

Coleção
Perspectivas Actuais

A pilotagem

Gilbert de Landsheere

dos sistemas de educação

Como garantir a
Qualidade da Educação?



EDIÇÕES ASA

A pilotagem dos sistemas de educação

Como garantir a
Qualidade da Educação?



EDIÇÕES ASA

COLEÇÃO PERSPECTIVAS ACTUAIS/EDUCAÇÃO

Volumes publicados

PLANIFICAÇÃO E DESENVOLVIMENTO CURRICULAR NA ESCOLA
Miguel A. Zabalza

ALMEIDA GARRETT E A INTRODUÇÃO DO PENSAMENTO EDUCACIONAL DE ROUSSEAU EM PORTUGAL
Fernando Augusto Machado

ESTA VIDA DE PROFESSOR...
Bernard Houot

AS FADAS NÃO FORAM À ESCOLA
Maria Augusta Seabra Diniz

MAPAS CONCEPTUAIS – Uma técnica para aprender
Vários autores

O ENSINO LIVRE – Uma fronteira da hegemonia estatal
Coordenador Roberto Carneiro

EDUCAÇÃO E FORMAÇÃO
Viviane de Landshere

MULTIPROFISSIONALISMO E INTERVENÇÃO EDUCATIVA
Maria Adelina Abreu Garcia

À DESCOBERTA DOS NÚMEROS – Contar, cantar e calcular
Institut National de Recherche Pédagogique

DO PERFIL DOS TEMPOS AO PERFIL DA ESCOLA
Fernando Augusto Machado

OS TESTEMUNHOS ORAIS NA ESCOLA
Luís Vidigal

NOS 10 ANOS DA LEI DE BASES – Memórias e Projectos
Vários autores

OS NÓS DA REDE – O problema das escolas primárias em zonas rurais
José Maria Azevedo

IMAGENS ORGANIZACIONAIS DA ESCOLA
Jorge Adelino Costa

EDUCAÇÃO: UM TESOURO A DESCOBRIR
Unesco

O ESTATUTO DA CARREIRA DOCENTE
Maria Conceição Castro Ramos

A RACIONALIDADE DO SENTIMENTO – O papel das Artes na Educação
David Besi

PROFESSORES, FAMÍLIAS E PROJECTO EDUCATIVO
Ramiro Marques

Coleção
Perspectivas Actuais

A pilotagem dos sistemas de educação

Como garantir a
Qualidade da Educação?

Gilbert de Landsheere



TÍTULO: A PILOTAGEM DOS SISTEMAS DE EDUCAÇÃO

AUTOR: GILBERT DE LANDSHEERE

TRADUÇÃO DE: JOSÉ CARLOS EUFRÁZIO

COLEÇÃO: PERSPECTIVAS ACTUAIS/EDUCAÇÃO

DIRECÇÃO DE: LUIS ALBERTO MARQUES ALVES

© EDIÇÕES ASA, 1997

DEPÓSITO LEGAL N.º 111680/97

Setembro-1997/1º Edição

Execução Gráfica/EDIÇÕES ASA – Divisão Gráfica



EDIÇÕES ASA

SEDE

R. Mártires da Liberdade, 77
Apartado 4263 / 4004 PORTO CODEX
PORTUGAL

DELEGAÇÃO EM LISBOA

Av. Dr. Augusto de Castro, Lote 110
1900 LISBOA • PORTUGAL

SUMÁRIO

PREFÁCIO	- 9 -
INTRODUÇÃO	11
CAPÍTULO I – NATUREZA DA PILOTAGEM	17
1. Tipos de pilotagem	17
1.1 Pilotagem administrativa ou monitoragem de conformidade	17
1.2. Pilotagem formativa ou pilotagem diagnóstica	17
1.3. Pilotagem do rendimento escolar	18
2. Modelo tecnológico – Modelo humanista	19
2.1. Sobre o modelo tecnológico	20
2.2. Sobre o modelo humanista	21
2.3. Atitudes tecnocráticas – Atitudes democráticas	22
3. Macropilotagem – Micropilotagem	24
3.1. Macropilotagem	24
3.2. Micropilotagem	24
3.3. Uma complementaridade	25
4. Sobre o valor acrescentado	26
5. Coerência, qualidade, equidade	29
5.1. Recapitulação das questões	32
6. Definir e chegar a acordo sobre as finalidades da educação	33
7. Desenvolver os serviços de estatística	34
8. Pilotagem da pilotagem	35
CAPÍTULO II – INDICADORES	36
1. Noção de indicadores	36
2. Necessidade de um modelo	41
3. Número ideal de indicadores?	42
4. Qualidades pretendidas	43
5. Directivas para uma utilização construtiva	44
6. Síntese dos problemas gerais	45
CAPÍTULO III – EXEMPLOS DE REALIZAÇÕES	48
INDICADORES NACIONAIS	48
1. Estados Unidos	48
1.1. A "National Assessment of Educational Progress" (NAEP)	48
1.1.1. Relatórios de Investigação	50
1.1.2. Acervo de publicações	60

1.1.3. The National Education Goals Reports	60
1.1.4. Outras avaliações nacionais importantes: estudos longitudinais	61
1.1.5. Sistemas de indicadores dos próprios Estados	64
2. Inglaterra	66
2.1. Unidade de Avaliação do Rendimento Escolar	66
2.2. Currículo Nacional e sua avaliação: uma pilotagem sistemática	67
2.3. Indicadores de rendimento	72
3. França: um desenvolvimento exemplar	73
3.1. Avaliação das aquisições durante a escolaridade	74
3.2. Evolução da política educativa	78
3.3. Produção e difusão de instrumentos de avaliação	78
3.4. Alguns esclarecimentos sobre a testagem em grande escala	78
3.5. Notável conjunto de publicações	79
4. Canadá	81
4.1. Quebec: indicadores do ensino primário e secundário	81
4.2. Indicadores nacionais do Canadá	83
5. Suécia	85
5.1. Sistema precursor: harmonização liberal da atribuição de notas	85
5.2. Avaliação nacional	87
6. Austrália	88
6.1. Pilotagem das escolas e dos sistemas escolares nos Estados	88
6.2. Projecto nacional	90
7. Outros países	91
INDICADORES INTERNACIONAIS	91
1. Associação Internacional de Avaliação do Rendimento Escolar (IEA)	92
1.1. Primeiros tempos da IEA: Um estudo de exequibilidade	93
1.1.1. Estudo sobre o rendimento do ensino da matemática (1961-1967)	93
1.1.2. Survey de seis disciplinas (1966-1973)	97
1.1.3. Segundo estudo sobre o ensino da matemática (1980-1986)	100
1.1.4. Segundo estudo sobre ciências (1984)	100
1.1.5. Estudo sobre a utilização pedagógica dos computadores (1987-1990)	102
1.1.6. Estudo sobre a educação pré-escolar: qualidade de vida das crianças de quatro anos (1986-1996)	102
1.1.7. Investigação sobre compreensão da leitura (1990-1993)	104
1.1.8. Terceiro estudo sobre o ensino da matemática e das ciências (1991-1995)	107
1.2. Papel histórico e contribuições metodológicas da IEA	107
1.2.1. Um processo inevitavelmente longo	109
2. International Assessment of Educational Progress (IAEP)	110
2.1. Estudo IAEP I (1987-1988)	110
2.2. Estudo IAEP II (1990-1991)	111
2.3. Crítica fundamental	111

3. Projecto da OCDE: International Indicators and Evaluation of Educational Systems (INES) – Indicadores internacionais do ensino	113
3.1. Primeira experiência: indicadores sociais	113
3.2. Projecto INES	115
3.2.1. Problemas particulares relativos às atitudes, concepções e expectativas	123
3.2.2. O Futuro	125
CAPÍTULO IV – MICROPILOTAGEM	
Generalidades	127
1. Realizações	129
1.1. Inglaterra: um exemplo de <i>checklists</i> de auto-avaliação	129
1.2. França: Proposta de Peretti	131
1.3. Dinamarca	134
1.4. Nova Zelândia	134
2. Recurso a peritos externos	135
3. Conclusão	135
CAPÍTULO V – OS PRÓS E OS CONTRAS	
1. Benefícios auferidos	137
1.1. Ruptura do isolamento	137
2. Perigos	138
3. Desde que saibamos servir-nos...	140
4. ... Retendo o que nos convier	143
ANEXO: ASPECTOS TÉCNICOS	
Introdução	145
1. A Avaliação das opções tecnológicas na perspectiva da sociedade	146
2. A Avaliação de programas	149
3. Psicologia cognitiva e novas perspectivas	151
3.1. Implicações na melhoria dos testes de rendimento escolar	152
3.2. Em conclusão	153
4. Validade e fidelidade das medidas	153
4.1. Capacidade de generalização	156
4.2. Estimativa da ocasião de aprender	156
4.3. Provas de rendimento ou provas de memorização?	156
4.3.1. Necessidade de processos de moderação	158
5. A administração das provas	159
5.1. Amostragem multimatricial	159
5.2. Comparação de testes	159

6. A interpretação dos resultados	160
6.1. Noção de padrão	160
6.2. Determinação da competência mínima	161
6.2.1. Fixação da nota de sucesso	161
6.3. Análise da percentagem de sucesso	166
6.4. Probabilidade de resposta correcta	167
6.4.1. Algumas particularidades sobre a teoria dos traços latentes	168
6.5. Fixação de escalas	170
6.5.1. Exemplos de escalas de competência	171
6.5.2. Crítica da teoria das respostas aos <i>itens</i> (Goldstein e Wood, 1989)	173
6.6. Funcionamento diferencial dos <i>itens</i>	174
Conclusão	174
PRINCIPAIS ABREVIATURAS	177
BIBLIOGRAFIA	179

PREFÁCIO

Arrastados pela crise cultural, de que a nova revolução industrial é apenas um aspecto, os países industrializados vêm, sucessivamente, instalando a pilotagem dos seus sistemas de ensino.

Estas inovações questionam o mundo da educação, dadas as questões de fundo que suscitam, e os problemas técnicos de grande complexidade que colocam.

Com este livro, pretende-se analisar a questão e apresentar um quadro das principais realizações – macroscópicas e microscópicas – na esperança de daí retirar lições e inspiração. Pretende-se, em especial, incitar os responsáveis pela recolha e publicação de indicadores, a concentrarem o essencial dos seus esforços na avaliação dos produtos da educação. O que os alunos aprendem, não será o mais importante? Cada professor na sua aula está tão implicado neste processo, como qualquer dos decisores e responsáveis máximos, ou qualquer dos outros membros da comunidade educativa.

Devido a um dos principais objectivos a alcançar – estimular a tomada de decisões – o tratamento deste tema assume, muitas vezes, a forma dum *dossier*, isto é dum conjunto ajustado de documentos e comentários, destinados a apresentar o aspecto geral duma questão e a fornecer, directa ou indirectamente, pistas para a acção.

A escolha de alguns exemplos, recolhidos num conjunto de trabalhos que manipulam centenas de milhões de dados, é arbitrária. Tem apenas valor ilustrativo. A bibliografia orientará o leitor desejoso de maior rigor e aprofundamento.

Finalmente, pode acontecer com frequência que os temas tratados, aparentemente simples, exijam de facto o recurso a teorias e técnicas complexas. Algumas delas são ainda objecto de controvérsia entre os especialistas. Uma obra de introdução como esta, não é o local adequado para debates deste tipo. Porém, seria perigoso ignorá-los totalmente. Esta a razão por que se resumem, em anexo, algumas destas teorias e técnicas.

O presente trabalho foi muito facilitado pela colaboração nas investigações, durante mais de trinta anos, da Associação Internacional de Avaliação do Rendimento Escolar (IEA), e pela participação nas actividades do grupo, de peritos que, com Norberto Bottani, prepararam tecnicamente o lançamento do projecto "Indicadores do Ensino", da Organização de Cooperação e Desenvolvimento Económico (OCDE). O capítulo sobre indicadores apoia-se, em grande

parte, no conjunto de documentos reunidos sobre este assunto, pelo Centro de Investigação e Inovação no Ensino (CERI), da OCDE.

Beneficiei, ainda, dos estímulos e das críticas construtivas de vários colegas do Instituto Jules Destrée, em especial de Gérard Fourez, que não me regateou as suas preciosas sugestões, e de M. Crahay, J.-P. Pourtois, H. Desmet, e J.-E. Charlier. Estou-lhes profundamente agradecido! Agradeço, também, aos numerosos colegas e amigos que puseram à minha disposição e autorizaram a utilização de documentos por vezes inéditos.

INTRODUÇÃO

"Outrora embiocada, guardando ciosamente os seus segredos e zonas sombrias, avara de resultados e desconfiando dos números, guiada por um santo horror aos 'quadros de honra', arrepiada perante a possibilidade de se compararem, em cada região, município ou cidade, os resultados escolares dos diferentes estabelecimentos de ensino, a educação nacional sofreu, nos últimos anos, uma pequena revolução. Sob o impulso dos utentes da escola, a quem o mito da igualdade de todos os alunos no cadiño da escola já não ilude, pressionada também pelos eleitos embarcados, de bom ou de mau grado, neste navio, mas renitentes em pagar sem ver, a educação aprendeu, progressivamente, a nada dissimular." (C.Garin, *Le Monde*, 14 de Janeiro de 1993).

"Só instrumentos complexos são capazes de dominar a complexidade da realidade" (R. Saint-Geours, *Elogio da Complexidade, Economica, 1987*).

Etimologicamente a ideia de pilotagem, de condução, é inerente à de educação – dado que esta palavra deriva de *ducere*. De facto, educar é *conduzir* a um determinado estado pretendido. Quem guia deve, não só conhecer o caminho, como ter o cuidado de não se afastar, correndo o risco de se perder. Deve, além disso, respeitar a liberdade do viajante, porque, devido a esta liberdade, o educando surge como sujeito com o mesmo valor que o educador, particularidade que fundamenta a própria noção de educação no sentido nobre do termo.

Contudo, o respeito pela liberdade tem os seus limites, uma vez que a cultura em que a criança se deve, paulatinamente, inserir, não está inscrita nos seus genes. Terá, pois, de se iniciar na linguagem, nos conhecimentos, nas técnicas e, numa maneira geral, em tudo o que constitui a sua própria civilização, desenvolvendo o seu espírito crítico, a sua capacidade de divergir, a sua independência.

De uma maneira mais directa, a pilotagem pode definir-se como uma tomada de decisão, a nível macroscópico (isto é, de um sistema ou de um subsistema educativo), ou microscópico (isto é de um estabelecimento, ou até mesmo de uma aula), na base de dados verificados tão objectivamente quanto possível, relativos à situação, ao funcionamento ou aos produtos dos sistemas. Estes dados não falam por si mesmos: são o resultado final da construção de um parecer, a partir de informações qualitativas e quantitativas que constituem os

indicadores. Deste modo, as decisões podem ser tomadas com melhor conhecimento daquilo que está por detrás das opções tomadas, e dos objectivos que se pretendem com elas.

A comparação com a tecnologia médica é, em parte, esclarecedora. A *monitoragem* faz-se a partir de sinais, de "piscadelas" que os dispositivos electrónicos emitem, quando um indicador (por exemplo, o ritmo cardíaco) atinge ou ultrapassa um limite pré-definido, ou quando o dispositivo se avaria, devido a uma anomalia de funcionamento ou de alimentação. Porém, a comparação tem limites, dado que o fenómeno educativo contém vários aspectos que nos escapam do ponto de vista racional.

