piloto - a 5 professores do	NM/NA	- 77 aos professores	NM/NA	NM/NA	NM/NA

		sec. De CFQ e TLQ e 2 escolas sec do Distrito do Porto - semi-estruturadas, individuais e gravadas					
2000	FERREIRA, Odette	- 6: 3 a cada prof, semi-estruturadas, 30 a 50', gravadas e transcritas.	- 12 aulas: 6 em cada turma de Port.	NM/NA	- textos produzidos pelas prof. Com comentários acerca da avaliação; - vários instrumentos de avaliação: testes, fichas de registo de observação, pautas com as classificações.	NM/NA	NM/NA
2000	GOUVEIA, Manuel	NM/NA	NM/NA	- 364 construído, fechado.	NM/NA	NM/NA	NM/NA
2000	PERES, Ana	3- Semi-diretiva, com guião, individuais, transcritas e revistas (1 a cada prof)	Com protocolo de observação, naturalistas (30 aulas entre 9/11 de cada professor	NM/NA	Fichas de avaliação, de trabalho, de registo de informação.	NM/NA	NM/NA
2000	SILVA, José	- 6: a 3 alunos de cada uma das 2 turmas; - 4: duas a cada prof..	NM/NA	- 34: a todos os alunos das duas turmas, maioritariamente fechado.	NM/NA	- reflexões individuais; - 3 momentos de avaliação formativa em grupo: 1º resultados sob a forma de relatório escrito; 2º sob a forma de relatório individual; 3º sob a forma de apresentação oral pelo grupo. - 1 av sumativa individual com construção de relatório.	- diário de bordo
2000	TORRES, M. Júlia	NM/NA	NM/NA	NM/NA	Fichas biográficas alunos; pautas; Dossiers de DT de	NM/NA	NM/NA

					cada Turma; Dossier do grupo de Mat; Livros de ponto; Registo biográfico dos professores.		
2001	DIOGO, Ana	- 20: 2 a cada prof, em fases diferentes, com guião pré-existente, cc 50', gravadas e transcritas, 2ª estruturada; 5: de grupos pequenos aos alunos, com guião, cc 1 tempo lectivo.	NM/NA	- 10: aos profs; - 25: aos alunos; - ambos de resposta fechada e condicionada; - o dos alunos foi adaptado de outro pré-existente.	- 11: um exemplar de cada teste sumativo aplicado na turma do 12ºnunciados de testes.	NM/NA	NM/NA
2001	RODRIGUES, Ana	NM/NA	NM/NA	NM/NA	Provas de Av de ano propedêutico:1; Provas Específicas: 5; Provas Escritas:3; Provas de Aferição:2; Provas de Exame Nacional:7; Provas 235 e equivalentes:11; total =29	NM/NA	NM/NA
2002	ARAÚJO, Filomena	semi-estruturadas, validadas, transcritas, individuais, aos alunos; TOTAL:113	NM/NA	Construído, de resposta aberta, um por professor; TOTAL:155	NM/NA	Conversas informais para construção dos instrumentos de recolha de informação	NM/NA
2002	SANTO, Gisela	NM/NA	NM/NA	- 391 (333 professores e 58 formadores)	NM/NA	NM/NA	NM/NA
2002	SANTOS, M. de Fátima	- 6: duas a cada, semi-estruturadas, com	NM/NA	NM/NA	- instrumentos de avaliação, instrumentos	- pensamento em voz alta; - instrumento pré-	NM/NA

		guião, 60 a 75', áudio-gravadas (as 2ºs duraram cc de 1h, transcritas e dadas a ler).			de aprendizagem.	existente para registo da planificação diária da av. formativa das professoras.	
2003	FERREIRA, Cristina Leite	- 1 entrevista individual a dois professores que responderam ao questionário (2)	NM/NA	67 aos professores	NM/NA	NM/NA	NM/NA
2003	ROCHA, Carla	- informais para construção do questionário	NM/NA	- construído a partir de entrevistas informais; pré-teste - de resposta fechada e aberta - 110 distribuídos, 6 eliminados TOTAL: 104	NM/NA	NM/NA	NM/NA
2003	SANTOS, M.	10 individuais a profs de 1 escola, com protocolos (entrevistas piloto para construir o questionário)	NM/NA	Pré-questionário a 18 profs de 3 escolas 114 profes com 85 respondentes de 17 escolas	NM/NA	NM/NA	NM/NA