

## Contornos de uma Experiência de Avaliação Desenvolvida no Instituto de Inovação Educacional (1990-1993)

Domingos Fernandes\*

### *Resumo*

A avaliação é uma componente imprescindível de qualquer processo de reforma porque permite identificar discrepâncias entre o que se concebeu e o que, de facto, acontece; contribui para que se «produzam» visões integradas dos elementos do sistema; ajuda a compreender os «pontos de vista» dos intervenientes e a «tomar o pulso» ao que acontece nos mais diversos níveis; facilita a correcção do que foi identificado como inadequado; ou permite a identificação e a divulgação de inovações.

Entre 1990/1991 e 1992/1993 o autor concebeu e coordenou, no então Departamento de Avaliação Pedagógica do Instituto de Inovação Educacional, um conjunto de trabalhos de investigação cuja principal finalidade era a de avaliar/acompanhar o processo de experimentação dos novos programas no âmbito da reforma do sistema educativo.

Neste trabalho discutem-se contornos dessa experiência em quatro aspectos essenciais: princípios observados; principais finalidades; estudos desenvolvidos; e opções metodológicas utilizadas.

O sistema educativo português atravessa, desde 1986, um período em que praticamente todos os seus elementos ou componentes já experimentaram, ou estão a experimentar, algum tipo de modificação. Professores, alunos, escolas, administração, pais e sociedade em geral têm

---

\* Universidade de Aveiro.

Boletim SPEF, n.º 10/11 Verão/Outono de 1994, pp. 7-32.

vivido, de forma mais ou menos interveniente, os avanços e recuos de uma reforma multifacetada, de contornos complexos, que pode constituir uma real oportunidade de melhoria geral do sistema.

A avaliação, não sendo de modo algum a panaceia para os «males» que nos possam atormentar, é uma componente imprescindível de qualquer processo de reforma pois, entre outras coisas:

- a) permite analisar discrepâncias entre o que gostaríamos que acontecesse e o que, de facto, acontece;
- b) contribui para que se «produzam» visões integradas dos diversos elementos do sistema;
- c) ajuda a compreender melhor os «pontos de vista» dos diversos intervenientes e a «tomar o pulso» ao que vai acontecendo nos mais diversos níveis;
- d) facilita a correcção daquilo que foi identificado como inadequado;
- e) permite a identificação, caracterização e a divulgação de inovações.

Em última análise, a avaliação dá um contributo inestimável ao desenvolvimento de *cultura educacional* porque permite conhecer em profundidade a natureza dos discursos dos diferentes intervenientes (por exemplo, professores, alunos, administração); descreve com detalhe os «ambientes» e os contextos em que decorrem o ensino e a aprendizagem; ou porque contribui para a criação de memória acerca das «coisas» da educação.

Talvez por estas razões o Ministério da Educação tenha decidido apoiar, entre 1990/1991 e 1992/1993, através do Instituto de Inovação Educacional (IIE), a concepção, o desenvolvimento e a divulgação de estudos de avaliação<sup>1</sup> acerca dos processos de lançamento da chamada Reforma do Sistema Educativo.

Neste trabalho relatam-se e discutem-se os principais contornos daquela experiência de avaliação pela qual fui responsável no então Departamento de Avaliação Pedagógica (DAP)<sup>2</sup> do IIE e que se encontra descrita mais detalhadamente em Fernandes (1992).

Em primeiro lugar fazem-se considerações referentes às dificuldades associadas ao desenvolvimento do trabalho de avaliação em geral e dão-se as principais características das seis abordagens mais utilizadas. Numa segunda parte discutem-se os princípios que orientaram a concepção e o desenvolvimento do trabalho do DAP. Em terceiro lugar descrevem-se as principais finalidades da avaliação e propõe-se um «modelo» que oriente o trabalho de avaliação no futuro. Finalmente, apresenta-se o espectro de estudos desenvolvidos, ou que se previam desenvolver de acordo com as finalidades previamente enunciadas e indicam-se as audiências preferenciais que então se julgava ser necessário servir.

## *Dificuldades Iniciais*

Estabelecer uma estrutura de avaliação com a responsabilidade de identificar e caracterizar os pontos fracos, os pontos fortes e as dinâmicas de um processo de reforma é, em si mesma, uma tarefa de grande complexidade que é substancialmente agravada por factores tais como:

- a) a ausência de uma teoria consistente;
- b) a ausência de qualquer «escola» ou tradição de avaliação no nosso país;
- c) as expectativas e exigências da administração, dos professores e da comunidade educativa em geral.

Quanto à ausência de uma teoria consistente, Madaus, Stufflebeam & Scriven (1991) consideram que não nos podemos referir a uma «teoria da avaliação» ou mesmo a «modelos de avaliação». Na opinião destes autores, as abordagens que têm sido utilizadas são ainda demasiado gerais e demasiado vagas e, conseqüentemente, não são suficientemente consistentes para nos permitirem interpretações e explicações suficientemente sólidas de fenómenos educacionais complexos. Pode afirmar-se que estas abordagens vão desde as de tipo utilitário até às chamadas abordagens intuicionistas e pluralistas. De facto, a literatura da especialidade (e. g., Cronbach, 1980, 1982; Eisner, 1979; Guba & Lincoln, 1981; House, 1991; Levine, 1974; Popham, 1975; Scriven, 1991; Stake, 1978; Stufflebeam & Webster, 1980; Tyler, 1967, 1991) permite identificar seis abordagens principais em avaliação educacional:

- a) Avaliação orientada por objectivos;
- b) Avaliação orientada para a decisão;
- c) Avaliação orientada para os consumidores;
- d) Avaliação desenvolvida por painéis de especialistas;
- e) Avaliação por exploração de contradições;
- f) Avaliação naturalista e orientada para os participantes.

Na Figura 1 sintetizam-se as características principais de cada uma destas abordagens.

Pode assim dizer-se que existem disponíveis um conjunto de categorias, de questões para reflexão, de descrições e procedimentos de diferentes abordagens de avaliação educacional mas não propriamente uma teoria de avaliação.

A segunda dificuldade, ausência de uma tradição de avaliação educacional no nosso país, impedia que no DAP tivéssemos acesso a experiências ou a conhecimentos previamente desenvolvidos. De facto, não existiam estudos de avaliação do sistema ou de programas que se integrassem num plano organizado e coerente e que reflectissem a realidade

Avaliação por Objectivos (Ralph Tyler)	Avaliar o grau de consecução dos objectivos; utilização de instrumentos objectivos; relevo à fiabilidade e validade; fácil de utilizar; foco nos produtos; visão simplista de educação.
Avaliação para a Decisão (Daniel Stufflebeam)	Obter informação útil para a tomada de decisões; avaliação de necessidades de informação dos decisores; visão sistémica da avaliação; modelo bem operacionalizado; ênfase no modelo produtivo das organizações; racionalista.
Avaliação para os Consumidores (Michael Scriven)	Obter informação útil acerca de produtos educacionais; utilização de listas de verificação; definição exhaustiva de critérios; avaliação formativa e sumativa; influência no desenvolvimento de produtos; pode limitar a criatividade e a inovação.
Avaliação Baseada no Conhecimento e na Crítica Educacional (Elliot Eisner)	Obter apreciações de grande qualidade, com base nos conhecimentos e experiência de especialistas; utilização da crítica educacional; exige poucos meios de realização; espectro de análise alargado; depende das qualificações do especialista.
Avaliação por Confronto de Opiniões (Murray Levine)	Integrar diferentes ideias e opiniões; inspirada nos métodos forenses; análise dos pontos fortes e fracos; apresentação de múltiplas perspectivas; orientada para a resolução efectiva de problemas; custos elevados; demorada.
Avaliação Naturalista (E. Guba e Y. Lincoln)	Compreender as complexidades e os contextos das situações; reflecte múltiplas realidades; utilização do raciocínio indutivo; foco na descrição, interpretação e compreensão; tendência para focar no que é discrepante; custos elevados.

Figura 1 — Estratégias de avaliação: proponentes e características principais

pedagógica portuguesa. No entanto, puderam referenciar-se alguns estudos de avaliação dos currículos do ensino unificado e da formação em serviço (GEP, 1979, 1980a, 1980b, 1981a, 1981b, 1991; Loureiro, Matias, Veiga, Pardal, Silveira & Boavida, 1986).