A instauração duma pilotagem deve ser sempre precedida por uma profunda reflexão sobre o projecto educativo, tanto no mundo político como na sociedade civil. É dos resultados, sempre provisórios, desta reflexão e da clarificação dos objectivos políticos que devem depender as linhas de força da pilotagem.

Esta clarificação diz respeito, em primeiro lugar, aos fins e objectivos, cuja escolha depende, em última instância, de juízos de valor. Ora, em matéria de ensino os interesses divergem, bem como os valores. Divergem, não apenas ao nível dos poderes organizadores, mas também ao nível dos que reflectem sobre a função da escola na sociedade e ao nível dos consumidores directos que são os alunos. Só mediante negociação política, podem surgir acordos ou compromissos sobre os fins da educação.

Embora não possam forjar decisões, cabe aos especialistas ajudar a definir os objectivos e imaginar soluções a propor aos actores sociais. Como sublinha G. Fourez (comunicação pessoal), "o especialista não pode funcionar adequadamente como um perito neutro, que apresentasse dificuldades absolutas e soluções únicas, mas tem o seu lugar como 'representante' de dificuldades – no sentido de que as torna presentes – e criador de soluções novas no debate sociopolítico."

A ideia de pilotagem não é nova. Quando o presidente dos Estados Unidos criou o primeiro Ministério da Educação em 1867, determinou-se que este deveria publicar, todos os anos, "estatísticas e observações que informassem sobre a situação e o progresso da educação". Sabemos hoje, que um longo caminho havia a percorrer, não só nos Estados Unidos (que publicam, actualmente, todos os anos um estudo estatístico, relativo a sessenta indicadores, intitulado *Situação da Educação*), mas também nos outros países.

Até há pouco tempo, as estatísticas relativas à educação eram, essencialmente, concebidas para a planificação e gestão. Partindo de dados demográficos e de opções gerais, como a instauração da escolaridade obrigatória até determinada idade, tratava-se sobretudo de determinar o nível de recursos orçamentais,

materiais e humanos, necessários (essencialmente dados de *input*); a produtividade dos sistemas media-se, principalmente, em termos de diplomados (de quem se ia analisando a origem socioeconómica).

As recentes evoluções da ciência, da tecnologia, da economia, e as conquistas sociais como a reivindicação da igualdade de oportunidades, o direito ao sucesso nos estudos, juntamente com o direito à diferença, deram origem a cada vez maior número de interrogações sobre a **qualidade da educação**. Grande parte do interesse desloca-se, assim, das variáveis de *input* para as variáveis de *tratamento* ou *processo*, e de *output*. A pilotagem da qualidade, em cada momento do *processo* e nos aspectos essenciais dos *produtos*, é o princípio chave da gestão da qualidade geral.

A instauração da pilotagem, num número cada vez maior de países, tem uma explicação cultural, orçamental, económica, política e social.

No sistema educativo, não se produzem nunca inovações importantes, por acaso. De facto, ele evolui em função do contexto cultural e das forças que dominam a sociedade. Assim, o rápido incremento da pilotagem do sistema educativo nos Estados Unidos explica-se, por um lado, pelo choque psicológico provocado pelo lançamento do Sputnik pela URSS (afinal a ciência americana não era a primeira!), pela crescente inquietação suscitada por um rendimento escolar gravemente insuficiente, e também pelo progresso dos direitos cívicos que garantem, em princípio, uma educação de qualidade para todos, pobres e ricos, negros e brancos.

De um modo geral, as sociedades avançadas actuais caracterizam-se por dois fenómenos principais: a intelectualização, cada vez maior, de todos os sectores de actividade e, nos casos mais favoráveis, uma crescente adesão aos valores fundamentais da democracia.

A referência à explosão da ciência e da tecnologia, tornou-se hoje um lugar comum. Ninguém mais duvida que o saber e o saber fazer de alto nível (o factor *k* – referido a *knowledge* – tão caro aos economistas contemporâneos) são as chaves do futuro económico e social.

“É claro que a terra e o capital continuam a existir, só que os seus valores dependem de processos que se desenrolam num outro espaço: o do conhecimento. (...) Em todas as esferas da economia, e não apenas nos serviços, a gestão inteligente de informações e de saberes em permanente transformação, surge como uma condição de sobrevivência. (...) Apesar de não termos esse hábito e de ser difícil, torna-se indispensável fazer um esforço para repensar o saber. Daqui em diante, o saber surge como uma infra-estrutura. É nele que se baseia a prosperidade e situação económica das nações, é ele que constitui factor decisivo entre empresas em competição, e é por isso que já não podemos deixar de pensar o conhecimento” (R.D. Collin, 1993:1-2).

A situação não é isenta de perigos. "O crescimento quase cancerígeno das tecnologias ameaça a própria existência da sociedade, ao mesmo tempo que cria prodigiosas oportunidades para a humanidade" (Gonod, 1990:2). Assim se explica a importância a dar à *avaliação do impacto social destas tecnologias*, e a pilotagem tal como é concebida hoje em dia é uma delas.

As razões orçamentais que incitam à pilotagem são bem conhecidas: já lá vai a época da riqueza dos *Golden Sixties*, que permitiu tantas iniciativas generosas, principalmente em matéria de igualdade de oportunidades. A expansão económica no ocidente afrouxou. Para conservar o seu lugar no mercado mundial as nações mais avançadas esforçam-se por promover a excelência e por aumentar a qualidade geral do seu ensino, limitando sempre as despesas ao máximo.

Finalmente, no plano da política geral, a luta pelo respeito dos direitos do homem, em especial o direito a uma educação de qualidade, começa a dar os seus frutos. Um dos aspectos principais deste direito diz respeito ao desenvolvimento máximo de cada um, no respeito pela sua diferença.

Em resumo, há duas preocupações dominantes que importa conciliar:

1. Permite a educação que cada qual realize o seu projecto pessoal?
2. Será que a educação equipa, convenientemente, as nações e os indivíduos chamados a enfrentar a nova revolução industrial e cultural que se está a dar à escala universal?

Como sublinha J.-E. Charlier (comunicação pessoal), não é possível evitar uma tensão ideológica entre os que acham que a escola é uma instituição encarregada de fornecer à sociedade aquilo de que ela necessita, se preciso for violentando os seus utentes, e os que consideram que ela deve dar resposta às necessidades dos alunos. E é utopia pensar que as duas posições podem caminhar a par, ou combinar entre si harmoniosamente.

A problemática da busca da qualidade deve estar nos nossos horizontes. De facto, como sublinha a OCDE na sua obra *Une éducation et une formation de qualité pour tous* (1992), a acusação de que o "nível" da educação está a baixar, ou já não é adequado, poderia levar a pensar que o remédio para esta situação se deve procurar nos objectivos do passado e na maneira de ensinar de outrora. Ignorar-se-iam, deste modo, as alterações económicas, sociais, culturais e pedagógicas acabadas de referir.

Quem pretende fazer pilotagem deve começar por definir, tão objectivamente quanto possível, a situação da educação, e sobre isto é importante estabelecer um acordo. Este primeiro passo é, de facto, a condição primordial duma reflexão colectiva, dum debate frutuoso e construtivo. "Enquanto os interlocutores divergirem sobre a situação concreta, dificilmente poderão convergir na busca de

soluções" (W. Hutmacher, 1993:147). Ora, sabe-se como é acesa a controvérsia a propósito do "nível" dos "produtos" da escola e da sua evolução.

Uma vez de posse das informações pretendidas, uma vez feita a "vistoria", há que tirar conclusões e decidir que fazer. Para assegurar o respeito pelo projecto educativo de cada um, há que observar, de modo contínuo, o comportamento global do sistema educativo (aspecto macroscópico) e determinar, ao mesmo tempo, em que medida ele oferece a cada um as desejadas ocasiões de aprender e, portanto, em que medida responde à procura individual (aspecto microscópico). Ambos os aspectos são igualmente importantes. O sistema educativo funciona bem quando consegue servir, simultaneamente, a colectividade e cada um dos seus membros em particular.

O que surpreende ainda, hoje em dia, quando se reflecte sobre a pilotagem do sistema educativo é, no fim de contas, as poucas informações precisas e seguras que se possuem sobre o seu funcionamento e sobre o que produz por si mesmo, e em comparação com outros sistemas.

É certo que é praticamente impossível explicar de maneira completa a génesis, as causas do rendimento escolar, tantos são os factores determinantes, em grande parte aliás exteriores à escola. Tal complexidade abrange tudo o que é humano. Não é porém razão suficiente para deixar de reflectir e de agir.

O passado

No passado, os responsáveis pelo ensino não foram certamente indiferentes ao sucesso do seu trabalho. Em especial os alunos há muito tempo que estão sujeitos a exames assim como os professores à inspecção.

Desde há muito tempo que vêm sendo denunciadas as limitações, e até mesmo as injustiças, destes tipos de controlo. Contudo, se nos situarmos no quadro social do passado, as coisas não correram assim tão mal, no sentido em que, se de facto houve muitos estragos, talentos desperdiçados e erros de apreciação, também as necessidades da sociedade puderam ser satisfeitas, é certo que muitas vezes não equitativamente.

Fazer a pilotagem dum sistema educativo é algo mais do que acumular indicadores. A pilotagem deve necessariamente incluir três componentes: a recolha regular de informações, a avaliação destas informações e a sua tradução em acções institucionais.

Uma pilotagem que informe de modo contínuo e que permita fazer as opções educativas mais favoráveis, torna-se indispensável. Em última análise, pode ser considerada como uma condição *sine qua non* de sobrevivência da nossa sociedade actual e, sobretudo, da sociedade de amanhã.

Um acto democrático

Determinar, tão objectivamente quanto possível, em que medida são atingidos pelo ensino os objectivos que a comunidade educativa lhe consignou e informar todos os interessados dos resultados desta observação, é uma obrigação. De facto, em democracia, qualquer cidadão tem o direito de ser informado sobre tudo o que lhe diz respeito: aqueles a quem os cidadãos confiaram o poder ou uma missão devem prestar-lhes contas rigorosas e inteligíveis. Em princípio, este direito estende-se à própria escolha dos instrumentos e das modalidades de avaliação.

Já não basta, pois, anunciar que Pedro ou Paulo é o primeiro ou o último da turma (*padrão social*), há que informar se progrediu, efectivamente, no seu projecto pessoal (*padrão individual*), e em que medida domina ou não os conhecimentos, o saber fazer e o saber ser, indispensáveis ao seu desenvolvimento e à sua inserção na vida social (*padrões criteriais*).

“A avaliação por indicadores, supõe que se faz do utente do sistema educativo um seu membro a tempo inteiro, que ele está plenamente integrado nesse mesmo sistema. Aliás, não se vê bem como é que os indicadores de resultados e de rendimento podem ser definidos sem a opinião dos utentes do sistema educativo: alunos, pais, empregados, eles próprios instituídos como sujeitos e já não como objectos dos sistemas de educação. Está na lógica do movimento dos indicadores conduzir a um reforço do poder dos utentes dentro dos sistemas educativos” (F. Ewald, 1991:13). A avaliação das tecnologias na perspectiva da sociedade é chamada a desempenhar, aqui também, um papel mais importante. Voltaremos a este assunto.

Uma necessidade pedagógica

De maneira mais ou menos consciente, todo o professor reinterpreta os programas oficiais, por um lado em função da sua personalidade e qualificação profissional e, por outro lado em função das características dos seus alunos. Dispõe, em geral, de poucos elementos que lhe permitem saber se os seus alunos atingem, realmente, os conhecimentos mínimos necessários. Ainda por cima, não é suficientemente apoiado na identificação das dificuldades de aprendizagem dos seus alunos; e contudo, sem diagnóstico correcto, não pode haver remediação adequada.

CAPÍTULO 1 – NATUREZA DA PILOTAGEM

1. TIPOS DE PILOTAGEM

No campo do ensino, há a distinguir três tipos de pilotagem (C.E.Richards, 1988; V. de Landsheere, 1992, p.475):

1.1. Pilotagem administrativa ou monitoragem de conformidade

Vela pelo respeito dos dispositivos regulamentares (*compliance monitoring*). Trata-se de determinar se as escolas atingem um nível de exigência pré-estabelecido. Neste caso, a unidade de análise é, muitas vezes, o estabelecimento ou a área escolar. As questões postas são do tipo:

- Os professores possuem as devidas habilitações?
- Os currículos praticados correspondem ao projecto pedagógico?
- Os equipamentos (bibliotecas, salas de computadores,...) são satisfatórios?
- Os transportes escolares estão bem assegurados?
- Etc.

Este tipo de pilotagem apresenta duas características principais.

- Trata principalmente do *input*.
- Tem como fim principal assegurar um nível de serviço adequado. Os critérios de conformidade são definidos por leis, regulamentos ou instruções administrativas. Geralmente a avaliação é feita no terreno, recorrendo sobretudo a *checklists*.

1.2. Pilotagem formativa ou pilotagem diagnóstica

Baseia-se nos indivíduos. Por exemplo, a aplicação periódica de testes centrados nos objectivos, abrangendo o essencial dum programa de estudos, permite identificar problemas de aprendizagem que podem surgir em determinado ambiente escolar e procurar solução para eles.

Este tipo de pilotagem faz-se, sobretudo, com a ajuda de provas criteriais, provas para as quais a definição dos principais domínios abrangidos pelo currículo é de primeira importância (Ver a este respeito V. de Landsheere, 1988:59 e seguintes).

O desenvolvimento deste tipo de pilotagem processa-se, em traços largos, do seguinte modo:

1. Os alunos são submetidos a um pré-teste para determinar o seu nível de conhecimentos antes de iniciarem a aprendizagem.
2. Os professores são convidados a concentrar os seus esforços nos objectivos em que os alunos revelaram mais dificuldades.
3. Os professores propõem tarefas para realizar, tanto na escola como em casa, para ajudarem os alunos a atingir estes objectivos. Os alunos são informados, logo que atingem um nível de aprendizagem satisfatório.
4. Em seguida, os alunos são submetidos a um novo teste.
5. Em função dos resultados obtidos, o professor elabora, eventualmente, um plano de remediação.

Os princípios fundamentais da pedagogia para a mestria – que se aplica tanto ao ensino colectivo como ao ensino individualizado ou semi-individualizado – ficam assim respeitados.

1.3. Pilotagem do rendimento escolar

Até há pouco tempo realizava-se, em geral através de testes normativos que, de modo comparativo, iam dando conta do que “produziam” as escolas e cursos duma região, duma província ou dum país. Assiste-se actualmente a numerosas tentativas de avaliações criteriais.

A pilotagem do rendimento pode conceber-se sob duas perspectivas diferentes. Em nossa opinião, deve essencialmente fornecer informações comparativas destinadas, por um lado, aos poderes organizadores preocupados em conhecer o nível geral e o grau de homogeneidade do produto do sistema e em informar a comunidade educativa sobre o assunto, e por outro lado, a fornecer *confidencialmente* aos professores informações sobre a situação das suas turmas em relação às outras. De uma maneira geral, de acordo com este ponto de vista, a publicação de resultados nunca permite identificar as escolas.

Nos países em que estas informações se tornaram públicas e em particular nos Estados Unidos (outros países começam já a seguir este exemplo), pretende-se aplicar a **lei do mercado** ao sistema escolar.

No mercado a concorrência actua em pleno: os consumidores utilizam o máximo de informações para avaliarem a qualidade e o preço dos produtos que lhes são oferecidos. Dum modo semelhante, também os pais bem informados

sobre as qualidades respectivas das escolas podem optar em consequência. E se a escola frequentada pelos filhos se apresenta numa situação desfavorável em relação às outras, os pais irão, eventualmente, exercer pressão para que a situação melhor. Por vezes, recorre-se à transferência para um estabelecimento que consiga melhores resultados. Ou, pelo contrário, dá-se preferência a uma escola menos exigente se se trata de alunos com dificuldades...

Chegados a este ponto, importa sublinhar quanto pode ser injusta uma comparação simplista entre estabelecimentos de ensino. Qual o que tem mais mérito? O que recebe alunos com dificuldades e consegue fazê-los progredir significativamente, ou o que apenas recebe ou conserva os melhores alunos e as turmas que ocupam os primeiros lugares nos quadros de honra?

De facto, duas questões fundamentalmente diferentes se colocam.

Por um lado, uma nação ou outra entidade política importante podem querer conhecer o que o seu sistema educativo produz, tanto para adaptar a sua política às necessidades como para avaliar **o capital humano** de que dispõem. Neste caso, são os indicadores globais que adquirem mais importância (**macropilotagem**).

Por outro lado, a questão pode dizer respeito à qualidade das escolas consideradas individualmente. Neste caso, é o valor acrescentado em cada uma delas que oferece mais interesse (**micropilotagem**).

2. MODELO TECNOLÓGICO – MODELO HUMANISTA

De uma maneira esquemática, podemos opor entre si dois tipos de gestão dos sistemas educativos: o modelo tecnológico (ou modelo de tendência sócio-económica) e o modelo humanista (ou modelo de tendência psicologisante).

Numa perspectiva puramente tecnocrática da gestão do ensino, faz-se de algum modo a economia da condição humana: face a um produto educativo pretendido, faz-se um esforço por reunir as matérias primas em melhores condições (recrutamento de pessoal e aquisição de meios), para desencadear um processo ou tratamento que permita obter o resultado desejado. Na prática procura-se, de forma mais ou menos empírica, que “as coisas andem o melhor possível”, sem nos preocuparmos muito com as consequências individuais e sociais.

Pelo contrário, a perspectiva humanista tem como primeiro critério o respeito pela pessoa humana que não pode nunca ser considerada como um simples instrumento, como um actor escravizado. Procura-se, sempre, respeitar a sua liberdade e ter em conta os seus projectos pessoais, desde que estes não firam a liberdade e dignidade dos outros.

De facto, a realidade não surge assim tão dividida: o modelo tecnológico não exclui um certo debate democrático, enquanto que o modelo humanista não é incompatível com a utilização das tecnologias. Mas as tecnologias utilizadas por uma parte e pela outra não são sempre as mesmas.

2.1. Sobre o modelo tecnológico

O modelo tecnológico *entrada, tratamento, saída, feedback ou retroacção* é hoje bem conhecido. Visa optimizar a produtividade. A sua maior preocupação é produzir com a melhor qualidade e ao menor preço.

Esta concepção de pilotagem assemelha-se, por vezes, ao *behaviourismo skinneriano*, em especial à engenharia *comportamental*. É evidente que o gestor de um sistema que diz respeito a centenas de milhar de pessoas deve, necessariamente, utilizar duma maneira racional, instrumentos adaptados ao macroscópico. Contudo, o modelo tecnológico não implica forçosamente que se encare sob um aspecto mecânico o fenómeno humano por excelência que é a educação. A sua aplicação exige adaptações e reservas, e pode aliás entrar em linha de conta com uma margem de indeterminação.

Esquematicamente o modelo tecnológico pode descrever-se da seguinte maneira.

Depois de uma análise das necessidades onde o factor económico ocupa geralmente um lugar importante, são assinalados objectivos mediante um programa mais ou menos obrigatório. Procuram-se os meios mais eficazes e menos dispendiosos para o pôr em prática. O rendimento, isto é, o nível dos resultados alcançados pelos alunos é avaliado no fim de cada etapa de estudos. (É preciso contudo não esquecer que uma análise das necessidades nunca é neutra: é influenciada pela opção da sociedade e pelos instrumentos de análise e quadros teóricos utilizados.)

Uma gestão deste tipo é, geralmente, realizada por um poder central que decide sobre os programas e sobre a maneira de os incrementar. Aos que ocupam o fundo da pirâmide hierárquica resta-lhes executar o trabalho (*modelo centro-periferia*).

O recurso às tecnologias, pode ter efeitos sociais indesejados e, por vezes, irreveresíveis. Por exemplo, o perigo de descentralizar o currículo, submetendo-o aos vários pontos de focalização da pilotagem, não deve ser subestimado. A problemática dos efeitos de refluxo dos exames, tão bem estudados por Madaus e Kellaghan, vai servir-nos como primeiro exemplo.

Madaus e Kellaghan (1922:143) distinguem duas grandes categorias de avaliação. Na primeira, situam as avaliações que se realizam diariamente no decurso

do processo de ensino (simples controlos, que não constituem qualquer ameaça, de saberes ou competências; avaliações formativas,...), sem um objectivo importante. Na segunda categoria agrupam as avaliações cujos resultados podem ter consequências graves para o indivíduo (acesso ao ensino superior, obtenção de um diploma de fim de curso,...). Neste segundo caso observa-se, geralmente, um importante *efeito de refluxo negativo*: a preocupação pelo sucesso nas provas comanda a escolha dos conteúdos e da didáctica, mais do que o desejo de garantir uma educação da melhor qualidade possível.

Os mesmos autores observam (p.144) que, já em 1868, uma comissão de inquérito inglesa concluía: "É o exame que determina o currículo e não o inverso. Desencoraja a experimentação pedagógica, limita a escolha das matérias a ensinar, entrava a liberdade no uso de metodologias e não contribui para a adesão da turma aos mais altos valores."

De facto, um ensino que, por exemplo, vise prioritariamente o sucesso em exames que se baseiem em conhecimentos factuais, corre o risco de o fazer à custa do desenvolvimento social, estético, afectivo e físico dos alunos. Esta influência pode exercer-se não só sobre o ano escolar a que o exame diz respeito, mas também sobre anos precedentes (esforço individual de um professor para defender a sua reputação, ou esforço colectivo para defender a imagem da escola). Por seu lado, os alunos logo que têm essa possibilidade optam, muitas vezes, mais em função das exigências de determinado programa do que dos seus interesses reais.

Por outro lado, as provas nacionais estandardizadas podem prejudicar a adopção de currículos locais, adaptados às necessidades de determinados meios particulares.

Pelo contrário, também é verdade que o efeito de refluxo pode ser positivo, incitando professores e alunos a darem mais valor a conteúdos ou a processos importantes que tenham, eventualmente, tendência a negligenciar.

Há outro aspecto a que se deve prestar muita atenção: a democratização do domínio dos instrumentos de pilotagem. O problema é delicado porque actualmente, a maior parte dos professores não estão suficientemente familiarizados com os métodos e técnicas – cada vez mais sofisticados – de avaliação.

2.2. Sobre o modelo humanista

Sem esquecer o inevitável encadeamento: *objectivo – input – tratamento – output – feedback*, a perspectiva humanista tem em conta outras prioridades e dá muito mais importância ao aspecto qualitativo e ao aspecto ético.

Para o psicólogo humanista os principais objectos de estudo são a experiência, tal qual ela é vivida por cada pessoa e o sentido que esta lhe confere. A preocupação constante é a dignidade e valor de cada indivíduo; cada qual deve poder desenvolver o seu potencial próprio. É dada grande importância ao que poderíamos chamar um sã desvio, isto é, à expressão dos talentos pessoais e à faculdade de seguir, livremente, o caminho que cada um escolheu para realizar a sua própria actualização. Ser eu próprio é, de facto, poder divergir. Neste contexto, os sentimentos e valores têm sempre direito a ser respeitados e o indivíduo fica livre para construir o seu futuro. A liberdade, o amor, o dom de si mesmo, a adesão a valores, a autodisciplina, a lealdade, a criatividade, a esperança passam a estar em primeiro plano.

"A finalidade do domínio social da tecnologia" escreve Gonod (1990:10), "é a sua regulação em função das necessidades da humanidade, se nos colocarmos no plano global das sociedades e das colectividades e, em última análise, dos indivíduos."

O *modelo periferia-centro*, anda associado à opção humanista. Tende a descentralizar a gestão do sistema educativo, e a dar um poder de iniciativa máximo às "bases".

É nítida esta tendência de descentralização a favor das regiões, das comunidades educativas locais, mas também dos estabelecimentos de ensino que são estimulados a formular projectos educativos próprios, o que implica também uma certa independência na gestão, na utilização de recursos, na concepção dos currículos, assim como na formação ou aperfeiçoamento do pessoal. "A autonomia é, sem dúvida, a expressão mais adequada da modernidade, a chave de abóbada de qualquer obra de renovação. (...) E só existirá, realmente, se transparecer na gestão quotidiana das escolas" (Dupont, 1988:150).

Actualmente observam-se duas tendências contraditórias. Concede-se mais autonomia às escolas mas, em contrapartida, reforçam-se os dispositivos de avaliação, de modo a levar as escolas a prestar contas explícitas sobre a qualidade do seu trabalho. Estimula-se a autopilotagem mas, ao mesmo tempo, desenvolve-se a macropilotagem de modo a determinar em que medida as escolas, no seu conjunto, atingem os objectivos julgados essenciais, e portanto prioritários, pelo poder organizador ou a comunidade educativa global.

2.3. Atitudes tecnocráticas – Atitudes democráticas

G. Fourez gosta de realçar que no debate que agora nos ocupa, a dimensão que engloba os aspectos relacionados com a sociedade é, em última análise, a

mais decisiva. Isto porque, embora quase sempre preocupadas – em graus diversos, é certo – com os valores sociais e humanos, todas as pilotagens fazem apelo às tecnologias. Algumas apresentam, contudo, um carácter tecnocrático muito vincado, no sentido em que partem do princípio que os saberes científicos e técnicos devem fornecer directamente à sociedade, soluções para os seus problemas: o debate político característico das democracias fica, assim, escamoteado. “Outras pilotagens seguem uma lógica decisional, para nos servirmos das categorias de J. Habermas, deixando aos decisores políticos a escolha dos objectivos, pretendendo situar-se apenas ao nível dos meios. Esta maneira de abordar a questão, já deixa mais espaço aos debates e às negociações políticas, pelo menos em relação às finalidades, mesmo que se mantenha a perspectiva tecnocrática em relação aos meios. Finalmente, o modelo pragmático de Habermas, alerta para a necessidade de um vaivém contínuo entre análises científicas e técnicas, por um lado, e a determinação dos fins, por outro. Tal atitude dá lugar ao debate político e, inclusive, à contribuição de especialistas sobre os assuntos. É essencial ao desenvolvimento da democracia numa sociedade tecnológica.” (G. Fourez, comunicação pessoal).

Estas diferentes atitudes não se baseiam na utilização ou no desprezo pelas tecnologias. Pode haver apoiantes do debate democrático, a todos os níveis, que aceitem plenamente tudo o que as tecnociências lhes oferecem. Assim como pode haver tecnocratas muito preocupados com as consequências humanas e sociais das decisões. A diferença reside, essencialmente, no lugar reservado aos peritos. Num perspectiva tecnocrática, a sua prática é considerada como socialmente neutra, enquanto que no modelo pragmático a palavra dos especialistas aparece enquadrada num debate de sociedade.

Devido à sua componente tecnológica, a pilotagem do ensino deve, como as outras tecnologias, ser objecto de uma *avaliação na perspectiva da sociedade*. Esta avaliação diz respeito à importância, às incidências e às vantagens relativas de determinado progresso técnico, assim como às consequências, a curto e a longo prazo, na economia, na política e, duma maneira geral, na sociedade global.

“A avaliação das opções tecnológicas” escreve Vinck (1988:7) “implica ter em conta as repercussões e riscos para a sociedade”. Por exemplo, em paralelo com os grandes serviços que nos prestam, os automóveis poluem e fazem com que se abandone o hábito de andar a pé, muito útil à saúde. “Mas”, continua Vinck, “há outro risco mais subtil: o de impor, progressivamente, uma forma de organização intolerável para determinados grupos.” (Sobre a avaliação das opções tecnológicas, na perspectiva da sociedade, ver o Anexo).

3. MACROPILOTAGEM – MICROPILOTAGEM

3.1. Macropilotagem

Qualquer poder organizador de ensino atribui objectivos ao sistema de que está encarregado e de que é responsável. Em democracia esta opção resulta de uma concertação entre todas as partes interessadas, concertação que se apoia, ou acaba por resultar, numa adesão aos valores que fundamentam as decisões. Os objectivos dizem respeito ao saber, ao saber fazer e ao saber ser.

Que haja um esforço para verificar em que medida os objectivos são atingidos, e para tomar a todos os níveis as medidas adequadas para que tal aconteça, é não só um acto racional mas também um dever cívico: a comunidade tem direito a que lhe prestemos contas dos actos cujo sucesso condiciona profundamente o futuro dos jovens e da sociedade na sua globalidade. Se surgirem carências ou insuficiências devem procurar-se as causas e devem aplicar-se os remédios convenientes.

Esta fase geral faz parte da pilotagem macroscópica. Tende a focalizar-se sobre o *input* e o *output* e é por isso, muitas vezes, denominada *tecnocrática*. Quer com isto dizer-se, que a concepção técnica dos problemas educativos corre o risco de fazer-se em detrimento das consequências humanas e sociais.

Uma acusação destas pode ser injustificada. Por exemplo, não é essencial determinar a percentagem da população escolar que adquiriu uma satisfatória capacidade de leitura sabendo, concretamente, que o nível médio de legibilidade dos textos que aparecem na imprensa diária ou nas empresas é *x* ou *y*. Se os jovens não atingirem esta capacidade de leitura têm poucas hipóteses de encontrar um emprego, as empresas ficarão carenciadas de mão-de-obra qualificada e os cidadãos não ficarão suficientemente habilitados para desempenharem plenamente o seu papel na sociedade.

A problemática aparece assim formulada de uma maneira técnica. O que não é razão suficiente para a condenar. Apesar da complexidade de alguns aspectos avaliativos eles fazem parte integrante da missão de educar. É importante, porém, que os educadores tenham sempre uma palavra a dizer sobre a concepção e execução das operações e sobre a interpretação dos resultados.

3.2. Micropilotagem

Inserido no sistema educativo global, cada local de educação possui a sua especificidade, a sua personalidade, o seu meio ambiente. Inserido no conjunto do corpo docente, cada professor é uma pessoa única, assim como cada aluno é

único entre os seus pares. Estamos assim perante a célebre frase de Popham: "A qualidade de aprendizagem que procuramos alcançar em determinada situação de ensino é função dum método *particular*, usado por um professor *particular*, para alunos *particulares* que tentam atingir um fim *particular*".

Nesta situação única, o aluno é – dada a sua liberdade – considerado sujeito, pelas mesmas razões que o é o educador, e devido à sua própria coerência pessoal, o professor – como os seus alunos – articula, constantemente, os seus princípios teóricos com o modo como os vive na sua existência concreta (Soëtard, 1983:127).

É partindo desta base, que V. de Landsheere (1992:3) escreve: "Se se admitir como primeiro princípio da educação o respeito pela pessoa, então, nem as leis psicológicas do desenvolvimento da inteligência e da personalidade, nem as leis da sociologia, nem a lógica interna dos vários ramos do saber resistem, duma maneira geral, uma vez franqueada a porta da aula. Pelo menos ajudam o educador a formular a sua própria lei *hic et nunc*". Será que o reconhecimento desta realidade implica a condenação implacável da pilotagem macroscópica? Converteza que não.