Finalmente, sabe-se que a comunidade educativa (e. g., administração, professores, gestores escolares) tem grandes expectativas relativamente à avaliação. Por um lado, tende a ver a avaliação como uma solução ou uma panaceia para os males que afectam o sistema educativo e, por outro, tende a considerar que é sempre apropriado avaliar e que todo e qualquer aspecto de qualquer acontecimento ou processo educativo tem que ser sujeito a avaliação. Esta tentação, que se pode considerar interessante ou talvez ideal, ignora muitas realidades práticas. Decidir o que avaliar e encontrar os meios para o fazer são problemas não rotineiros cujas soluções dependem do conhecimento e da experiência dos avaliadores, do empenho da administração, e de todos os que de algum modo estão envolvidos num processo de mudança, como o que uma reforma pode induzir no sistema.

## *Princípios de Avaliação*

Sem termos propriamente uma tradição de avaliação em Portugal, foi a partir de uma revisão de literatura de diferentes perspectivas de avaliação e dos conhecimentos e experiência das diferentes equipas de investigação que se adoptaram três princípios orientadores:

- a) O princípio do ecletismo das abordagens;
- b) O princípio da relevância dos pensamentos e conhecimentos dos professores;
- c) O princípio da diversificação dos avaliadores.

### *Ecletismo das Abordagens*

Cronbach (1982), Greene, Caracelli e Graham (1989), Madey (1982), Talmage (1982) e Worthen (1977, 1981), perante a diversidade de abordagens à avaliação de programas, defenderam a necessidade de se dedicar mais atenção ao desenvolvimento de avaliações com características ecléticas. Segundo aqueles investigadores, a possibilidade de combinar selectivamente aspectos de diferentes abordagens permite um conhecimento mais profundo e detalhado do objecto de avaliação. No entanto, esta posição é contestada, entre outros, por Guba e Lincoln (1981, 1989, 1991), por Lincoln e Guba (1985) e por Stake (1978) que consideram a impossibilidade de qualquer compatibilidade entre abordagens com pressupostos filosóficos e epistemológicos antagónicos.

Aquele antagonismo tem-se reflectido no debate existente em torno da objectividade ou subjectividade da avaliação ou ainda em torno do seu carácter quantitativo ou qualitativo (Howe, 1985; Maxwell e Lincoln, 1990; Page e Stake, 1970). É um debate relativamente pouco frutuoso que tem contribuído para a polarização dos avaliadores e investigadores (Madaus, Scriven e Stufflebeam, 1991).

No caso das abordagens objectivas, a avaliação é considerada «cientificamente objectiva». Esta objectividade obtém-se através da utilização de instrumentos, tais como questionários e testes, construídos com elevado rigor técnico e cujos resultados são analisados com base em técnicas de natureza exclusivamente quantitativa. Um dos pontos chave desta abordagem tem a ver com a possibilidade dos estudos poderem ser replicados. Para tal, o problema, as questões, as opções metodológicas, os instrumentos e os procedimentos de recolha e análise de dados são descritos exaustivamente.

Alguns dos principais pressupostos epistemológicos das abordagens objectivas são o da uniformidade da natureza, o da possibilidade de se formularem leis através do método indutivo e o da semelhança metodológica entre as ciências naturais e sociais, ressaltando-se o facto de, estas

últimas, serem de natureza mais complexa. Desta forma, os factos observáveis podem ser «facilmente» agrupados em categorias tal como o que se passa nas chamadas ciências experimentais. Torna-se possível, através de técnicas apropriadas, a definição de variáveis e o estabelecimento de relações entre elas e, uma vez assumido o princípio da uniformidade da natureza, podem induzir-se leis para as ciências sociais tal como para as ciências da natureza.

Assim, na sua forma mais radical, os estudos objectivos excluem tudo o que não é quantitativo; a tradução, em números, dos factos observados é sinónimo de objectividade. O principal indicador desta objectividade é o chamado acordo intersubjectivo entre observadores ou avaliadores. Note-se que a intersubjectividade implica que o trabalho de avaliação desenvolvido por uns possa ser experimentado, da mesma forma, por outros. Consequentemente, a objectividade é obtida através da explicitação exaustiva de todas as referências do processo de avaliação para permitir a replicação. É por isto mesmo que a fiabilidade é um conceito tão relevante em avaliação objectiva. Quanto mais fiáveis forem os instrumentos e técnicas utilizados, mais exactamente se podem determinar eventuais diferenças entre observadores independentes e mais objectiva se torna a avaliação. Logo, nesta perspectiva, o conceito de objectividade está directamente relacionado com os procedimentos que permitem determinar a intersubjectividade de apreciação dos avaliadores.

Nas chamadas abordagens subjectivas, tais como a naturalista, a natureza dos procedimentos utilizados é mais geral, não existindo qualquer preocupação em explicitar todos os seus pormenores para que outros possam observar exactamente as mesmas coisas. As leis ou proposições mais gerais são formuladas pelos avaliadores como resultado das experiências e conhecimentos que vão adquirindo no estudo de casos devidamente contextualizados. No entanto, o seu principal objectivo não é a formulação de proposições generalizáveis; existe sempre a preocupação em obter asserções válidas tendo em conta o quadro de referência dos participantes. Assim, o que é verdade num dado contexto, pode não o ser noutro. Por isso, as interpretações dos participantes envolvidos num dado programa ou actividade são mais valorizadas do que as dos avaliadores pois estes poderão não ter um conhecimento aprofundado do que se está a passar. Procuram assim obter-se múltiplas perspectivas que, naturalmente, não serão coincidentes.

Embora existam diferenças fundamentais nestas duas abordagens, é claro que ambas estão orientadas para o conhecimento rigoroso dos fenómenos ou factos envolvidos. Na verdade, o debate epistemológico não tem impedido o desenvolvimento de trabalhos de reconhecida qualidade em qualquer dos paradigmas.

Como acima já se referiu, a ideia dos benefícios obtidos através da integração selectiva de métodos ou técnicas de diferentes abordagens,

nomeadamente de tipo qualitativo e quantitativo, tem sido referida e posta em prática por diversos investigadores (Greene, Caracelli e Graham, 1989; Madey, 1982; Smith, 1980; Worthen, 1981). Tal integração é apontada como vantajosa nas fases do desenho da investigação e nas de recolha e análise de dados, particularmente quando se trata de avaliar programas de grande complexidade.

No Departamento de Avaliação Pedagógica do Instituto de Inovação Educacional foi sempre clara a existência daquele debate epistemológico e foi sempre assumida a necessidade de aprofundar o estudo e a reflexão em torno dos seus principais contornos.

Dada a dinâmica e a natureza do processo de experimentação dos novos programas, sentiu-se que restringir a avaliação a métodos quantitativos (ditos «objectivos») limitaria o significado dos resultados obtidos. Por outro lado, reduzir a avaliação a métodos exclusivamente interpretativos (ditos «subjectivos») poderia ter, entre outros, o problema da dificuldade em obter conclusões aplicáveis às decisões de política geral de educação. Uma abordagem eclética, tanto quanto possível integrada, pareceu ser a mais adequada tendo em conta os objectivos que se pretendiam alcançar e o contexto em que a avaliação iria decorrer.

Tal opção foi tomada e aceite no pressuposto de que a escolha de uma dada estratégia de avaliação deve depender mais do seu objectivo concreto e das suas potencialidades para gerar a informação que é necessária e muito menos de ideias pré-concebidas acerca do que são abordagens legítimas de avaliação.

### *Relevância dos Pensamentos e Conhecimentos dos Professores*

Um dos dilemas de qualquer movimento reformista que vise a introdução de mudanças significativas na estrutura e funcionamento do sistema foi formulado por Max Weber da seguinte forma: *Como é que, simultaneamente, se poderá controlar a maximização da eficiência e eficácia dos participantes minimizando, simultaneamente, a sensação de frustração e antipatia que esse mesmo controlo lhes provoca?*

No momento que actualmente se vive em Portugal, aquele dilema parece ajustar-se perfeitamente. De facto, pretende-se que os professores melhorem o funcionamento das escolas, ensinem melhor, promovam a qualidade do sistema ou utilizem novas formas de avaliação e novas metodologias. Estas exigências, induzidas pelos novos objectivos e pelos imperativos económico-sociais, são, obviamente, fontes de discrepâncias de entendimento face a uma grande diversidade de questões e, conseqüentemente, fontes de conflitos. Na verdade, o desenvolvimento da reforma impõe, por um lado, um certo grau de centralização e de normalização e, por outro, o reforço e a generalização da autonomia das escolas. Debate-se entre o problema da gestão do sistema por controlo

hierárquico e a necessidade de dar mais responsabilidade e condições de desenvolvimento profissional aos gestores e professores; entre a introdução de mudanças por despacho ou decreto, medidas meramente prescritivas, e a necessidade de criar condições para que as mudanças sejam o resultado do trabalho empenhado das escolas e professores.