3.3. Uma complementaridade

A pilotagem microscópica ao nível da aula e até do aluno, não é incompatível com um projecto educativo global e, muito menos, com um projecto de escola. Trata-se na realidade de um meio cada vez mais subtil. Se não for possível é porque existe qualquer defeito ou incoerência ao nível de um dos escalões do projecto educativo, pelo menos.

É evidente que o método de avaliação irá sendo diferente, de acordo com os níveis, pois vai-se tornando cada vez mais qualitativo à medida que se aperfeiçoa a observação. As duas afirmações seguintes estão, de facto, muito distantes uma da outra:

- Segundo os critérios adoptados pela Associação Nacional para o Ensino da Leitura, só 32% dos alunos com catorze anos, podem ser considerados bons leitores.
- Aos catorze anos, Paulo apenas consegue ler pequenos textos com frases simples que contenham informações puramente factuais.

É claro que, para cumprir a sua missão, o professor tem necessidade de informações como as que são veiculadas pela segunda afirmação. Mas é também evidente que é à avaliação como a da primeira afirmação que recorrerá um

ministro da educação, quando tiver de tomar uma decisão sobre programas de ensino ou sobre a formação de professores.

Os aspectos comuns das perspectivas assentam na vontade de objectividade, racionalidade e eficácia da acção educativa, a todos os níveis, o que não implica nem mecanização do processo, nem recusa da subjectividade, da intuição e da construção individual do sentido que cada qual confere, ou julga conferir, às suas acções. A preocupação dominante deve ser, sempre, prestar o melhor serviço possível aos educandos.

Os relatórios de pilotagem devem servir para esclarecer adequadamente o debate pedagógico e político em matéria de acção escolar. "Se não se articularem estas duas dimensões, temo que caminhemos para o fracasso ou, pelo menos, para a não-pertinência: o debate político que não tenha em conta o contributo dos pedagogos (assim como o dos economistas ou outros), permanece nas nuvens; e o mesmo acontecerá com o debate pedagógico, se não se articular com os modos como a sociedade negoceia as suas prioridades e a articulação das suas instituições" (G. Fourez, comunicação pessoal).

E assim como a investigação, por si só, não garante o desenvolvimento, assim também é necessário que haja um entendimento correcto dos serviços que a pilotagem pode prestar. É ingenuidade pensar que esta compreensão se vai dar de um dia para o outro. Será igualmente ingênuo esperar que os responsáveis pelas decisões utilizem todos os dados postos ao seu alcance, mesmo os mais seguros.

4. SOBRE O VALOR ACRESCENTADO

Uma avaliação do *valor acrescentado* consiste na interpretação dos resultados escolares (em sentido lato), em função do que os alunos sabiam, ou de como eles estavam ao entrar na escola ou na turma. Pode, também, ter-se em consideração as diferenças de atitude, bem como a qualidade do apoio prestado pelas famílias. É mais fácil de dizer do que de fazer. É claramente impossível provar que para determinada disciplina passar dum nível de sucesso de 50 para 60% equivale a passar de 70 para 80%, ou de 85 para 90%...

A solução adoptada na Califórnia é a seguinte: todas as escolas públicas são submetidas a um mesmo teste e os resultados duma escola particular são comparados aos resultados obtidos pelo conjunto de 10% das escolas, cujo índice sócio-económico estiver mais próximo do índice da escola em questão. É evidente que um índice desta natureza não explica tudo; constitui, contudo, um ponto de referência importante.

Uma inovação recente da Direcção de Avaliação e Prospectiva (DEP) do Ministério de Educação Nacional, de França, insere-se nesta mesma perspectiva: as taxas de acesso regionais ao diploma de conclusão do ensino secundário, em 1991, foram comparadas com "uma taxa de acesso prevista, tendo em conta a estrutura social de cada região" (DEP, *L'état de l'école*, Outubro de 1992:23).

Sem subestimar o interesse deste tipo de comparações, não se pode contudo ignorar como é difícil de apreender a realidade da "escola eficaz", mesmo se limitarmos o âmbito dessa eficácia às aprendizagens escolares tradicionais.

Liensol e Meuret (1987), citados por A. Grisay (1988:5), estudaram em duzentos e dez estabelecimentos de ensino secundário das academias de Grenoble, Metz-Nancy e Toulouse, ao longo de três anos consecutivos, os resultados alcançados pelos alunos no final do ensino secundário. A influência da origem sócio-económica dos alunos e dos seus atrasos escolares anteriores, é evidente.

Contudo, num mesmo estabelecimento de ensino, a taxa de sucesso está muito longe de ser estável. "Um liceu em três, pode apresentar resultados radicalmente diferentes dum ano para o outro. O mesmo se passa quando se avaliam os valores da taxa de sucesso na obtenção do diploma de conclusão do ensino secundário." Importa, pois, definir de maneira mais rigorosa o conceito de eficácia ou de rendimento das escolas.

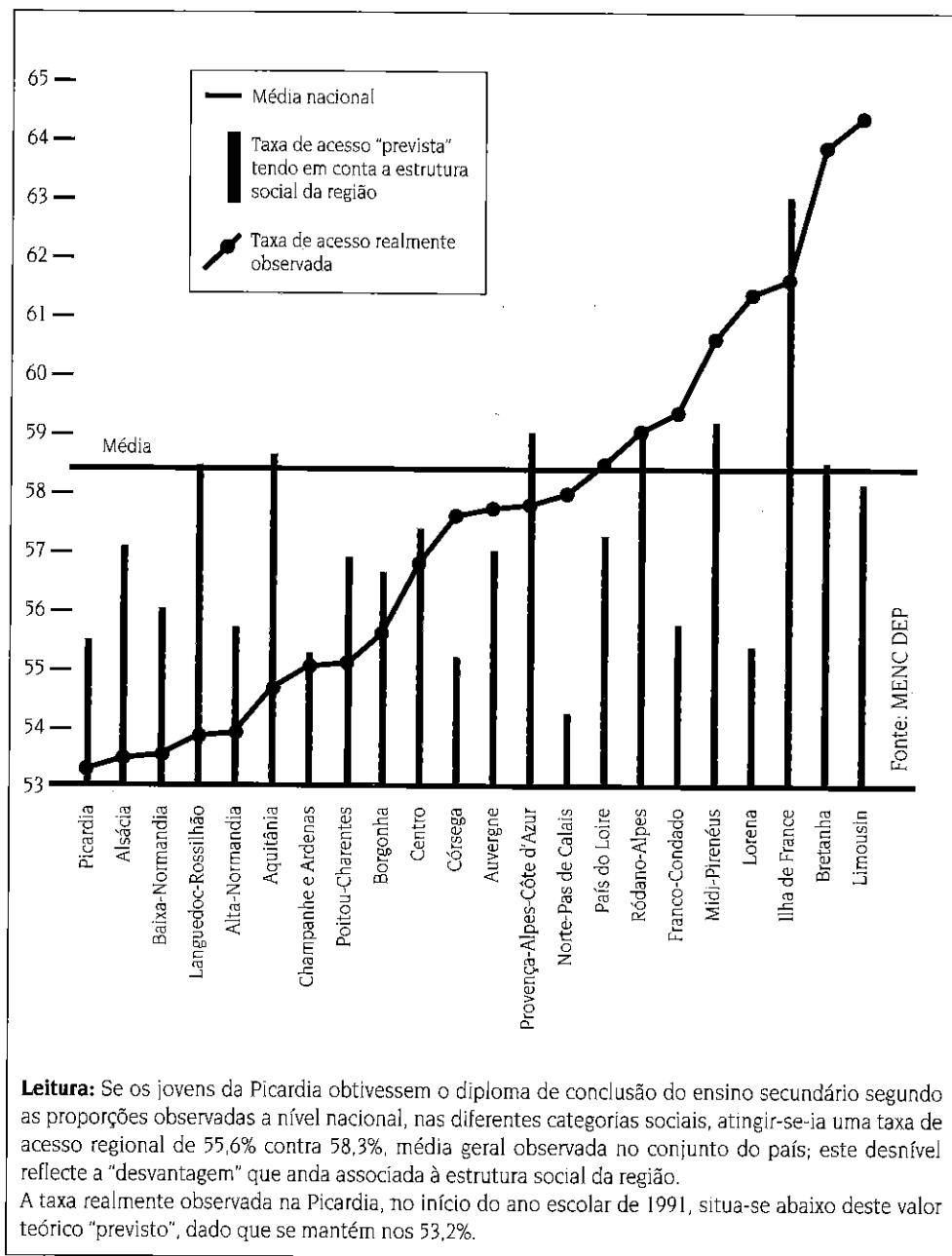
O balanço que A. Grisay faz das investigações quantitativas, que se inscrevem na linha do Relatório Coleman (investigações sobre a relação entre *input* e *output*) e também as investigações mais qualitativas relacionadas com o *modelo ecológico*, revela bem a dificuldade da tarefa e a pobreza de conhecimentos seguros de que dispomos actualmente sobre o assunto.

O que torna exasperantes os resultados das investigações quantitativas *input-output* sobre rendimento escolar, conclui A. Grisay (p.17), é o facto de "as diferenças de rendimento entre escolas (que nada têm a ver com o recrutamento e são, pois, em princípio, devidas à qualidade do meio envolvente escolar) existem sem que nos possamos aperceber delas (ou muito raramente) através de diferenças objectiváveis de 'tratamentos' seguidos pelos alunos. De facto as observações provam que: *há escolas que conseguem melhores resultados do que outras.*"

As conclusões gerais deste primeiro tipo de investigações são (p.36):

- A "boa escola" não se reconhece logo à primeira vista. Há que ter cuidado, sobretudo, em não confundir "boas escolas" e escolas "bem frequentadas". Apesar dos resultados elevados, estas últimas não são, necessariamente, as melhores.
- O bom rendimento duma escola anda sobretudo associado ao modo como é explorado o tempo de aprendizagem.

Figura 1
Taxas regionais de acesso ao diploma de conclusão
do ensino secundário (1991)
Efeito da estrutura social



– “Não é clara a existência dum “perfil” de escola com alto rendimento válido sempre e em toda a parte. A realidade é mais complexa: existem, provavelmente, características escolares que dão bons resultados com *certas* populações escolares e com *certas* disciplinas e que dão resultados menos bons com outros alunos e outros domínios do saber...”

O balanço das investigações de tipo ecológico também não é mais esclarecedor.

O processo geral consiste principalmente numa observação etnometodológica de estabelecimentos de ensino, considerados tanto de alto como de fraco rendimento: tenta-se então descobrir, através duma observação profunda e sistemática, os processos (sociais, organizacionais, pedagógicos,...) que podem explicar a eficiência ou a não eficiência.

Factores como a qualidade da disciplina, as expectativas positivas em relação aos alunos, a capacidade de liderança do director do estabelecimento estão, claramente, entre os mais importantes. Mas há que ser prudente nas conclusões a retirar daqui: “Os resultados actuais ainda não nos dizem o que são escolas de alto rendimento. Dão-nos, apenas, um determinado número de indicações sobre os factores associados a uma subida da taxa de sucesso em testes estandardizados” (p.57). Há, pois, a possibilidade de pôr em causa a própria validade dos critérios, dado que não só se devem explorar melhor os efeitos cognitivos como também se deve dar ao aspecto afectivo muito mais atenção, do que aquela que se lhe dá actualmente.

A conclusão geral vem ao encontro daquela a que se chegou a propósito da investigações *input-output*: há que explorar de maneira mais aprofundada as razões que levam *certo* estilo de escolas a obterem mais rendimento do que outras.

Confirma-se, pois, uma vez mais, a imensa complexidade da avaliação dos efeitos da escola. Perante semelhante dificuldade, ou se baixam os braços e se renuncia a qualquer esforço de objectivação, ou se dá provas de possuir a “coragem da imperfeição”: nas ciências humanas a complexidade e diversidade são tais que só mediante grandes aproximações se pode trabalhar. O importante é estar, permanentemente, consciente disso e, do pouco que se conhece, retirar o proveito máximo e o maior número de ensinamentos.

5. COERÊNCIA, QUALIDADE, EQUIDADE

Pode-se tentar controlar a produtividade dum sistema educativo no âmbito do saber, do saber-fazer e do saber-ser, quer independentemente dos objectivos

do programa (faz-se o apanhado de todas as aquisições que for possível identificar), quer tendo em conta os objectivos principais que o ensino se propôs atingir.

É antes de mais nesta segunda hipótese que se fala de pilotagem. Pretende-se saber se se está a avançar na direcção escolhida e, para tal, dá-se maior importância a indicadores que poderão tornar-se, eventualmente, factores de vacilação, no caso de surgirem problemas particulares. A repetição periódica das avaliações, permite distinguir evoluções e tendências.

A pilotagem deve determinar em que medida o sistema educativo dá resposta a **três exigências: coerência, qualidade e equidade**.

Existe *coerência*:

- se os *objectivos* a alcançar se inscrevem, efectivamente, quer na linha do projecto educativo global, democraticamente formulado e adoptado (macro-pilotagem), quer na linha de projectos particulares formulados por diversos grupos ou até mesmo projectos individuais (micropilotagem);
- se o *currículo implantado* – isto é, aquilo que efectivamente se ensina e aprende – corresponde aos objectivos fixados;
- se a natureza e modalidades das *avaliações* estão em função destes mesmos objectivos.

A *qualidade* pode ser avaliada de maneira absoluta (Permite a educação um desenvolvimento excelente e o mais favorável de cada indivíduo?), ou relativa (Estão os nossos alunos a ser tão bem formados como os outros?).

O critério de *equidade* exige, em especial, que todos os alunos tenham acesso de forma igualitária aos recursos disponíveis o que não significa, de modo nenhum, que devam tirar partido desses recursos da mesma maneira e na mesma quantidade (*discriminação positiva*). Convém, sobretudo, determinar se a qualidade e nível das aquisições não terão mais a ver com o estabelecimento de ensino frequentado, do que com as reais capacidades daqueles que o frequentam. Em democracia os sistemas educativos têm como ideal a igualdade de sucesso que pode revestir várias modalidades conforme as aptidões, as necessidades e os interesses de cada um.

Por exemplo, um dos aspectos da equidade a garantir é a igualdade entre sexos. As investigações avaliativas sobre este assunto abordaram especialmente os seguintes aspectos:

- Tendências gerais dos resultados alcançados pelas raparigas e comparação com os obtidos pelos rapazes.

- Comparação dos resultados obtidos por rapazes e raparigas em francês e matemática.
- Diferenças na escolha dos cursos.
- Orientação profissional das raparigas.
- Percentagem de raparigas e de rapazes que prosseguem estudos, após o secundário.
- Diferenças de expectativas.

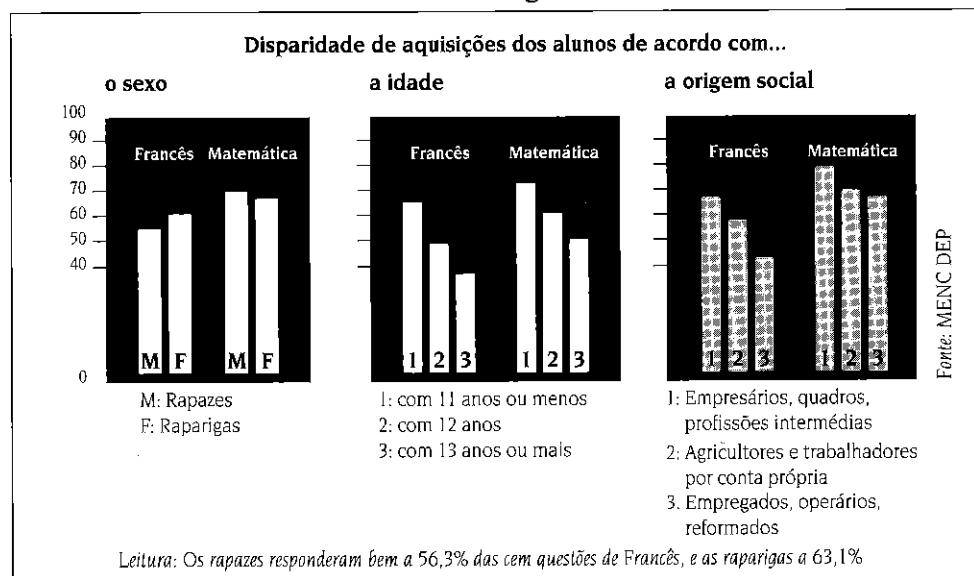
Em *L'état de l'école* (2, 1992:39) a Direcção de Avaliação e Prospectiva do Ministério da Educação Francês dá sobre este assunto as seguintes informações:

Figura 2 a
Percentagem de sucesso nas questões de Francês e Matemática
no início do sexto ano de escolaridade (Setembro de 1991)

	No seu conjunto	10% dos alunos mais fracos	10% dos melhores alunos
Francês	60,6	24,4	85,7
Matemática	70,1	35,9	89,7

Leitura: em média, os alunos responderam correctamente a 60,6% das questões de Francês

Figura 2 b
Disparidade de aquisições dos alunos de acordo com o sexo,
a idade e a origem social



Há, pois, muitas questões a exigirem respostas claras e bem fundamentadas. Que aprendem realmente os alunos? Há equidade no sistema educativo? Os programas correspondem às exigências do mundo contemporâneo e do futuro? Os métodos de ensino são adequados? São estas as questões evocadas no quadro seguinte.

5.1 Recapitulação das questões

	Coerência	Qualidade absoluta	Qualidade relativa	Equidade
OBJECTIVOS DOS PROGRAMAS	O ensino dá resposta às necessidades dos alunos?	Os programas são suficientemente exigentes, mas sem exageros?	São os alunos tão bem formados como os de outros países?	Os programas estão concebidos de modo a assegurarem a igualdade de oportunidades e as discriminações positivas?
PRÁTICAS EFECTIVAS NAS AULAS	Correspondem aos objectivos a atingir?	O currículo adoptado é válido?	Dentre os vários métodos, quais se revelaram mais eficazes? Que lições retirar do que se passa na aula?	Não estará o ensino orientado em função das origens sociais ou étnicas, do sexo, da implantação geográfica?
AQUISIÇÕES DOS ALUNOS	O perfil de competências obtido é o que se pretendia?	As aquisições são suficientes para garantir o desenvolvimento pessoal, o exercício da cidadania responsável?	Em que situação se encontra: – o aluno médio – o aluno de um escalão superior, em relação aos dos outros países?	Verificam-se diferenças importantes entre as categorias acima indicadas? Quais as taxas de insucesso?
EVOLUÇÃO	Apercebemo-nos dela?	O nível de rendimento baixa?	O sistema evolui melhor ou pior do que o dos outros?	O sistema evolui para uma maior democratização?

Segundo D.H. Hargreaves (1988) a pilotagem só terá sucesso se estiver de acordo com as quatro condições seguintes:

1. Ser credível, não só para quem ensina mas também para os outros membros da comunidade educativa.

2. Respeitar o anonimato quando as circunstâncias o exigirem. Daí que os responsáveis pelas operações devam gozar de certo grau de independência em relação às autoridades.
3. Combinar duas funções: a da pilotagem e a do desenvolvimento. De facto, não basta estabelecer contactos, há que partir para uma melhoria real do sistema.
4. Entusiasmar e apoiar os professores e os responsáveis pela direcção das escolas.

6. DEFINIR E CHEGAR A ACORDO SOBRE AS FINALIDADES DA EDUCAÇÃO

Todas as considerações precedentes assentam numa condição primordial: definição clara de finalidades e objectivos e acordo sobre esta matéria.

Por exemplo, nos Estados Unidos, as agências a quem se confiaram as operações de pilotagem tentaram, em vão, pôr de acordo sobre as finalidades da educação, professores, especialistas das várias disciplinas, representantes dos pais e parceiros sociais. Por vezes, os responsáveis pela pilotagem chegaram mesmo a ser acusados de quererem forjar um consenso que não existia de facto.

Tais divergências existem, mais ou menos, por todo o lado. Estão relacionadas com os fins específicos mas também com as finalidades gerais das disciplinas: a matemática deve ser estudada para resolver os problemas da vida quotidiana ou pela ciência em si? Qual a importância relativa das aprendizagens propostas pela escola? Estamos de acordo num mínimo de desacordo?, questiona-se V. de Landsheere (1988:231).

Este fenómeno não deve ser considerado como anormal ou como uma fraqueza. Todos os seres humanos são diferentes. Excepto quanto às necessidades primárias ligadas à sobrevivência, não adoptam os mesmos valores, não sentem as mesmas necessidades, o que provoca uma grande variedade de objectivos a atingir em determinado momento.

Tal facto deve despertar a nossa atenção mas não deve, contudo, impedir a acção educativa colectiva que, apesar dos nossos esforços, continuará inevitavelmente desadaptada se continuarmos a pensar em termos absolutos. Importa é tentar orientar a pilotagem para uma espécie de máximo denominador comum cultural. E estar, portanto, bem consciente dos limites da tarefa. É o que se passa com todas as acções importantes que impliquem a participação do homem. O denominador comum deve buscar-se, prioritariamente, ao nível dos valores.

Uma coisa é certa, não cabe apenas à escola fixar os fins e objectivos da educação. É à comunidade educativa na sua globalidade – e em última análise esta confunde-se com a sociedade – que incumbe determiná-los, através de debates *realmente* democráticos. J.-E. Charlier (comunicação pessoal) conclui que a questão principal se situa a montante da operacionalização da pilotagem; antes de mais, há que ter capacidade para elaborar um projecto coerente para o ensino. A experiência mostra que esta tarefa está muito longe de ser fácil. De facto, as condições económicas, políticas, culturais e sociais actuais tornam praticamente impossível a emergência dum projecto educativo aceite plenamente por todos. Há, contudo, que desconfiar duma espécie de perfeccionismo que nos impede de avançar e dar aceitação a acções, mesmo que apenas parcialmente satisfatórias, desde que se inscrevam na direcção certa.

Num plano mais técnico observa-se que a formulação de algumas finalidades se refere a determinado critério externo de êxito. Por exemplo, uma das seis finalidades que o presidente Bush queria que o seu país atingisse era: "No ano 2000, os resultados dos alunos dos Estados Unidos em ciências e matemática devem ser os melhores do mundo." Outro exemplo, desta vez na Inglaterra: "A mão-de-obra britânica deve possuir conhecimentos e competências que a tornem competitiva nos mercados europeus e mundiais" (*National Commission of Education*, 1990).

Como nota T.N. Postlethwaite (1993), "No passado, as finalidades da educação eram formuladas num contexto nacional, sem referência comparativa explícita a padrões nacionais ou internacionais. (...) Em contrapartida, desde meados dos anos oitenta, verifica-se uma tendência para formular as finalidades em termos de alvos a atingir". O mesmo se passa nas empresas.

7. DESENVOLVER OS SERVIÇOS DE ESTATÍSTICA

Para fazer a pilotagem do sistema educativo e prestar contas ao país é indispensável a recolha regular de dados e de estatísticas a nível local, regional, nacional e internacional. Tais estatísticas devem abranger não só os recursos e duma maneira geral tudo o que entra no sistema, mas também a maneira como o sistema funciona (processo) e o que ele produz.

Praticamente todos os países dispõem de serviços de estatística da educação, mais ou menos desenvolvidos. Contudo, os dados recolhidos variam consideravelmente tanto de natureza como de qualidade. Assim se explicam, por exemplo, as dificuldades sentidas pelas grandes organizações internacionais, como a UNESCO ou a OCDE, para conseguirem estatísticas suficientes, fiáveis e

comparáveis. São actualmente desenvolvidos enormes esforços para melhorar esta situação.

8. PILOTAGEM DA PILOTAGEM

A pilotagem é uma técnica: trata-se, de facto, de um procedimento aplicado sistematicamente, para obter um resultado determinado, neste caso a optimização da acção educativa.

A própria concepção de óptimo levanta uma questão crucial: quem o define e como? Autocraticamente ou democraticamente? Há que afirmar, desde já, que optamos decididamente pelo segundo processo o que implica o domínio social da pilotagem.

Para fazer a pilotagem de um sistema educativo há que mantê-lo na trajectória e, portanto, reduzir a sua liberdade submetendo-o a uma norma que permita determinar se se afasta do caminho traçado ou se a finalidade é de facto atingida.

Contudo os padrões e normas podem ser redutores e travar a inovação. E isto é especialmente verdade no campo da educação dado que a aprendizagem de um mesmo saber, saber-fazer ou saber-ser, se pode fazer por caminhos e modalidades muito diferentes. Além disso, se a pilotagem for utilizada para levar, de bom ou de mau grado, à obtenção de um objectivo autoritariamente fixado, então a norma e o padrão tornam-se "uma arma redutora, um instrumento de vassalagem e de exclusão" (Gonod, 1992:8). Daí que as normas e padrões, também devessem, sempre, ser objecto de negociação social.

Ora, como veremos, a realização de uma pilotagem – sobretudo na perspectiva macroscópica – levanta problemas de grande tecnicidade perante os quais muitos professores se sentem desarmados. "A democratização do controlo postula uma democracia participativa associada à emergência de uma cultura técnica de massas, mediante a democratização dos conhecimentos e da informação tecnológica" (Gonod:8). Vemo-nos assim, uma vez mais, perante a necessidade de elevar o nível científico da formação de professores.

Por outro lado, a política e a prática da pilotagem devem ser objecto de permanente discussão entre todas as partes interessadas. Em anexo é feita referência aos diferentes métodos de prevenção dos perigos que podem advir dum a aplicação descontrolada das tecnologias e aos métodos de procura de consenso neste campo.

CAPÍTULO 2 – INDICADORES

1. NOÇÃO DE INDICADORES

Devido ao uso considerável que a pilotagem faz dos indicadores, há necessidade de definir com precisão aquilo de que estamos a falar. V. de Landsheere (1992, p.478) escreve:

"Sendo uma medida destinada a servir a pilotagem, um indicador é uma estatística directa e válida que dá informações sobre o estado e as mudanças de grandeza e natureza, ao longo do tempo, de um fenómeno social tido como importante. Em educação informa, de uma maneira especial, sobre a saúde e qualidade de funcionamento do sistema, sobre os conhecimentos dos alunos, sobre as suas capacidades, sobre a evolução positiva ou negativa destas condições (daí a atenção a dar à observação de características duráveis), e sobre as diferenças significativas que podem existir, em determinado momento, entre áreas geográficas ou instituições. Um sistema de indicadores deverá também dar informações sobre o modo como os indicadores medidos conjugam a sua influência, de modo a dar origem à situação e à evolução do sistema educativo."

Um *dado* é uma informação qualitativa ou quantitativa, tratada estatisticamente e útil à tomada de decisões. Uma *estatística* é um dado quantitativo. Os indicadores são estatísticas (resultados de testes, taxas de abandono escolar,...). Os sistemas de pilotagem podem utilizar dados qualitativos e estatísticas.

Instalar um sistema de indicadores da educação é uma tarefa ambiciosa. Não estaremos a exigir-lhe mais do que pode dar, no estado actual dos nossos conhecimentos? De facto, estamos muito longe de saber com segurança o modo como as componentes decisivas de um macrosistema ou de um microsistema educativo, combinam esforços para produzirem os resultados observados. Contudo, isso não é razão suficiente para nada fazer e, sobretudo, para deixar de desenvolver sistemas de indicadores que ajudem a descrever e, se possível, a explicar, nem que seja só em parte, este fenómeno. É também importante que, em paralelo com este esforço, se desenvolvam investigações quer sobre o modo de produzir indicadores individuais, como por exemplo a medida dos resultados intelectuais de alunos a quem foram cometidas tarefas de resolução de problemas, quer sobre a natureza das principais componentes do sistema educativo que conjugam os seus esforços para produzirem o estado actual do mesmo sistema (Odden, 1990:34).

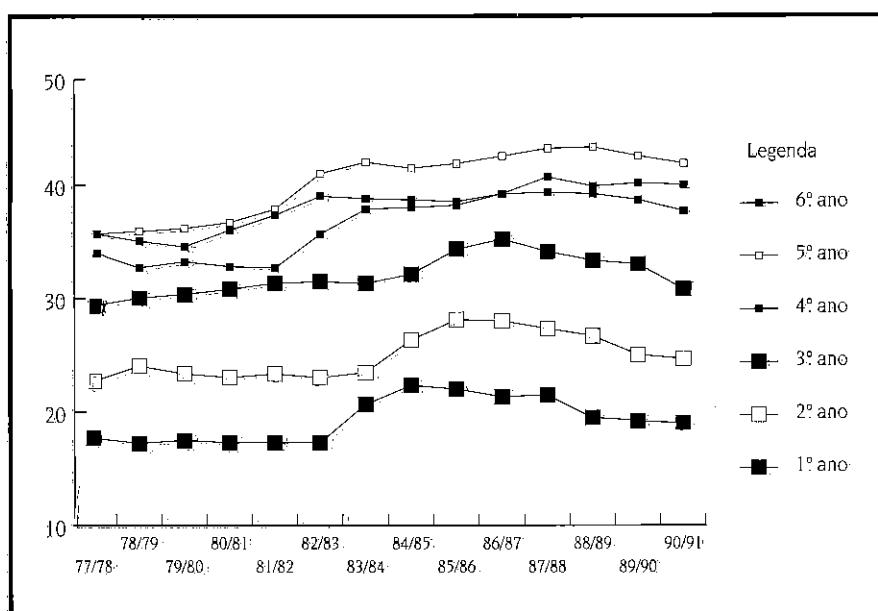
Em boas condições, os indicadores são o reflexo da política educativa e dão, eventualmente, informações sobre modificações a operar nessa mesma política.

Devem ser consistentes, isto é, não devem ser demasiado sensíveis às flutuações accidentais. Têm, pois, de referir-se a características do sistema educativo que se manifestem de uma maneira geral e duradoura. Se o seu significado for facilmente compreendido por todos os membros da comunidade educativa, fica facilitado o diálogo entre políticos, cidadãos e professores.

Um indicador pode servir de sinal de alerta para procurarmos as causas. Desperta a atenção e convida à ação (pilotagem). É, também, um instrumento de comunicação.

É interessante acompanhar, por exemplo, certos fenómenos educativos ao longo de certo período de tempo e tentar compreender as profundas modificações que se vão dando. Por exemplo, o gráfico seguinte mostra a evolução da percentagem de insucessos escolares, durante um período de catorze anos, no ensino secundário, entre a comunidade flamenga da Bélgica (De Groof *et al.*, 1993:115)¹⁰.

Figura 3
Evolução da percentagem de alunos repetentes no ensino secundário
1977/78 – 1990/91



(1) Para melhor compreensão do sistema educativo francês (e belga) veja-se, mais adiante, o capítulo 3. (N.T.).