Lidar com estes e outros dilemas faz parte dos esforços que é necessário empreender para melhorar as escolas e o próprio sistema. Como é óbvio, os professores são elementos centrais na resolução destes problemas. Por isso mesmo, conhecer as suas perspectivas sobre a forma como a experimentação dos novos programas está a decorrer, sobre a Área-Escola ou sobre as suas necessidades de formação é, seguramente, uma forma de compreender em que medida é que as mudanças de política educacional estão a afectar o seu empenhamento. E é sabido que o seu grau de empenhamento nessas mudanças depende, em boa medida, das percepções que têm sobre a forma como todo o processo se desenvolve.

Por outro lado, partiu-se do pressuposto de que é nas salas de aula e nas escolas reais, que está o cerne do processo de ensino e aprendizagem. Os professores, obviamente, têm um papel fundamental no desenvolvimento deste processo e, por isso mesmo, muito do que é preciso mudar e melhorar passará necessariamente pelas suas atitudes e pelas suas tomadas de decisão. Ou seja, nada, ou quase nada, no processo de desenvolvimento da reforma se poderá concretizar efectivamente sem os professores. Consequentemente, conhecer a forma como eles percebem o que está a acontecer nas salas de aula é uma forma de identificar e reforçar o que possa estar bem e procurar corrigir o que possa estar mal.

Vários investigadores (e. g., Clark e Peterson, 1986; Schremer, 1991; Shavelson e Stern, 1981) têm referido o papel chave desempenhado pelos professores nos processos de mudança e a importância que, em termos de avaliação, pode ter a análise dos seus pensamentos e conhecimentos.

Nestas condições, decidiu-se que as opiniões dos professores eram, nesta fase, um elemento crucial para conhecer o processo de experimentação dos novos programas e, como tal, deveriam ser um dos principais elementos a considerar nos estudos de avaliação a efectuar.

### *Diversificação dos Avaliadores*

Um dos problemas com que uma equipa de avaliadores se confronta é o da credibilidade do seu trabalho de avaliação. Em Portugal, tal problema é substancialmente agravado por duas ordens de razões.

Por um lado, parte importante da comunidade educativa, por motivos de natureza diversa, tende a descrever do trabalho realizado no



âmbito de instituições e organismos do Ministério da Educação e a desvalorizar a sua relevância e qualidade técnico-científica.

Por outro lado, a falta de hábito em confrontar conhecimentos baseados na experiência pessoal, ou na tradição, com outros resultantes de investigação, induz desconfiança e ainda provoca reacções negativas. Este facto prende-se, naturalmente, com a nossa falta de tradição no desenvolvimento de trabalhos de avaliação.

Uma das formas de contribuir para tornar mais credíveis os trabalhos de avaliação é a de envolver avaliadores externos diversificando, deste modo, as perspectivas utilizadas. Pensou-se que esta diversificação poderia ter as seguintes vantagens:

1. Aumentaria a credibilidade da avaliação tomada como um todo pois os avaliadores externos, não fazendo parte da administração, são considerados menos sujeitos a possíveis influências ou interesses relacionados com o sucesso ou insucesso dos objectos da avaliação.
2. Aumentaria o espectro dos objectos de avaliação, colmatando limitações físicas, técnicas e científicas dos avaliadores do DAP e contribuindo para desenvolver o campo da avaliação em Portugal.
3. Permitiria o aparecimento de estudos que incluíssem perspectivas diferentes e optassem por outras abordagens conducentes à identificação de outras questões.
4. Aumentaria a possibilidade de se obterem dados com mais qualidade pois sabe-se que os participantes numa dada experiência manifestam, por vezes, mais disponibilidade em fornecer informações relevantes a avaliadores externos (neste caso ao «Ministério») do que a avaliadores internos.
5. Contribuiria para o estabelecimento, junto das escolas superiores de educação e das universidades, de observatórios do sistema educativo. Este objectivo é relevante já que, como se referiu, o trabalho de investigação e de produção teórica na área da avaliação de sistemas e programas é praticamente inexistente nas universidades e escolas superiores portuguesas.

Por estas razões, e ainda pela vantagem em promover o intercâmbio com instituições do ensino superior, foi decidido adoptar o princípio da diversidade.

Foi na sequência deste princípio que se apoiaram financeiramente investigadores da Universidade de Lisboa (Ponte, Matos, Guimarães, Leal e Canavarro, 1991; Matos, Ponte, Guimarães e Leal, 1993; Canário, Barroso, Oliveira e Pessoa, 1994), da Universidade do Minho (Serafini, Magalhães e Castro, 1991; Castro, Afonso, Pacheco e Magalhães, 1993) e da Universidade de Aveiro (Arroteia, Pardal e Martins, no prelo) para desenvolverem estudos de avaliação que se integrassem nas finalidades delineadas pelo DAP do IIE.

## *Finalidades da Avaliação*

A avaliação do sistema educativo e, muito particularmente, da reforma educativa consiste, no fundo, em conceber, desenvolver e aplicar um conjunto de procedimentos que permitam analisar criticamente uma política educativa e, muito particularmente, as suas relações com as estruturas que o constituem e com as pessoas que nelas trabalham. Este processo, complexo e multidimensional, exige que se descrevam as formas como aquela política se traduz no terreno, identificando aquilo que funciona como previsto e aquilo que não funciona, e se interpretem a grande variedade de situações e contextos em que se desenvolvem todo o tipo de processos educativos. Por exemplo, em relação ao currículo, estudam-se diferenças entre o currículo proposto pela administração ou por um conjunto de especialistas, o currículo efectivamente posto em prática e ainda o currículo que «passa» para os alunos, ou seja, aquilo que de facto eles aprendem. Ao mesmo tempo, analisam-se os contextos que acompanham o processo de aprendizagem (e. g., condições de funcionamento, formação de professores, materiais de apoio) e estudam-se os mecanismos que se desenvolvem (e. g., dinâmicas de desenvolvimento, processos de comunicação, soluções encontradas, níveis de participação dos intervenientes).

As consequências de qualquer processo de avaliação devem ser a despistagem, ou correcção, do que está mal e o reforço do que está bem. Trata-se de um processo que visa, acima de tudo, informar para promover a reflexão e a melhoria. Para tal, torna-se necessário mobilizar a administração, as escolas, os professores, os alunos e os respectivos encarregados de educação.

A experiência e os conhecimentos adquiridos nos primeiros dois anos (1990 e 1991) pelos avaliadores do Instituto de Inovação Educacional, incluindo, naturalmente, as limitações que reconheciam nos estudos realizados, permitiram, em 1992, identificar as vertentes segundo as quais os estudos de avaliação dos anos seguintes se deviam desenvolver, assim como as orientações metodológicas a que deveriam obedecer. Tratou-se do que então se considerou ser um contributo para o desenvolvimento de um *sistema integrado e global de avaliação do sistema educativo* e, em particular, tendo em conta o contexto de então, da chamada reforma educativa.

Por uma questão de simplicidade e clareza os objectivos gerais desse sistema aparecem aqui sob a forma de questões organizadas em cinco grandes áreas:

- a) Integridade e coerência das medidas propostas;
- b) Adequação e exequibilidade das medidas propostas;
- c) Reacção dos intervenientes às medidas propostas;
- d) Resultados da aprendizagem;
- e) Participação dos intervenientes.