O gráfico revela duas tendências. Por um lado, verifica-se que a percentagem de repetências aumenta no quinto e sexto ano, sobretudo a partir de 1981-82. É, provavelmente, devido ao facto de menos alunos abandonarem prematuramente o ensino. Por outro lado, a partir de 1983-84 verifica-se um aumento de insucesso no primeiro ano; a situação melhora um pouco nos anos seguintes mas torna a deteriorar-se novamente. Tais fenómenos podem atribuir-se ao prolongamento da escolaridade obrigatória.

Nem toda a estatística relativa à educação vale como indicador. Assim, por exemplo, conhecer o número de escolas de uma região não nos diz nada sobre a qualidade do que aí se faz. Um número só adquire significado se puder ser interpretado em relação a um ponto de referência: evolução de uma mesma estatística, comparação com medições semelhantes feitas noutras locais, normas definidas para a sociedade,... O mais importante é dispor de meios para ser correctamente informado sobre a qualidade da educação.

Conceber indicadores é, pois, uma operação em que intervêm considerações filosóficas (juízos de valor), políticas e técnicas. Pretende-se que sejam importantes, válidos, de fácil utilização, credíveis e que conduzam à melhoria das situações.

Exercer a pilotagem é mais do que acumular indicadores. Ela tem, necessariamente, três componentes: recolha regular de informações, avaliação destas informações e, sobretudo, a tradução em acções institucionais ou em sanções, isto é, em aprovações ou desaprovações. Pode-se, por exemplo, chegar à conclusão que determinado objectivo foi ou não atingido de maneira satisfatória. Em resumo, enquanto as estatísticas que se publicam habitualmente sobre educação fornecem informações quantitativas (orçamento, número de diplomados, etc.), os indicadores ajudam a perceber melhor o funcionamento qualitativo dos sistemas e os resultados por estes alcançados.

Segundo Oakes (1986:496), os indicadores devem fornecer pelo menos, uma das informações seguintes:

- Descrição das qualidades do sistema educativo que levem à consecução dos objectivos a atingir.
- Informação sobre as características relacionadas com os resultados alcançados (ponto estabelecido pela investigação experimental).
- Descrição das características fundamentais do sistema. Deve conduzir a uma melhor compreensão do funcionamento.
- Informação sobre problemas que se venham a colocar.

- Informações importantes para a política educativa, especialmente para a decisão de reformas.
- Descrição de características gerais, isto é, que podem ser observadas, de uma forma ou de outra em todo o sistema, e que podem ser comparadas em diversos contextos.
- Informações relativas a características permanentes do conjunto, ou de subconjuntos do sistema, formuladas de modo a permitir a comparação ao longo do tempo.

Por seu lado, Kaagan e Smith (1985) sublinham o importante papel que os indicadores podem desempenhar na pilotagem das reformas:

- Indicar mudanças de aspecto importantes como a qualidade do ensino e os resultados obtidos pelos alunos e, assim, chamar a atenção para eventuais problemas.
- Avaliar a influência dos esforços de reforma.
- Desenvolver esforços relativamente a outros países ou entidades políticas.
- Chamar a atenção para estabelecimentos de ensino que devem ser objecto de melhorias.
- Servir de catalisadores de novas ideias.

Para cobrir o sistema educativo as informações devem abranger:

- O que entra no sistema (*input*): recursos, qualidade dos alunos, dos professores, etc..
- Os processos educativos (*treatment*): pedagogia praticada, modalidades de exames, etc..
- Os “produtos”, o que sai do sistema (*output*): aquisições escolares, número de diplomados, etc..

Estas três componentes mantêm relações estreitas e multidireccionais. Importa pois examinar como é que os indicadores se associam para explicar o efeito global.

Nesta matéria, como em todas as que digam respeito ao homem, não estamos a lidar com uma ciência integralmente exacta. Os indicadores baseiam-se quer em conhecimentos científicos, quer no bom senso, quer em preocupações políticas. Além disso, os indicadores não estão de modo nenhum, ao abrigo da crítica. Critica-se-lhes, muitas vezes, o facto de apenas apresentarem um conjunto de informações díspares. O exemplo dos indicadores económicos mostra,

contudo, como eles podem ser úteis, apesar desta imperfeição. As falhas foram ainda muito mais graves no passado aquando da tentativa de definir indicadores sociais.

Com efeito, muito antes do movimento relacionado com os indicadores da educação ter adquirido a importância que tem actualmente, **os indicadores sociais** destinados a avaliar os efeitos da prosperidade económica dos anos sessenta sobre a qualidade de vida e sobre o progresso da equidade, suscitarão um interesse considerável. Pretendia-se fazer a pilotagem da evolução sócio-económica. A obra de R. Bauer, *Social Indicators* (1966) é, a este respeito, a referência principal. A OCDE muito cedo se interessou por este movimento na esperança de medir o efeito da educação na sociedade. Era a grande época da investigação quantitativa em sociologia, notavelmente retratada por R. Boudon (A Matemática em Sociologia, 1971) e por J. S. Coleman e seus pares no célebre relatório (1966) sobre a igualdade de oportunidades na educação.

Ao entusiasmo sucedeu rapidamente a desilusão. Aspectos particulares de muitas situações e efeitos desapareciam sob o efeito de indicadores nacionais globais, o que em nada ajudava a encontrar as soluções pretendidas. Por outro lado, as relações causais mantinham-se muito ao nível da correlação e não da compreensão. Além disso, estava-se muito longe de ter conseguido um acordo sobre a definição dos principais conceitos. Não havia planificação nem decisões que pudessem apoiar-se em bases assim tão pouco seguras.

Da experiência adquirida durante esta primeira época, T.J. Wyatt (1988:5) retirou cinco grandes lições:

1. Os indicadores devem ser definidos de maneira simples e operatória.
2. Há que evitar misturar medidas díspares num indicador global que prenda representar o produto da educação.
3. Os indicadores podem revelar problemas mas não fornecem soluções.
4. A interpretação dos indicadores é pelo menos tão importante como a sua elaboração. Para se fazer uma ideia de conjunto de uma situação há que aprender a articular os indicadores observando as tendências de cada um ao longo do tempo.
5. É importante que os indicadores ajudem a formular a política educativa.

Além disso, pode temer-se que os indicadores exerçam um efeito de refluxo sobre o ensino. Já aludimos a este fenómeno. Basta-nos, pois, recordar com Porter (1988:504), e novamente a título de exemplo, que: "Se existem bons indicadores relativamente ao ensino da matemática e das ciências e não para as disciplinas de letras é porque as escolas terão, provavelmente, a tendência a dar

mais importância à matemática e às ciências do que ao resto. Se os indicadores dizem respeito a competências consideradas separadamente e ao conhecimento dos factos, mais do que à compreensão profunda e à resolução dos problemas, então a escola corre o risco de se orientar neste mesmo sentido."

De facto, os indicadores podem tornar-se algo mais do que informações sobre a saúde do sistema escolar; podem tornar-se definições do que é essa boa saúde.

Compete à comunidade educativa velar para que não se cometam tais erros.

2. NECESSIDADE DE UM MODELO

É necessário existir um modelo coerente de funcionamento do sistema educativo para decidir quais os indicadores a eleger. A escolha deste modelo determina, em grande parte, as informações que o sistema educativo irá fornecer. No capítulo anterior utilizou-se o modelo sistémico *entrada - tratamento - saída - feedback*. O modelo geral adoptado pela OCDE (CERI), inscreve-se nesta linha: num determinado meio, estudam-se os *recursos*, os *processos* e os *efeitos*.

Os modelos estão sempre orientados de acordo com uma perspectiva epistemológica particular (económica, sociológica,...) do sistema educativo. Assim, se explica que não haja uma aceitação unânime de nenhum modelo geral do processo educativo.

Scheerens (1989, citado por Wyatt, 1991:10) distingue três tendências:

1. A passagem de estatísticas descritivas relativas ao *input* (recursos,...) a conjuntos de indicadores relativos aos níveis de rendimento alcançados, ou aos resultados (*output*). Esta tendência vem dar resposta às pretensões dos responsáveis pelas decisões políticas que desejam saber o que o sistema produz e desejam que as escolas sejam levadas a prestar contas, não tanto do passado, mas mais dos objectivos que atingem.
2. Tendência a privilegiar sistemas educativos mais amplos, mais centrados nos resultados e medidas de contexto, tendência acompanhada dum interesse crescente pelos factores de *input* e pelas características de processos maltráteis.

Os responsáveis pelas decisões não querem apenas saber o que se passa, mas pretendem, também, saber por que é que é assim e o que podem fazer para remediar as falhas observadas. Dá-se cada vez mais importância ao que é produzido realmente nas escolas (daí a insistência no "tratamento" e nos resultados). Há uma recusa em continuar a deixar os aspectos do

"tratamento" fechados numa "caixa negra", pretende-se descobrir-lhes os segredos, numa primeira fase sem teoria pré-existente. Trata-se, antes de mais, de abrir os olhos para a realidade dos processos.

3. Tendência a considerar a totalidade do sistema: *input* – tratamento ou processo – *output*.

Este terceiro modelo é o mais utilizado. R.R. Herriott [1979, citado por Van Herpen (1992:40)], de uma maneira muito interessante, tornou mais subtil ainda este modelo, não se concentrando apenas sobre o processo mas tendo também em conta aspectos sociológicos da estrutura e da cultura, assim como características do meio ambiente.

Herriott distingue seis componentes:

1. *Meio envolvente*. Os objectos e as formas de relação que existem a todos os níveis fora do sistema educativo. Ou este influencia o meio ou é influenciado por ele.
2. *Estrutura*. Padrões (*patterns*) relativamente estáveis de interacções sociais, que integram diversos elementos de um sistema educativo, a todos os níveis.
3. *Cultura*. Sentimentos, capacidade de captar e viver sensações, compromissos que têm a ver com a experiência de vida dos diferentes actores em presença (professores, alunos, responsáveis pela administração,...), e com os objectivos, as actividades e os problemas do sistema educativo.
4. *Output*. Conhecimentos, capacidades e orientações exigidas pelo meio envolvente e possuídas pelos alunos quando, a qualquer momento, abandonam o sistema educativo.
5. *Throughput*. "Materiais" brutos (fundamentalmente os alunos) que passam pelo sistema, e sobre os quais o sistema opera para produzir um *output* adequado ao meio.
6. *Input*. Instrumentos (manuais escolares,...), pessoal (professores, serviços, responsáveis pela administração) e informações (conhecimentos e valores), que um sistema educativo vai buscar ao meio envolvente e que utiliza para transformar o *throughput* em *output*.

3. NÚMERO IDEAL DE INDICADORES?

Não existem respostas pré-fabricadas para esta questão.

Acabamos de ver que existem múltiplos indicadores para cada aspecto importante da educação e, face a cada um deles, não é nada fácil decidir quantos

eleger. É certo que quanto menos são os indicadores mais atraem a nossa atenção, mas esta redução pode prejudicar a sua validade: o conjunto pode passar a não ser representativo do sistema avaliado. Ora, esta representatividade é necessária, pois um indicador não é apenas uma medida de si mesmo; desempenha o papel de uma variável que sintetiza um conceito mais alargado.

Dito isto, é preciso não perder de vista que os indicadores “não dizem tudo o que respeita à educação. Permitem apenas entrever as grande linhas de determinada situação. (...) O seu grau de generalidade pode dissimular situações muito diferentes existentes nos subgrupos. Eis a razão por que os indicadores devem poder ser decompostos, a fim de haver possibilidade de examinar as subpopulações.” (T. Wyatt, 1991:9).

Em última análise, a resposta à questão deve ser ditada pelo objectivo a atingir. Por exemplo, a resposta será diferente conforme se pretender fazer sensacionalismo ou se quiser avaliar uma situação de modo sereno e bem informado...

4. QUALIDADES PRETENDIDAS

Embora especialmente desejáveis nos indicadores relacionados com a escola, as qualidades que a seguir se enumeram dizem também respeito à maior parte dos outros indicadores (Gray e Jesson, 1968, citados por Nuttal, 1991):

- Avaliarem directamente os níveis de rendimento escolar.
- Estarem sobretudo atentos aos processos de ensino e aprendizagem.
- Serem quantificáveis, o que não quer dizer estritamente mensuráveis.
- Serem unidimensionais (evitando as sobreposições).
- Cobrirem os objectivos prioritários no seu conjunto.
- Tratarem dos aspectos mais importantes do sistema educativo, capazes de serem modificados por mudanças de política.
- Serem verificáveis.
- Poderem ser postos em relação com outros.
- Serem económicos tanto quanto possível. A recolha de indicadores deve poder levar-se a cabo durante um determinado período de tempo, a um preço e a um nível de especialização razoáveis.
- Serem fornecidos em tempo útil (especialmente para a tomada de decisões).
- Prestarem informações relativas a problemas.
- Serem facilmente compreendidos pelo grande público.
- A validade e fidelidade dos indicadores devem ser reconhecidas por todos.

5. DIRECTIVAS PARA UMA UTILIZAÇÃO CONSTRUTIVA

L. Darling-Hammon (1991:23 e seguintes) dão as seguintes indicações:

1. Os números não bastam. O conhecimento do ensino, da aprendizagem e do funcionamento das instituições educativas, deve fazer com que todas as partes interessadas os interpretem correctamente.
Não é por serem rigorosos que os números passam a ser válidos.
2. Só uma multiplicidade de indicadores pode dar a imagem correcta de uma situação. Os aspectos de educação, alvos directos da política educativa, não são os únicos a despertar interesse. Importa sobretudo ter em atenção as características duradouras do sistema educativo.
3. Os sistemas de indicadores devem ser redundantes e continuamente revistos. Reduz-se o perigo de corrupção de um indicador se se proceder a várias avaliações do mesmo aspecto do sistema.
4. Haja ou não indicadores, a política deve basear-se numa sa compreensão da natureza das motivações individuais e das alterações organizacionais e da influência destes factores nos níveis de rendimento.
5. Na avaliação de uma situação, os indicadores devem ser utilizados como a água em relação ao moinho e não como factos suficientes para desencadear acções que poderiam ser ineficazes. O interesse dos indicadores é emitirem sinais a propósito de questões importantes e de problemas que poderão surgir. E quando se der este caso impõe-se um exame profundo da situação.

Dito isto, há que ter cuidado em não falsificar o significado dos indicadores. Por exemplo, indicadores de níveis de rendimento pouco favoráveis não significam, *ipso facto*, que o programa seja mau.

Por outro lado, dado o significado que os indicadores podem assumir para os responsáveis políticos, para a reputação dos estabelecimentos de ensino, para certos financiamentos, para o sucesso de determinados programas, etc., devem tomar-se todas as precauções para garantir a integridade, a qualidade e a neutralidade das análises interpretativas dos indicadores.

O ideal é recorrer a organismos ou entidades apolíticas não ligadas a grupos de pressão que sejam institucionalmente credíveis, estejam ao corrente dos problemas da educação e da gestão, metodologicamente competentes, capazes de sintetizar as informações e de as fazer chegar aos investigadores, aos especialistas em avaliação, aos professores, aos funcionários "generalistas", aos meios de comunicação social e ao público (Guthrie, 1989).