## *Integridade e coerência das medidas propostas*

O sistema educativo está a experimentar um conjunto de mudanças supostamente conducentes à sua modernização e à melhoria da qualidade da educação. Para tal têm sido introduzidas medidas tendentes a adaptar o sistema às necessidades de desenvolvimento do país e às profundas alterações económicas, sociais, tecnológicas, culturais e políticas que este tem experimentado nos últimos 20 anos. Assim, definiram-se novos currículos para os ensinos básico e secundário, um novo sistema de gestão para as escolas, um novo sistema de avaliação dos alunos, criaram-se escolas básicas integradas de 9 anos e consignou-se a autonomia das escolas. Todas estas medidas estão explicitadas num edifício legislativo que, em princípio, se espera inteligente. Ou seja, é de esperar que haja consistência nas suas fundações e articulação e flexibilidade entre as suas partes para que os utilizadores e beneficiários se sintam seguros e possam percorrer o edifício sem problemas.

O papel da avaliação neste caso deve ser o de acompanhar e analisar o funcionamento do edifício e das suas partes à medida que este for iniciado nas funções para que foi concebido. De facto, só o seu funcionamento permitirá determinar com rigor os ajustamentos necessários para que as pessoas se sintam confortáveis e saibam exactamente como e onde se dirigir para resolver os problemas. Nestas condições, e neste caso, a avaliação deve dar respostas a questões tais como:

- Em que medida é que há coerência e consistência no conjunto de medidas propostas?
- Em que medida é que a formulação de conceitos essenciais é clara para todos os intervenientes (e. g., escola básica integrada de 9 anos, avaliação aferida, currículo, programa)?
- Será que todas as medidas propostas fazem sentido no terreno? Haverá falta ou excesso de especificação nessas mesmas medidas?
- Que comparações é possível fazer entre as medidas de política educativa, tal como elas surgem formalmente, e a sua aplicação em escolas e salas de aula reais?
- Há ou não discrepâncias entre as concepções de quem propõe o currículo e as novas medidas e as daqueles que, na prática, têm de lidar com os problemas da sua implementação? Qual a natureza dessas discrepâncias?

## *Adequação e exequibilidade das medidas propostas*

A reforma educativa pressupõe um conjunto de mudanças que vão desde a estrutura do próprio sistema até às metodologias utilizadas na sala de aula. Assim, torna-se necessário saber se tais mudanças são

compatíveis com as condições existentes e conhecer em que medida as pessoas directamente envolvidas no processo estão receptivas para as concretizar. Só assim se poderão introduzir medidas de remediação ou proceder a alterações mais ou menos profundas que eventualmente sejam necessárias.

Por exemplo, é importante conhecerem-se respostas para questões tais como:

- Até que ponto as escolas têm que alterar os seus procedimentos para porem em prática as mudanças propostas? Qual a natureza dessas alterações? Que capacidade têm para o fazer?
- Quais os principais constrangimentos ao desenvolvimento da reforma educativa? Será que os objectivos propostos são exequíveis?
- Quais os principais efeitos, não necessariamente previstos inicialmente, do desenvolvimento da reforma educativa nas escolas e nas salas de aula? Qual a sua natureza? Que medidas são necessárias para estimular efeitos considerados positivos e para despistar os que se considerarem negativos?

### *Reacção dos intervenientes às medidas propostas*

As reacções (e. g., percepções, atitudes, motivações, opiniões) das pessoas envolvidas no processo educativo e, em particular, dos professores e responsáveis das escolas, são elementos fundamentais para se proceder a uma avaliação responsável acerca do que está, de facto, a acontecer. Só conhecendo o sentido dessas reacções se poderão prever e planear com alguma segurança necessidades de formação, necessidades materiais, apoios a conceder ou concursos de projectos que ajudem a minimizar carências detectadas. Ou seja, os processos de decisão a vários níveis poderão ser substancialmente melhorados pois podem responder mais eficazmente às necessidades percebidas pelos participantes e, conseqüentemente, melhorar as condições objectivas e subjectivas conducentes à aprendizagem.

A avaliação das reacções dos intervenientes a uma política educativa e ao conjunto de medidas que lhe são inerentes deve basear-se numa descrição detalhada dos contextos e processos envolvidos para que melhor se possa interpretar o que está a acontecer.

Nestas condições, devem procurar-se respostas para questões tais como:

- Como é que as pessoas envolvidas (e. g., professores, directores de escolas, presidentes e elementos de conselhos directivos) estão

a reagir às mudanças propostas nos documentos da reforma? Que soluções têm encontrado? Que dificuldades mais os afectam?

- Quais as principais vantagens e desvantagens que diferentes intervenientes no processo educativo (e. g., professores, alunos, pais, autarcas) vêem nas medidas que a reforma pressupõe? Qual é a sua percepção das mudanças que são necessárias e das condições que existem para as pôr em prática?
- Qual a natureza dos processos que se desenvolvem e qual a influência que os diferentes contextos exercem sobre eles?
- Qual é o valor que os diversos intervenientes atribuem à reforma educativa?

### *Resultados da aprendizagem*

Saber o que os alunos estão, de facto, a aprender é talvez a responsabilidade dos avaliadores que mais directamente interessa a um maior número de audiências. Os pais querem saber se os seus filhos estão a aprender, como estão a aprender e o que estão a aprender. Os professores querem e precisam de saber se os seus conhecimentos, os seus métodos e, em geral, a forma como desenvolvem o currículo, estão devidamente ajustados aos seus estudantes. As escolas e, em particular, os seus gestores administrativos e pedagógicos querem e precisam de saber até que ponto os seus planos de formação e o ambiente pedagógico existente está a responder às necessidades dos seus alunos. Os especialistas em educação (e. g., os autores dos currículos) querem e precisam de saber se o currículo proposto está a ser apreendido por professores e alunos conforme o previsto, para poderem investigar e propôr eventuais remodelações ou correcções. A administração central precisa de conhecer o que os alunos estão de facto a aprender pois trata-se de um importante indicador do funcionamento do sistema. Mais particularmente, a administração precisa de saber se o que os alunos estão a aprender serve as finalidades delineadas para o desenvolvimento social, económico e cultural da sociedade portuguesa.

E a lista poderia prosseguir citando mais interessados nos resultados da aprendizagem dos alunos. Por isso, um sistema de avaliação não pode deixar de prever um conjunto de procedimentos (e. g., a aplicação de testes, observações formais e informais, entrevistas) destinados a obter dados referentes à aprendizagem. Tais procedimentos deverão constituir um dos contributos importantes para dar resposta a questões tais como:

- Tendo como referência o currículo proposto, o que é que os estudantes estão de facto a aprender através desse currículo? Que

conhecimentos, competências, capacidades, atitudes e valores estão a ser desenvolvidos?

- Quais as principais dificuldades no desenvolvimento do currículo? Quais os seus pontos fortes? Quais os seus pontos fracos?
- Haverá grandes discrepâncias no desempenho dos estudantes ao longo das diferentes regiões do país? Nas diferentes classes sociais? Nas diferentes escolas? Qual a natureza dessas discrepâncias?
- Até que ponto os instrumentos e procedimentos utilizados influenciam, positiva ou negativamente, o desenvolvimento do currículo nas escolas e, conseqüentemente, a aprendizagem?

### *Participação dos intervenientes*

A experiência de avaliação do IIE em 1990-1993 mostrou que, em muitos casos, e independentemente da intervenção directa da administração central, há dificuldades e problemas que são resolvidos localmente, nas escolas e nas salas de aula. Por vezes as soluções encontradas revestem-se de originalidade e de inovação, mobilizando recursos que nem sempre tinham sido previstos. Os exemplos que a este respeito são mais elucidativos surgem no âmbito da Área-Escola em que os professores, embora referindo não possuir condições materiais para desenvolver aquela área, acabaram por encontrar soluções que permitiram ultrapassar, ou pelo menos minimizar, aquela dificuldade.

Desta forma, a avaliação deverá ter em conta a participação dos intervenientes na procura de soluções para resolver problemas existentes nas escolas e que, aparentemente, poderão estar a ser induzidas pela reforma educativa. A caracterização dessas soluções poderá constituir uma referência importante para ajudar a concretizar os objectivos previstos. Assim, os estudos de avaliação a desenvolver deverão ter em conta questões tais como:

- Que tipo de soluções são encontradas pelas escolas e pelos professores para resolver os problemas com que se confrontam? Qual é a natureza desses mesmos problemas? As soluções são encontradas no seio da escola (e. g., Conselho Directivo, Conselho de Grupo, Conselho Pedagógico?) ou com ajuda externa (e. g., Outras escolas, especialistas, autarquias, empresas comerciais e industriais)?
- Até que ponto é que as soluções eventualmente encontradas nas escolas, e que por isso devem em si mesmas reflectir o que é desejável e/ou o que é possível, podem contribuir para fundamentar decisões por parte da administração?

- Em que medida se tem estimulado a participação dos diversos intervenientes na procura de soluções para os problemas existentes? Qual é a natureza dessa eventual participação?
- Que tipo de motivações e que tipo de expectativas são inerentes à procura de soluções para os problemas postos pelo desenvolvimento da reforma?

### *Síntese*

Em suma, de forma esquemática, enunciaram-se os grandes objectivos para a avaliação do processo de desenvolvimento da reforma educativa formulados em 1992. A formulação daqueles objectivos baseou-se no pressuposto de que é necessário emitir juízos acerca dum vasto conjunto de elementos ou componentes que integram o sistema (a reforma) tais como as medidas propostas, os currículos, as actividades pedagógicas que se desenvolvem, os contextos de realização, o grau de participação e as reacções dos principais intervenientes e os resultados da aprendizagem dos alunos.

Tradicionalmente, questões daquela importância têm sido apenas abordadas superficialmente. O trabalho que se desenvolveu teve como principal desafio estabelecer um plano global, tanto quanto possível integrado, de avaliação do sistema (da reforma) que, genericamente, identificasse:

- a) as necessidades de informação;
- b) as fontes onde obter essa informação;
- c) os contextos e processos a avaliar;
- d) os principais intervenientes;
- e) produtos a avaliar.

A Figura 2 organiza e sintetiza as ideias atrás explicitadas. Trata-se de uma forma esquemática de apresentar uma abordagem que se julgou dever ser posta em prática depois de dois anos de experiência na concepção e desenvolvimento de mais de uma dezena de estudos de avaliação.

### *Estudos de Avaliação*

Após a identificação das questões da avaliação a realizar, tornou-se necessário desenvolver um plano que permitisse encontrar as respectivas respostas e que fosse coerente com os princípios enunciados e com o quadro organizador dos objectivos, dos processos e produtos a avaliar. Tal plano baseou-se essencialmente na enumeração do tipo de estu-

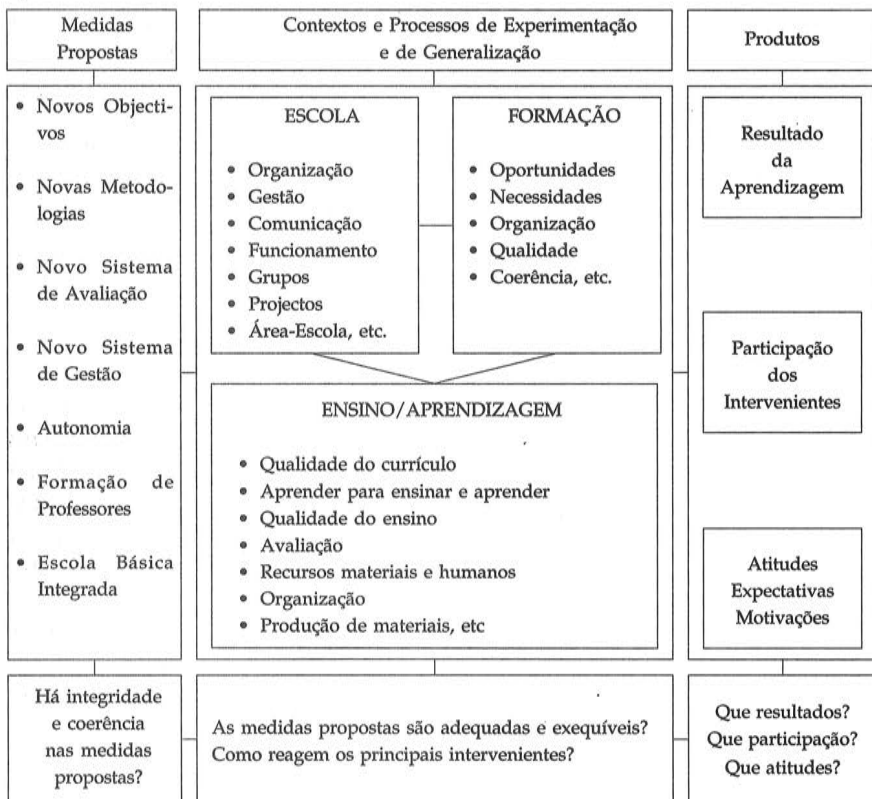


Figura 2 — Avaliação educacional no contexto da reforma educativa?

dos e de metodologias de recolha de informação que melhor pudessem contribuir para responder às questões de avaliação. Foram definidos os seguintes:

- a) Estudos de tipo sondagem;
- b) Estudos de caso;
- c) Estudos de avaliação da aprendizagem;
- d) Estudos de avaliação participada;
- e) Estudos baseados no conhecimento e crítica de especialistas.

### *Estudos de Tipo Sondagem*

Uma vez que um dos objectivos gerais da avaliação era o de conhecer a reacção geral dos participantes às medidas educativas em execução, os estudos de tipo sondagem constituíram uma base segura para a obten-



ção dessa informação. Na verdade, o estudo, ainda que intenso, de apenas alguns casos, não permite o conhecimento do que é mais relevante na maioria das escolas, a não ser que deliberadamente se aceitem margens de erro muito consideráveis.

Na realização destes estudos analisaram-se as opiniões emitidas por mais de 2000 professores experimentadores dos novos programas e por várias dezenas de conselhos pedagógicos, conselhos escolares, presidentes de conselhos directivos e directores de escolas. Desta forma acompanhou-se o processo de experimentação dos novos programas de todas as disciplinas do 1.º ao 9.º ano de escolaridade (Fernandes *et al* 1994a, 1994b; IIE, 1990, 1991a, 1991b, 1991c, 1992a, 1992b, 1992c) e dos 10.º e 11.º anos de escolaridade (Serafini, Magalhães e Castro, 1991 e Castro *et al*, 1993).

Foi também através de um estudo do tipo das sondagens desenvolvido por Fernandes *et al* (1991) que se investigaram as opiniões dos professores dos ensinos básico e secundário acerca das medidas propostas no primeiro projecto do novo sistema de avaliação dos alunos.

Os estudos de tipo sondagem, no entanto, não permitiram a obtenção de dados referentes à estrutura lógica de um processo de mudança como é o da experimentação dos novos programas; são em geral pouco interpretativos e iluminativos pois não fornecem dados substanciais acerca da forma como esses processos se desenvolvem no terreno. Por isso mesmo se decidiu que, em simultâneo com as sondagens, seria importante investigar as «particularidades» do processo de experimentação dos novos programas através de estudos de caso realizados em algumas escolas do país.

### *Estudos de Caso*

A realização de estudos de caso revelou-se imprescindível pois permitiu obter dados que doutra forma não teria sido possível obter. De facto, os estudos de caso têm potencialidades para relatar de forma exhaustiva o que se está, de facto, a passar no terreno, ao nível, por exemplo, das necessidades de formação, das necessidades de recursos, do desenvolvimento do currículo, da aprendizagem dos alunos, da motivação dos professores, das resistências aos processos de mudança ou da identificação de soluções e inovações.

Nestas condições, os estudos de caso tornaram-se um instrumento determinante na consecução dos objectivos preconizados no plano de avaliação que acima se sintetizou.

Ponte *et al* (1991) e Matos *et al* (1993) investigaram como decorreu o processo de experimentação dos novos programas de Matemática dos 7.º, 10.º e 11.º anos de escolaridade através de dois estudos de caso realizados em duas escolas de Lisboa e da Grande Lisboa.

A forma como decorreram quer a generalização do novo programa do 1.º ano de escolaridade numa escola de Lisboa (Fernandes et al, 1993; IIE, 1992e) quer o desenvolvimento e a aplicação da Área-Escola numa escola C + S do Alentejo (IIE, 1992d) também foram investigadas através de estudos qualitativos de caso.

Finalmente, foi ainda sob a forma de estudos de caso que Arroteia *et al* (no prelo) investigaram, numa escola secundária da cidade do Porto, dinâmicas suscitadas por medidas propostas na reforma do sistema educativo e que Canário *et al* (1994) analisaram a problemática do desenvolvimento de mediatecas escolares em quatro escolas do litoral-centro do país.