A experiência mostra que é possível encontrar uma solução para este difícil problema recorrendo aos serviços universitários especializados em matéria de investigação experimental em educação e avaliação. Contudo, é forçoso que estes serviços apoiem a sua actuação técnica numa reflexão permanente sobre as opções educativas e os debates políticos com elas relacionados. Também aqui é preciso ter em conta os debates sobre os objectivos culturais da escola: o que é que lhe exigimos em matéria de educação científica, histórica, ética, estética, social, cívica?... "Esta dimensão "sociopolítica" da pilotagem, ultrapassa o quadro do paradigma actual da pedagogia e, até mesmo, do paradigma dominante das ciências da educação. Porém, uma pilotagem que negligenciasse estas dimensões, corria o risco de se tornar, a longo prazo – se não muito rapidamente até – num instrumento tecnocrático pouco adequado a uma sociedade em constante transformação" (G. Fourez, comunicação pessoal).

Uma outra solução – que não exclui a precedente mas que até a completa – consiste na criação de uma "unidade independente de análise" que, integrada na administração educativa ou não, reuna representantes dos partidos políticos, interlocutores sociais, empresas, pais e público em geral. Uma unidade deste tipo deve gozar de muita liberdade em relação ao ministério da educação, ao partido político dominante ou, em geral, em relação a toda a autoridade que possa ter um interesse pessoal no assunto. Uma unidade assim é, normalmente, financiada com fundos públicos.

6. SÍNTESE DOS PROBLEMAS GERAIS

Colocam-se problemas por vezes difíceis a propósito dos indicadores. Já falámos da maior parte deles, e limitamo-nos agora a uma referência global ou à indicação de alguns aspectos ainda não referidos:

- Dificuldade de um acordo sobre as finalidades da educação e sobre os objectivos a atingir.
- Utilização apropriada dos dados.
- Comparabilidade dos valores obtidos na medição do rendimento escolar.

Para assegurar a comparabilidade máxima há que utilizar instrumentos de avaliação comuns (questionários, testes,...). Esta estandardização torna-se ainda mais importante quando os dados a recolher dizem respeito a comportamentos tais como atitudes, expectativas, traços de personalidade (capacidade de liderança, por exemplo) e modo de ensinar.

As categorias e a apresentação devem ser iguais em todo o lado, sobretudo se se trata da publicação de dados internacionais.

- Limitar a avaliação aos conhecimentos e capacidades cognitivas deixa de lado outros possíveis efeitos da educação, tanto no que se refere aos indivíduos como à sociedade.
- Os indicadores são estatísticas. Ora, é mais fácil atribuir um número à quantidade do que à qualidade. Devemos sempre verificar se os indicadores de rendimento escolar reflectem os conhecimentos reais dos alunos e o que eles são, efectivamente, capazes de fazer.
- Há características como a qualidade dos professores que são difíceis de avaliar em grande escala; variáveis apresentadas para substituir esta avaliação, como o tipo de formação ou a antiguidade, têm muito fraca validade.
- É injusto julgar os professores pelos resultados observados sem ter em conta as condições e factores do meio que os rodeia. Podem variar consideravelmente segundo as regiões, os bairros de uma cidade, etc.
- Quanto mais globais são os indicadores, maior risco correm de passar despercebidos fenómenos importantes.
- A publicação de indicadores tem um significado político e não é raro que surjam pressões tendentes a orientar a escolha ou a apresentação de resultados.
- Quanto menos definida é uma política ou um programa de educação mais difícil se torna escolher indicadores verdadeiramente representativos.
- Principalmente quanto aos indicadores internacionais, é preciso não perder de vista que os valores educativos e as opções políticas, bem como as instituições e modos de proceder, variam conforme os países e as culturas. O mesmo pode acontecer até a nível regional.
- As informações devem ser publicadas o mais rapidamente possível pois é importante poder reagir depressa, em caso de problemas graves; além disso, à medida que o tempo passa, o significado de um indicador pode mudar.
- Mesmo concebidos e reunidos com cuidado, os indicadores não fornecem informações de significado evidente. Este varia de acordo com os valores a que aderimos e ainda de acordo com as opções pedagógicas. As expectativas face ao sistema educativo variam, aliás, segundo os grupos: alunos, professores, famílias, empresas, autoridades religiosas.

A. Porter (1988) interroga-se, concretamente, sobre a questão da pilotagem prejudicar ou não a autonomia dos professores. De facto, a introdução da

pilotagem faz pender a balança a favor dos partidários de um controlo central do que é ensinado, do modo de o fazer, por quem e para quem. Deste modo, não estaremos a opor-nos à tão desejada descentralização e, em particular, à responsabilização dos professores e aos projectos de escola?

Serão os dois fenómenos incompatíveis? Não nos parece que assim seja, desde que os professores sejam associados à concepção e escolha dos indicadores e que a sua formação lhes permita agir com real conhecimento de causa.

O mesmo autor observa porém que permitir que os professores definam, por si próprios, os limites dentro dos quais pensam poder agir, não garante uma melhoria de educação. A participação dos professores no desenvolvimento dos indicadores poderá dar origem a indicadores que reflectam o denominador comum mais baixo da prática pedagógica na ocasião.

Em contrapartida, a existência de indicadores de grande qualidade pode favorecer a melhoria do estatuto dos professores: "Um bom sistema de indicadores pode ser uma arma política para os professores permitindo-lhes dar força à sua profissão, pois os indicadores não servem, apenas, para a pilotagem da prática pedagógica; também a influenciam."

CAPÍTULO 3 – EXEMPLOS DE REALIZAÇÕES

Há um número cada vez maior de países que já organizaram ou estão em fase de instaurar a pilotagem do seu sistema de ensino, em especial do seu rendimento, e que publicam periodicamente as conclusões deste trabalho. As observações que se seguem tratam essencialmente de dispositivos. São apresentados alguns exemplos de resultados, simplesmente para mostrar a natureza das informações recolhidas e dos modos de as apresentar. Não era possível apresentar uma descrição conjunta, num resumo deste género. Só os resultados dos estudos da Associação Internacional de Avaliação do Rendimento Escolar (IEA), de que nos ocuparemos a seguir, ocupam numerosos volumes... Na bibliografia o leitor poderá encontrar referenciadas as obras que tratam dos assuntos que, porventura, lhe interessem em especial.

INDICADORES NACIONAIS

1. ESTADOS UNIDOS

1.1. A “National Assessment of Educational Progress” (NAEP)

Decidida, fundamentalmente, em 1963 e efectivamente instaurada em 1969, esta *avaliação nacional do progresso educativo* abarca conhecimentos, capacidades e atitudes. Os objectivos a atingir são:

- Dispor, de forma contínua, de informações acerca dos níveis de rendimento escolar dos alunos.
- Medir a evolução dos seus níveis de rendimento, ao longo do tempo.
- Realizar sondagens especiais sobre certos aspectos dos níveis de rendimento escolar (por exemplo, quanto à compreensão da leitura), ou do comportamento dos jovens (por exemplo, os seus hábitos de consumo).
- Reunir dados, analisá-los e redigir sobre eles relatórios destinados a públicos variados: estudantes, professores, responsáveis pela administração, decisores, poderes organizadores,...
- Encorajar e apoiar as investigações sobre os dados recolhidos.
- Familiarizar os investigadores e os serviços de avaliação regionais e locais, com as avançadas técnicas utilizadas pela NAEP.
- Continuar a desenvolver e a aperfeiçoar as técnicas necessárias à recoleção e à análise dos dados obtidos.

Por ocasião de novo apelo lançado pelo Congresso em 1978, foi definido que o fim principal seria avaliar os níveis de rendimento de alunos e de jovens adultos nas competências básicas de: leitura, matemática e comunicação.

Para tal, a NAEP deveria:

- recolher dados, pelo menos de cinco em cinco anos, de modo a avaliar os níveis de rendimento de alunos de idades e níveis escolares diferentes, em leitura, em expressão escrita e em matemática;
- fazer periodicamente um relatório sobre as mudanças observadas, durante um dado período, no saber e no saber fazer relativos a estas disciplinas;
- realizar, se necessário, outras avaliações no domínio educativo;
- prestar assistência técnica aos serviços dos Estados e das comunidades locais, de modo a ajudá-los a compreender melhor os objectivos da avaliação nacional, a utilizar os dados relativos às competências básicas e a comparar os dados locais com os dados nacionais.

Primeiro a avaliação abrangeu quatro níveis de população: 9 anos, 13 anos, 17 anos e adultos entre os 26 e os 37 anos. Em seguida foram testados os quatro anos de escolaridade primária e o segundo e quinto ano do secundário. Em cada operação foram avaliados entre 75 000 e 120 000 jovens. Cada ano par realizavam-se avaliações. Com o tempo aumentou também o número de disciplinas examinadas. Foram acrescentadas: ciências, civismo, literatura, música, história, geografia, formação profissional e educação artística. Actualmente, a leitura e a matemática são avaliadas de dois em dois anos, a expressão escrita e as ciências de quatro em quatro anos, e as outras disciplinas de forma irregular.

É dada grande importância aos subprodutos destas operações: desenvolvimento de bancos de objectivos e de questões, criação de bases de dados que facilitem a construção de testes, etc.

De início os resultados publicados abarcavam o país inteiro e as quatro grandes regiões que ele compreende: Nordeste, Oeste, Centro, Sudeste. Evitava-se assim a comparação entre Estados; esta foi, contudo, exigida a partir de 1985; é um dado adquirido desde 1990.

A NAEP é financiada pelo *Office for Educational Research and Improvement* do Ministério da Educação dos Estados Unidos. Em vez de procurar constituir um serviço administrativo *ad hoc* de alto nível científico, o Ministério da Educação prefere recorrer a centros de excelência. Desde 1983 que a execução dos trabalhos foi confiada ao *Educational Testing Service* (ETS), em Princeton, o maior organismo mundial especializado na construção de instrumentos de avaliação. Embora mantenha relações estreitas com a Universidade de Princeton, o ETS é uma empresa privada. Foi fundado em 1947 como sociedade com fins não lucrativos.

A importância do seu volume de negócios permite-lhe contratar os melhores especialistas em avaliação... e dedicar à investigação em educação um orçamento superior ao total das despesas que a comunidade francesa da Bélgica lhe consagra.

O facto do ETS ter assumido a responsabilidade da NAEP (e ter criado, especialmente no seio da sua organização, um *Center for the Assessment of Educational Progress* - CAEP), traduziu-se, imediatamente, em notáveis progressos metodológicos e técnicos nos processos de recolha, análise e interpretação dos dados.

Uma das principais críticas dirigidas à NAEP, diz respeito ao seu carácter demasiado descritivo: "A NAEP é uma espécie de termómetro que indica a temperatura do corpo mas não nos indica qual a doença que dá origem à febre" (Richards, 1988:496). Mas não é por isso que os elementos descritivos e os sinais apresentados deixam de ter um interesse considerável.

1.1. Relatórios de Investigação

São apresentados de maneira muito cuidada e redigidos numa linguagem acessível ao grande público. Alguns destes relatórios tratam da avaliação feita entre diferentes avaliações. Indicam como os grupos (constituídos de acordo com a idade, o sexo, a região, o grupo étnico, o nível sociocultural dos pais, etc.) se comportam, tanto nas tarefas (*itens*) isoladas, como em conjuntos de tarefas.

Alguns exemplos apresentados a seguir, testemunham bem a grande quantidade de dados armazenados.

– *Leitura: primeiro exemplo.*

Para medir a capacidade de leitura em 1986, a NAEP utilizou grande variedade de textos e prestou particular atenção às estratégias de leitura.

Os textos a ler iam, desde frases simples exprimindo uma única ideia, até artigos complexos sobre assuntos de nível superior de ciências da natureza e ciências sociais. Esses textos tanto podiam constar de histórias, como de poemas ou de ensaios e de relatórios de análise, com origem em livros de leitura de nível elementar e de manuais do ensino secundário; figuravam, também, entre os documentos a ler, horários dos comboios e facturas do serviço de telefones.

A compreensão foi avaliada de vários modos: desde questões factuais de escolha múltipla, até questões abertas obrigando o aluno a reestruturar e interpretar a matéria lida.

O estudo foi concebido de modo a poder determinar as relações entre os níveis de competência e diferentes factores contextuais, considerados isoladamente ou em grupo.

Eis uma síntese dos resultados.

Quadro 1
Competência média na leitura dos alunos do 3º ano do ensino primário
e do 1º e 5º do secundário. Resultados nacionais

	3º ano da primária	1º do secundário	5º do secundário
Média nacional	38,1 (0,2) ^a	48,9 (0,1)	56,1 (0,2)
Brancos	39,8 (0,2)	50,3 (0,1)	57,3 (0,2)
Negros	33,4 (0,4)	45,2 (0,3)	51,5 (0,3)
Hispânicos	33,2 (0,3)	44,4 (0,4)	51,3 (0,3)
Nordeste	39,1 (0,3)	50,7 (0,3)	57,4 (0,5)
Sudeste	37,2 (0,3)	48,1 (0,2)	54,8 (0,3)
Centro	39,3 (0,4)	49,0 (0,2)	56,5 (0,5)
Oeste	36,9 (0,4)	48,0 (0,4)	55,4 (0,4)
Meio urbano desfavorecido	31,9 (0,5)	43,8 (0,4)	51,2 (0,6)
Meio urbano favorecido	41,2 (0,5)	51,6 (0,4)	59,5 (0,5)
Rapazes	37,3 (0,2)	47,5 (0,2)	54,5 (0,3)
Raparigas	38,9 (0,2)	50,3 (0,1)	57,7 (0,2)
Quartil superior	47,4 (0,1)	58,2 (0,1)	67,3 (0,1)
Quartil inferior	25,9 (0,1)	38,1 (0,1)	43,4 (0,1)

a. O erro-padrão de medida aparece entre parêntesis

Nos três níveis de ensino analisados, os alunos têm imensas dificuldades em comentar e em justificar a interpretação do que leram. Os autores concluem que o ensino devia dar atenção especial a estes dois parâmetros.

Entre outras observações de interesse, damos particular relevo às seguintes:

- Quando deparam com alunos com dificuldades na leitura, os professores insistem, sobretudo, nas estratégias de descodificação e muito menos na compreensão e na crítica.
- Os alunos fracos variam pouco as suas estratégias de leitura. No fim do ensino secundário encontram-se alunos que continuam a usar processos próprios do 3º ano da primária.
- No final do ensino secundário, os níveis de rendimento na leitura dos alunos pertencentes a minorias étnicas e a grupos sociais de risco aproxima-se do nível nacional do 1º ano do secundário.
- Os bons leitores pertencem aos cursos mais difíceis. Fazem com regularidade os deveres de casa e as famílias estimulam a leitura.

– Leitura: segundo exemplo

Neste segundo exemplo, os resultados aparecem expressos numa escala de competências com cinco níveis. As abordagens são feitas de acordo com a teoria da resposta aos *itens* (ver o artigo consagrado à abordagem das escalas, assim como dois exemplos pormenorizados de escalas, no anexo dedicado aos problemas técnicos). Esta complexa teoria progrediu muito, graças às investigações realizadas no *Educational Testing Service*, de Princeton; continua a ser objecto de controvérsia entre os especialistas.

Quadro 2

Percentagem de alunos que atingem ou ultrapassam um dos cinco níveis de competência na leitura – Anos: 1971 – 1984. – Idades: 9, 13 e 17 anos^a

Nível de leitura	Idade	1971	1975	1980	1984
Percentagem					
Rudimentar (150)	9	90,4	93,3	94,4	93,9
	13	99,7	99,6	99,8	99,8
	17	99,7	99,9	99,9	100,0
Básico (200)	9	58,3	61,7	65,1	64,2
	13	92,3	92,8	94,3	94,5
	17	96,6	97,5	97,9	98,6
Intermédio (250)	9	15,6	14,0	17,0	18,1
	13	57,0	57,5	59,3	60,3
	17	80,0	82,0	82,8	83,6
Apto (300)	9	1,1	0,7	0,8	1,0
	13	9,3	9,7	10,9	11,3
	17	37,2	36,1	34,8	39,2
Avançado (350)	9	0,0	0,0	0,0	0,0
	13	0,2	0,2	0,3	0,3
	17	4,9	3,5	3,1	4,9

a. *The NAEP reading scale*

Nível 150: Rudimentar – Capacidade para cumprir tarefas de leitura simples e isoladas.