### *Estudos de Avaliação da Aprendizagem*

Avaliar o desenvolvimento das capacidades, competências, atitudes, valores e conhecimentos dos estudantes é uma componente fundamental da avaliação de qualquer programa. De facto, permite a obtenção de indicadores importantes relativos à forma como se desenvolve o currículo e ao que os estudantes aprendem desse mesmo currículo. Em última análise, o desempenho dos estudantes constitui um importante indicador da qualidade da educação que lhes é oferecida pelo sistema educativo.

Um dos dilemas mais discutidos em avaliação educacional é o da credibilidade e da necessidade da avaliação externa, normalmente desenvolvida pela administração central, face à existência da avaliação interna que se desenvolve nas escolas. Se, por um lado, os professores e as escolas estão em melhores condições do que ninguém para avaliar os seus alunos, o sistema sente dificuldade em interpretar os dados dessa mesma avaliação porque os critérios e os instrumentos em que os professores se baseiam, variam significativamente de escola para escola. Talvez por esta razão esteja prevista a chamada avaliação aferida como um dos instrumentos de avaliação do próprio sistema. Na verdade, a avaliação externa baseia-se num conjunto de critérios pré-definidos, normalmente com base nos objectivos previstos no currículo, para desenvolver instrumentos que avaliem o desempenho dos estudantes. Tais critérios e instrumentos, sendo os mesmos para todos os estudantes, permitem a uniformidade de que o sistema necessita para fundamentar as suas análises.

A avaliação externa pode também contribuir para o desenvolvimento, nas escolas, de metodologias mais adaptadas aos currículos propostos pela reforma e para sensibilizar os professores para o tipo de questões que tais metodologias devem sugerir. Resnick e Resnick (1992), por exemplo, consideram que a avaliação externa pode constituir um importante meio para desenvolver as capacidades de resolução de

problemas e os processos mais complexos de pensamento dos estudantes. Esta posição é baseada no pressuposto, geralmente aceite, de que a avaliação externa pode induzir a forma como se ensina.

Em suma, importa contribuir para que a avaliação externa (aferida) venha a constituir um instrumento de avaliação do currículo, da qualidade de ensino e da aprendizagem dos estudantes assim como um veículo que pode induzir, nas escolas, novas formas de ensinar e de avaliar. Por estas razões se torna necessário investir na investigação de novas formas de avaliar tal como as que nos são relatadas em Berk (1986), Biggs e Collis (1982), Gifford e O'Connor (1992), Jzn (1987), Kulm (1990), Kulm e Malcom (1991), Mitchell (1992), Nutall (1986), Romberg e Stewart (1987) e Webb (1992). Estes trabalhos deverão constituir um dos pontos de partida para o estabelecimento, ou consolidação, de linhas de investigação na área da avaliação da aprendizagem dos alunos que respondam a necessidades relacionadas com:

- a) o desenvolvimento da avaliação aferida;
- b) o desenvolvimento de instrumentos de diagnóstico da aprendizagem;
- c) a utilização de instrumentos de avaliação da aprendizagem nas salas de aula.

Foi tendo em conta tais preocupações que Oliveira et al (1993, 1994) desenvolveram instrumentos de avaliação da aprendizagem em Matemática para o 6.º ano de escolaridade e que Bentes (1993) e Cristo e Bentes (1993) investigaram formas de avaliar, e avaliaram, a aprendizagem em Língua Portuguesa no 1.º e no 2.º ciclos do ensino básico, respectivamente.

Na área da avaliação das aprendizagens deve ainda assinalar-se a produção de um estudo em que se compararam os sistemas de quatro países europeus (IIE 1992f) e de um trabalho de natureza teórica em que, essencialmente, se fundamenta a necessidade de mudar as formas de avaliar e de ensinar (IIE, 1992g).

### *Estudos de Avaliação Participada*

A experiência tem demonstrado que o envolvimento mais profundo dos participantes (e.g., professores, elementos da gestão das escolas) no desenvolvimento dos estudos de avaliação é vantajosa por várias razões:

- 1) Facilita o acesso a fontes primárias de informação que doutro modo nem sempre estão acessíveis;
- 2) Ajuda a clarificar e a precisar os objectivos da avaliação;

- 3) Proporciona uma visão mais diversificada das questões em análise;
- 4) Tem um elevado poder mobilizador de meios humanos e materiais que interessam ao processo de avaliação;
- 5) Aumenta a credibilidade da avaliação junto dos professores e das escolas;
- 6) Facilita a geração de soluções para resolver os problemas encontrados.

Tendo também em conta que, nos próximos anos, se vão iniciar os processos de experimentação de medidas tais como o novo sistema de gestão e da autonomia das escolas, parece ser relevante a integração de professores e elementos da gestão das escolas nas equipas de avaliação. De facto, este tipo de procedimento poderá contribuir para aumentar o empenho das escolas e dos professores na criação de melhores condições e na procura de soluções para ultrapassar os problemas. Ao mesmo tempo, pode reduzir a intervenção dos organismos centrais ou regionais como protagonistas de soluções e indutores de mudanças que, muitas vezes, não têm condições para serem postas em prática.

No âmbito destes estudos, que essencialmente seriam de natureza qualitativa, deveriam estar envolvidas duas a três escolas em que decorresse o processo de experimentação de uma ou mais das novas medidas. Os avaliadores iriam confrontando os participantes com os resultados das suas observações e análises e fomentariam a discussão orientada para a identificação das dificuldades, procura de soluções e o desenvolvimento da inovação. Trata-se de um «modelo» que se inspira nas abordagens de tipo naturalista e pluralista e que poderá fornecer informação útil para a tomada de decisões.

### *Estudos Baseados no Conhecimento e na Crítica de Especialistas*

O julgamento de um programa feito por especialistas é talvez, como já foi dito atrás, um dos métodos mais antigos de avaliação e talvez um dos mais utilizados. É muito comum, por parte da administração, solicitar pareceres a reputados especialistas sobre aspectos particulares de um programa ou sobre os possíveis efeitos práticos de uma dada medida de política educativa. Outra prática comum entre nós é a designação de grupos de trabalho para estudar determinado assunto. Normalmente é escrito um relatório ou feito um relato oral em que se expõem os aspectos essenciais para responder às necessidades da administração.

Este método tem riscos relacionados com a dependência exclusiva nos presumíveis conhecimentos e experiência dos «juízes» e com a forte possibilidade de existência de conflitos de interesses, mas tem a vantagem de, com base nas capacidades de julgamento e apreciação humanas,

tal como o que sucede nos meios forenses, produzir avaliação com meios relativamente escassos. Por outro lado, utilizados com outros estudos, as opiniões e julgamentos de especialistas podem trazer mais profundidade, esclarecimento e enquadramento teórico à avaliação.

A avaliação a desenvolver nos próximos anos deveria contar com este tipo de contribuição, muito especialmente no que respeita à análise dos objectivos relacionados com a integridade e coerência do conjunto de novas medidas que estão a ser introduzidas.

A contribuição de painéis de especialistas que se reuniriam por um período limitado de tempo (um a dois dias) para analisar aspectos previamente identificados e que se obrigariam a produzir um documento escrito de apreciação, é uma das formas que se poderá utilizar. Outra poderá ser a consulta directa, individual, a especialistas a quem se podem solicitar opiniões e/ou pareceres acerca do processo de desenvolvimento de um dado aspecto do currículo ou de uma medida proposta pela administração.

Promover o confronto formal de opiniões acerca de aspectos mais críticos do processo educativo, é também uma metodologia que poderá contribuir para uma avaliação mais credível, útil e responsável.

Na Figura 3 estão organizadas as finalidades da avaliação para os próximos anos, o tipo de estudos a desenvolver e as audiências que, preferencialmente, os seus resultados devem servir.

### *Nota Final*

George Pólya, educador e matemático eminente, dizia que se aprende a resolver problemas, resolvendo problemas. Em 1990, quando se iniciou o trabalho de avaliação que aqui se descreveu, afirmava-se, parafraseando aquele matemático, que se aprenderia a avaliar, avaliando... Se é certo que muito há ainda a aprender, não é menos certo que o trabalho que se desenvolveu nos ensinou pelo menos uma coisa: é possível avaliar e contribuir para que a avaliação faça parte integrante da cultura educacional portuguesa.

Tal como no início deste trabalho referi, avaliar é, em si mesmo, um processo complexo, que exige uma grande diversidade de saberes e de competências e um trabalho de investigação sistemático e coerente. A experiência demonstrou claramente que havia ainda muitos saberes e muitas competências a desenvolver. No entanto, creio que o trabalho realizado no Departamento de Avaliação Pedagógica do Instituto de Inovação Educacional em 1990-1993 constitui uma base para que, sem improvisações e sem devaneios, se possa prosseguir e, naturalmente, melhorar o que já se consegui fazer. Na verdade, após toda a experiência que aqui sinteticamente descrevi, estou convencido de que a continuação do trabalho na área da avaliação educacional é uma exigência real da melhoria da qualidade do sistema educativo português. Só por

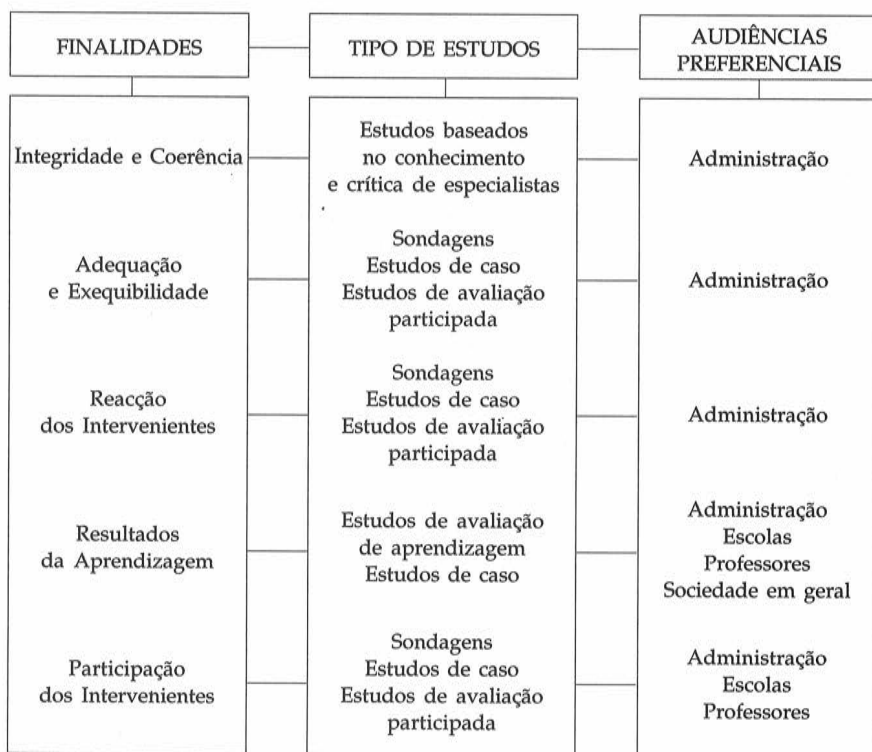


Figura 3 — Finalidades da avaliação, estudos a desenvolver e audiências preferenciais

razões estranhas aos interesses da educação ou por pura irresponsabilidade se poderia ignorar tal exigência.

### Notas

<sup>1</sup> Todos esses trabalhos foram concebidos e integralmente realizados naquele período. No entanto, muitos deles só muito recentemente foram publicados como é, por exemplo, o caso de Bentes (1993), Cristo e Bentes (1993), Castro, Afonso, Pacheco e Magalhães (1993), Fernandes, Góis e Cortez (1993), Fernandes, Assunção, Faria, Gil e Mesquita (1994a, 1994b), Matos, Ponte, Guimarães e Leal (1993), Oliveira, Pereira e Fernandes (1993, 1994). Outros trabalhos desse período ainda estão no prelo.

<sup>2</sup> Hoje trata-se da Direcção de Serviços de Avaliação do Sistema Educativo (DASE).

### Referências

ARROTEIA, J. C.; PARDAL, L. A., e MARTINS, A. M. (no prelo). *Análise da implementação e do impacto dos novos programas na Escola Secundária Filipa de Vilhena*. Lisboa: IIE.

- BENTES, M. C. (1993). *Avaliação da comunicação escrita em alunos do 1.º ciclo: Um estudo longitudinal*. Lisboa: IIE.
- BERK, R. (1986). *Performance assessment: Methods and applications*. Baltimore: John Hopkins University Press.
- BIGGS, J. e COLLIS, K. (1982). *Evaluating the quality of learning: The SOLO taxonomy (Structure of the Observed Learning Outcome)*. New York: Academic Press.
- CANÁRIO, R.; BARROSO, C.; OLIVEIRA, F. e PESSOA, A. (1994). *Mediatecas escolares: Génese e desenvolvimento de uma inovação*. Lisboa: IIE.
- CASTRO, R.; AFONSO, A.; PACHECO, J. e MAGALHÃES (1993). *Conteúdos e contextos da reforma curricular no 11.º ano de escolaridade*. Lisboa: IIE.
- CLARK, C. e PETERSON, P. (1986). Teachers' thought processes. In M. Wittrock (Ed.), *Handbook of Research on Teaching* (3rd ed.), (pp. 253-314). New York: MacMillan.
- CRISTO, F. e BENTES, M. C. (1993). *Avaliação da aprendizagem em Língua Portuguesa no contexto dos novos programas: 6.º ano de escolaridade*. Lisboa: IIE.
- CRONBACH, L. J. (1980). *Toward a reform of program evaluation*. San Francisco: Jossey-Bass.
- CRONBACH, L. J. (1982). *Designing evaluations of educational and social programs*. San Francisco: Jossey-Bass.
- EISNER, E. W. (1979). The use of qualitative forms of evaluation for improving educational practice. *Educational Evaluation and Policy Analysis*, 1 (6), 11-19.
- FERNANDES, D. (1992). *Prática e perspectivas de avaliação: Dois anos de experiência no Instituto de Inovação Educacional*. Documento policopiado não publicado. Lisboa: IIE.
- FERNANDES, D.; RAMALHO, G. e LEMOS, V. (1991). *Opiniões dos Professores dos Ensinos Básico e Secundário Acerca das Medidas Constantes no Projecto do Novo Sistema de Avaliação dos Alunos*. Lisboa: IIE.
- FERNANDES, D.; GÓIS, E. e CORTEZ, M. (1993). *A generalização do novo programa numa escola do 1.º ciclo: Um estudo de caso (Cadernos de Avaliação — 9)*. Lisboa: Instituto de Inovação Educacional.
- FERNANDES, D.; ASSUNÇÃO, M. I.; FARIA, P.; GIL, D. e MESQUITA, M. (1994a). *O processo de experimentação dos novos programas: 9.º ano de escolaridade (Cadernos de Avaliação — 11)*. Lisboa: IIE.
- FERNANDES, D.; ASSUNÇÃO, M. I.; FARIA, P.; GIL, D. e MESQUITA, M. (1994b). *O processo de experimentação dos novos programas: 4.º ano de escolaridade (Cadernos de Avaliação — 10)*. Lisboa: IIE.
- GEP (1979). *Ensino secundário unificado: Relatório de avaliação do 7.º ano de escolaridade*. Lisboa: GEP.
- GEP (1980a). *Ensino secundário unificado: Relatório de avaliação do 7.º e 9.º anos de escolaridade*. Lisboa: GEP.
- GEP (1980b). *Ensino secundário unificado: Inquérito a professores do 9.º ano de escolaridade*. Lisboa: GEP.
- GEP (1981a). *Ensino secundário unificado: A aprendizagem da Matemática — Influência da Escola e da Família*. Lisboa: GEP.

- GEP (1981b). *Ensino secundário unificado: Relatório de avaliação do 9.º ano de escolaridade-Entrevistas a encarregados de educação*. Lisboa: GEP.
- GEP (1991). *Ensino secundário unificado: Um diagnóstico da situação*. Lisboa: GEP.
- GIFFORD, B. e O'CONNOR, M. (1992). *Changing assessments: Alternative views of aptitude, achievement and instruction*. Boston, MA: Kluwer.
- GREENE, J.; CARACELLI, V. e GRAHAM, W. (1989). Toward a conceptual framework for mixed-method evaluation designs. *Educational Evaluation and Policy Analysis*, 11 (3), 255-274.
- GUBA, E. e LINCOLN, Y. (1981). *Effective evaluation*. San Francisco: Jossey-Bass.
- GUBA, E. e LINCOLN, Y. (1989). *Fourth generation evaluation*. Newbury Park, CA: Sage.
- GUBA, E. e LINCOLN, Y. (1991). Epistemological and methodological bases of naturalistic inquiry. In G. F. Madaus, M. Scriven & D. Stufflebeam (Eds.), *Evaluation models: Viewpoints on educational and human services evaluation* (pp. 311-334). Norwell, Massachusetts: Kluwer.
- HOUSE, E. R. (1991). Assumptions underlying evaluation models. In G. F. Madaus, M. Scriven. & D. Stufflebeam (Eds.), *Evaluation models: Viewpoints on educational and human services evaluation* (pp. 45-64). Norwell, Massachusetts: Kluwer.
- HOWE, K. R. (1985). The dogmas of educational research. *Educational Researcher*, 14 (8), 10-18.
- IIE (1990). *Relatório de avaliação da experimentação do programa do 1.º ano do 1.º ciclo: Ano lectivo 1989/1990*. Lisboa: IIE.
- IIE (1991a). *Opiniões dos professores e directores de escolas sobre a experimentação dos novos programas dos 1.º e 2.º anos de escolaridade*. Lisboa: IIE.
- IIE (1991b). *Opiniões dos professores do 5.º Ano de escolaridade acerca do processo de experimentação dos novos programas*. Lisboa: IIE.
- IIE (1991c). *Opiniões dos professores do 7.º Ano de escolaridade acerca do processo de experimentação dos novos programas*. Lisboa: IIE.
- IIE (1992a). *Os professores, os conselhos escolares e o processo de experimentação dos novos programas do 3.º Ano de escolaridade: Um estudo de avaliação*. Lisboa: IIE.
- IIE (1992b). *Os professores, os conselhos pedagógicos e o processo de experimentação dos novos programas do 6.º Ano de escolaridade: Um estudo de avaliação*. Lisboa: IIE.
- IIE (1992c). *Os professores, os conselhos pedagógicos e o processo de experimentação dos novos programas do 8.º Ano de escolaridade: Um estudo de avaliação*. Lisboa: IIE.
- IIE (1992d). *Área-Escola em experiência: Um estudo de caso numa escola C + S*. Lisboa: IIE.
- IIE (1992e). *A generalização dos novos programas do 1.º Ano de escolaridade — O que mudou na escola? (Um estudo de caso)*. Lisboa: IIE.
- IIE (1992f). *Estudo comparativo dos sistemas de avaliação dos alunos em quatro países*. Lisboa: IIE.
- IIE (1992g). *Avaliar é aprender*. Lisboa: IIE.
- JZN, J. (1987). *Mathematics, insight and meaning*. Utrecht University. Netherlands.



- KULM, G. (1990). *Assessing higher order thinking in mathematics*. Washington DC: American Association for the Advancement of Science.
- KULM, G. e MALCOM, S. (1991). *Science assessment in the service of reform*. Washington DC: American Association for the Advancement of Science.
- LEVINE, M. (1974). Scientific method and the adversary model. Some preliminary thoughts. *American Psychologist*, 661-677.
- LINCOLN, Y. e GUBA, E. (1985). *Naturalistic inquiry*. Newbury Park, CA: Sage.
- LOUREIRO, J.; MATIAS, B.; VEIGA, E.; PARDAL, L.; SILVEIRA, M. e BOAVIDA, V. (1986). *Profissionalização em exercício: Um estudo de avaliação*. Lisboa: GEP.
- MADAUS, G.; STUFFLEBEAM, D. e SCRIVEN, M. (1991). Program evaluation: A historical overview. In G. F. Madaus, M. Scriven. & D. Stufflebeam (Eds.), *Evaluation models: Viewpoints on educational and human services evaluation* (pp. 3-22). Norwell, Massachusetts: Kluwer.
- MADEY, D. L. (1982). Some benefits of integrating qualitative and quantitative methods in program evaluation, with illustrations. *Educational Evaluation and Policy Analysis*, 4 (2), 223-236.
- MATOS, J. F.; PONTE, J. P.; GUIMARÃES, H. e LEAL, L. (1993). *A aplicação do novo programa de matemática do 11.º ano: Um estudo de caso*. Lisboa: IIE.
- MAXWELL, J. A. e LINCOLN, Y. (1990). Methodology and epistemology: A dialogue. *Harvard Educational Review*, 60 (4), 497-512.
- MITCHELL, R. (1992). *Testing for learning*. New York: The Free Press.
- NUTALL, D. (1986). *Assessing educational achievement*. London: The Falmer Press.
- OLIVEIRA, I.; PEREIRA, J. e FERNANDES, D. (1993). *Desenvolvimento de instrumentos de avaliação da aprendizagem em Matemática*. Lisboa: IIE.
- OLIVEIRA, I.; PEREIRA, J. e FERNANDES, D. (1994). *Seis propostas de avaliação: Matemática*. Lisboa: IIE.
- PAGE, E. e STAKE, R. (1979). Should educational evaluation be more objective or more subjective? *Educational Evaluation and Policy Analysis*, 1 (1), 45-47.
- PONTE, J.; MATOS, J. F.; GUIMARÃES, H.; LEAL, L. e CANAVARRO, P. (1991). *O processo de experimentação dos novos programas de Matemática: Um estudo de caso*. Lisboa: IIE.
- POPHAM, W. J. (1975). *Educational evaluation*. Englewood Cliffs, New Jersey: Prentice Hall.
- RESNICK, L. B. e RESNICK, D. P. (1992). Assessing the thinking curriculum: New tools for educational reform. In B. R. Gifford e M. C. O'Connor (Eds.), *Changing assessments: Alternative views of aptitude, achievement and instruction* (pp. 37-76). Norwell, MA: Kluwer Academic Publishers.
- ROMBERG, T. e STEWART, D. (1987). *The monitoring of school mathematics: Background papers* (vol. II). Wisconsin: University of Wisconsin.
- SCHREMER, O. (1991). The teacher: A category in curriculum evaluation. *Studies in Educational Evaluation*, 17, 23-39.
- SCRIVEN, M. (1991). *Evaluation thesaurus*. London: Sage.
- SERAFINI, O.; MAGALHÃES, J. e CASTRO, R. (1991). *A experimentação dos novos programas para o 10.º ano de escolaridade: Um estudo de avaliação*. Lisboa: IIE.

- SHAVELSON, R. e STERN, P. (1981). Research on teachers' pedagogical thoughts, judgements, decision and behaviour. *Review of Educational Research*, 51 (4), 455-498.
- SMITH, N. L. (1980). Federal research on evaluation methods in health, criminal justice and the military. *Educational Evaluation and Policy Analysis*, 2 (4), 53-59.
- STAKE, R. E. (1978). The case study method in social inquiry. *Educational Researcher*, 7 (2), 5-8.
- STUFFLEBEAM, D. L. e WEBSTER, W. J.(1980). An analysis of alternative approaches to evaluation. *Educational Evaluation and Policy Analysis*, 2 (3), 5-20.
- TALMAGE, H. (1982). Evaluation of programs. In H. E. Mitzel (Ed.), *Encyclopedia of educational research* (5th edition). New York: The Free Press.
- TYLER, R. (1967). Changing concepts of educational evaluation. In R. E. Stake (Ed.), *Perspectives of curriculum evaluation*. New York: Rand McNally.
- TYLER, R. (1991). A rationale for program evaluation. In G. F. Madaus, M. Scriven & D. Stufflebeam (Eds.), *Evaluation models: Viewpoints on educational and human services evaluation* (pp. 67-78). Norwell, Massachusetts: Kluwer.
- WEBB, N. (1992). Assessment of students' knowledge of mathematics: Steps toward a theory. In D. A. Grouws (Ed.), *Handbook of Research on Mathematics Teaching and Learning* (pp. 661-686). New York: MacMillan.
- WORTHEN, B. R. (1977, April). *Eclecticism and evaluation models: Snapshots of an elephant's anatomy?* Paper presented in the Annual Meeting of the AERA. New York.
- WORTHEN, B. R. (1981). Journal entries of an eclectic evaluator. In R. S. Brandt (Ed.), *Applied strategies for curriculum evaluation* (pp. 58-90). Alexandria, VA: ASCD.