Nível 200: Básico – Capacidade para compreender informações isoladas ou sequenciais.

Nível 250: Intermédio – Capacidade para procurar informações específicas, para as relacionar e para generalizar.

Nível 300: Apto – Capacidade para compreender, resumir e explicar informações relativamente complexas.

Nível 350: Avançado – Capacidade para sintetizar e perceber, através da leitura, textos especializados.

FONTE: NAEP, *The Reading Report Card, Progress Toward Excellence in Our Schools* (Relatório N° 15-R-01), 1985.

Verifica-se que a diferença de nível de rendimento, entre os melhores leitores e os mais fracos, permanece relativamente constante em todos os níveis de ensino. Há que encontrar os meios para corrigir este desnível...

A figura seguinte, representando o bom leitor, mostra bem a qualidade gráfica das ilustrações destinadas a transmitir informações pedagógicas básicas ao grande público.

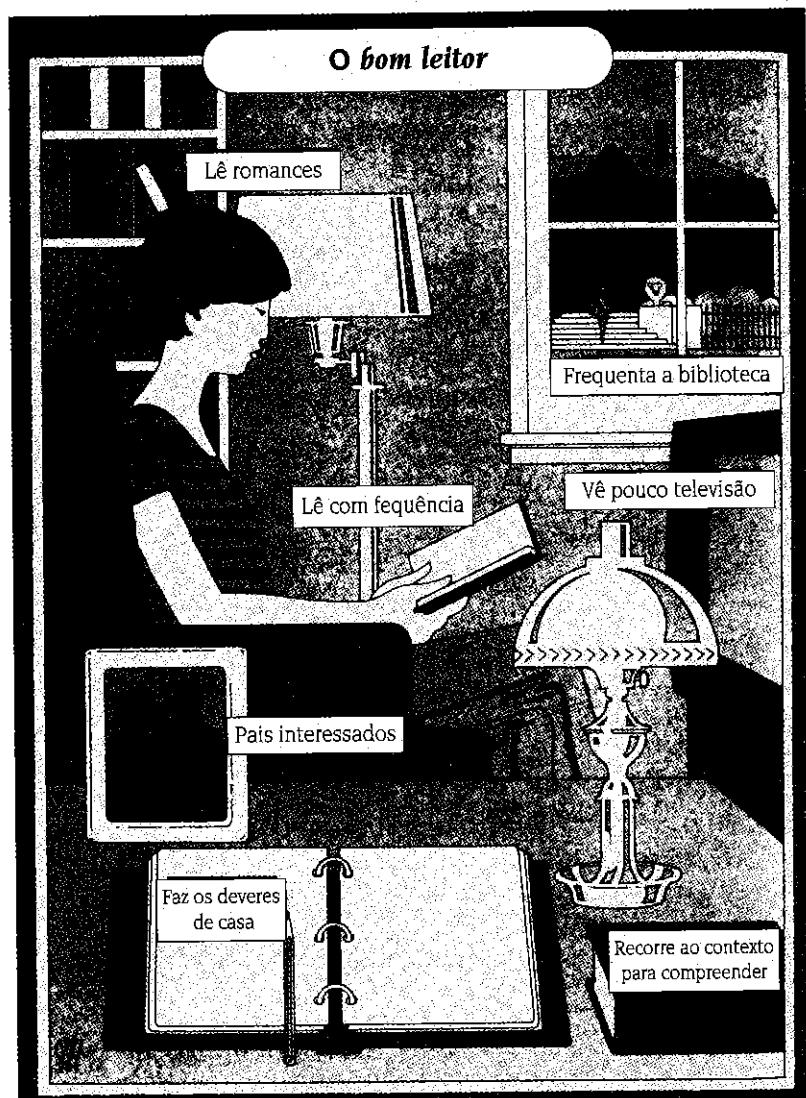


Fig. 4

– *Expressão escrita*

As conclusões desta avaliação realizada em 1984 em alunos do 4º ano do ensino primário e do 2º e 5º anos do secundário foram, também elas, consideradas alarmantes. A maior parte dos alunos, independentemente da sua origem (brancos, negros, asiáticos, hispânicos), são incapazes de se exprimir adequadamente por escrito, a não ser nos casos mais simples. Os resultados melhoraram bastante entre o 4º ano do ensino primário e o 2º ano do secundário, mas essa melhoria é muito menor entre o ensino secundário inferior e o superior.

No 5º ano do secundário não chegam a um quarto os alunos que conseguem exprimir-se, adequadamente, em tarefas que impliquem competências necessárias para prosseguir os estudos, entrar no circuito comercial ou exercer uma profissão liberal. A maior parte dos alunos americanos é incapaz de se exprimir, a não ser a um nível mínimo, sem garantia sequer de que a sua mensagem atinja o resultado pretendido (A.N. Applebee, J. Langer e I.V.S. Mullis, 1986:9).

Outra conclusão grave: esta avaliação revela que a maioria dos alunos apresentam deficiências relacionadas com as competências cognitivas superiores, em todos os níveis de ensino. Analisam mal um problema, e não têm poder persuasivo.

– *Matemática*

O quadro seguinte, diz respeito aos resultados em matemática; estão muito longe de ser brilhantes.

O nível 150 é atingido quase por todos, mas a partir do nível 200 começam a surgir as dificuldades: os alunos sentem dificuldades em efectuar operações que exijam mais do que um simples raciocínio numérico. Não chegam a um quarto os alunos de nove anos que atingem o nível 250 (domínio das quatro operações com números inteiros e raciocínio aritmético elementar) e, na globalidade, um terço dos alunos de treze anos não atingem este nível. Em relação aos dois níveis seguintes, os dados são pelo menos inquietantes.

Os autores concluem que os resultados são desencorajadores por duas razões. Primeiro, porque muito poucos alunos, seja qual for o seu grupo etário, atingem os níveis correspondentes, respectivamente, ao nível primário, secundário inferior e secundário superior. Depois porque, à medida que se avança na escolaridade, não cessa de aumentar o desnível entre os níveis de rendimento reais dos alunos e aqueles níveis que, legitimamente, se poderia esperar que eles alcançassem.

Quadro 3
Percentagem de alunos que atingiram ou ultrapassaram
os cinco níveis de competência em matemática

	Idade	1978	1982	1986
Nível 150 Adições e subtrações elementares	9 13 17	96,5 (0,2) 99,8 (0,0) 100,0 (0,0)	97,2 (0,3) 99,9 (0,0) 100,0 (0,0)	97,8 (0,2) 100,0 (0,0) 100,0 (0,0)
Nível 200 Início da competência e da compreensão	9 13 17	70,3 (0,9) ^a 94,5 (0,4) ^a 99,8 (0,0)	71,5 (1,1) 97,6 (0,4) 99,9 (0,1)	73,9 (1,1) 98,5 (0,2) 99,9 (0,1)
Nível 250 Domínio das quatro operações. Início da resolução de problemas.	9 13 17	19,4 (0,6) 64,9 (1,2) ^a 92,1 (0,5) ^a	18,7 (0,8) 71,6 (1,2) 92,9 (0,5) ^a	20,8 (0,9) 73,1 (1,5) 96,0 (0,4)
Nível 300 Processos de complexidade média. Geometria elementar. Raciocínio.	9 13 17	0,8 (0,1) 17,9 (0,7) 51,4 (1,1)	0,6 (0,1) 17,8 (0,9) 48,3 (1,2)	0,6 (0,2) 15,9 (1,0) 51,1 (1,2)
Nível 350 Resolução de problemas em várias fases. Álgebra.	9 13 17	0,0 (0,0) 0,9 (0,2) 7,4 (0,4)	0,0 (0,0) 0,5 (0,1) 5,4 (0,4)	0,0 (0,0) 0,4 (0,1) 6,4 (0,4)

a. Diferença estatisticamente significativa ao nível de 0,5 em relação a 1986

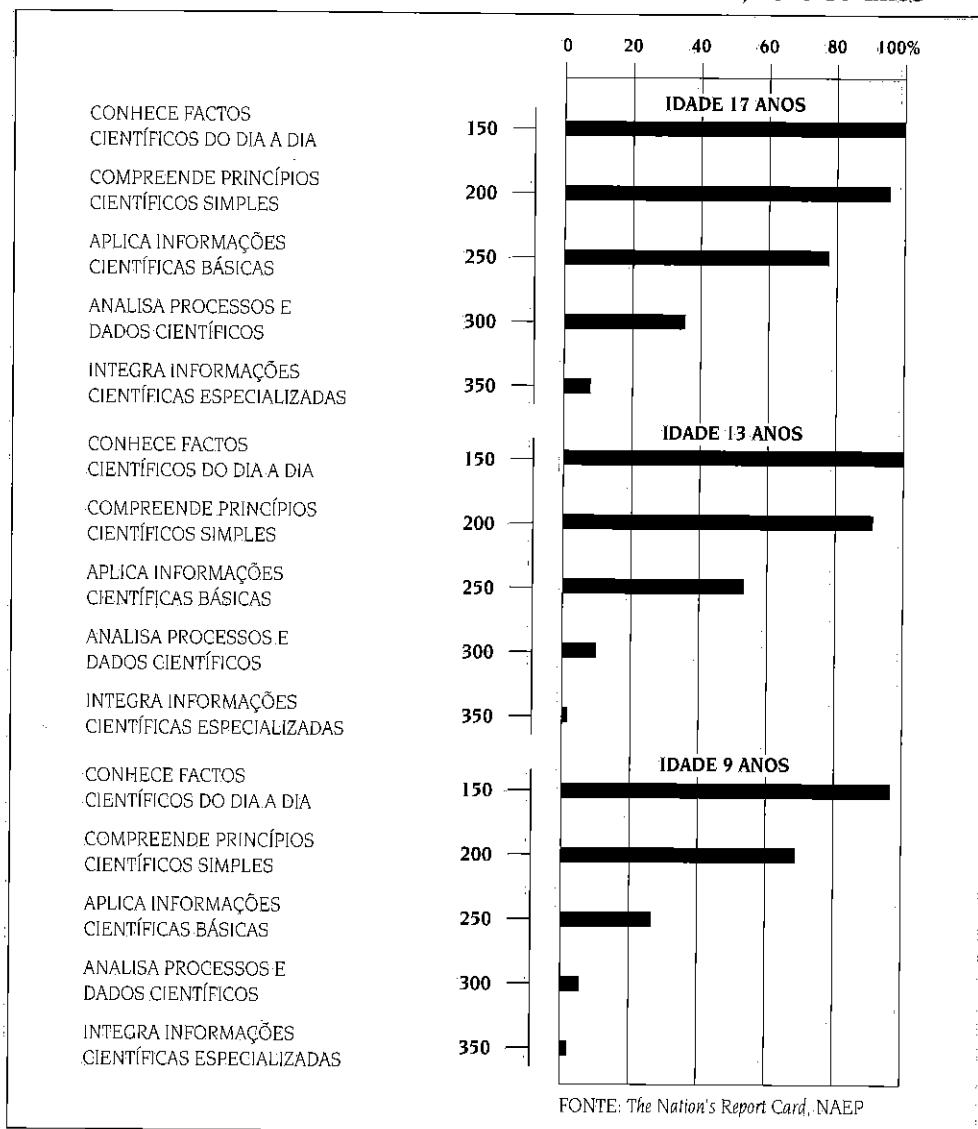
Como comentário aos dados de 1986, e tendo como referência as investigações internacionais da IEA e da IAEP, podem ler-se duas perturbantes observações no *Executive summary* do relatório da NAEP, que trata, em particular, dos resultados de 1986 (J.A. Dossey et al. s.d.):

“Os dados internacionais mais recentes relativos à aprendizagem da matemática, mostram que o aluno japonês médio atinge um nível superior ao de 50% dos alunos mais brilhantes dos Estados Unidos, que frequentam os cursos preparatórios de matemática na universidade. Recentemente, uma sociedade japonesa de produção de semi-condutores, teve de contratar estudantes com o primeiro ciclo universitário já terminado para efectuar controlos estatísticos de qualidade que são confiados, no Japão, a alunos acabados de sair do ensino secundário.”

– Ciências

Eis, antes da mais, a síntese dos resultados obtidos em ciências, em 1986, pelos alunos de 9, 13 e 17 anos, naturais da região Oeste dos Estados Unidos (Mullis e Jenkins, 1988:60).

Quadro 4
Competências em ciências. Percentagem de alunos que atingem ou ultrapassam os níveis fixados em 1986 - Idades: 9, 13 e 17 anos



Ainda a propósito das ciências, e para as mesmas idades, o quadro seguinte apresenta três pontos de referência (1977, 1982 e 1986) que permitem apreciar a evolução ao longo de dez anos.

Quadro 5
Percentagem de alunos que atingiram ou ultrapassaram os cinco níveis de competência em ciências
Idades: 9, 13 e 17 anos

	Idade	1977	1982	1986
Nível 150 Conhecimento rudimentar de plantas e animais. Descobre como utilizar aparelhagem doméstica simples	9 13 17	93,6 (0,5) ^a 98,6 (0,1) 99,8 (0,0)	95,0 (0,5) 99,6 (0,1) 99,7 (0,1)	96,3 (0,3) 99,8 (0,1) 99,9 (0,1)
Nível 200 Compreende princípios científicos simples relativos à vida animal ou vegetal	9 13 17	67,9 (1,1) ^a 85,9 (0,7) ^a 97,2 (0,2)	70,4 (1,6) 89,6 (0,7) 95,8 (0,4)	71,4 (1,0) 91,8 (0,9) 96,7 (0,4)
Nível 250 Início da compreensão do método experimental	9 13 17	26,2 (0,7) 49,2 (1,1) ^a 81,8 (0,7)	24,8 (1,7) 51,5 (1,4) 76,8 (1,0) ^a	27,6 (1,0) 53,4 (1,4) 80,8 (1,2)
Nível 300 Análise dos dados e dos métodos científicos	9 13 17	3,5 (0,2) 10,9 (0,4) 41,7 (0,8)	2,2 (0,6) 9,4 (0,6) 37,5 (0,8) ^a	3,4 (0,4) 9,4 (0,7) 41,4 (1,4)
Nível 350 Integra informações científicas especializadas	9 13 17	0,0 (0,0) 0,7 (0,1) 8,5 (0,4)	0,1 (0,1) 0,4 (0,1) 7,2 (0,4)	0,1 (0,1) 0,2 (0,1) 7,5 (0,6)

a. Diferença estatisticamente significativa ao nível de 0,5 em relação a 1986

De 1977 a 1986 verificou-se um ligeiro progresso geral. Contudo, os conhecimentos científicos e a capacidade de os utilizar continuam muito limitados. Além disso, dados que aqui se não reproduzem, indicam que os alunos pertencentes a minorias étnicas obtêm resultados claramente inferiores às médias nacionais. Por outro lado, as raparigas continuam a obter níveis de rendimento inferiores aos dos rapazes.

É interessante acrescentar aos dados precedentes, informações que, só muito raramente, constam das publicações relacionadas com os níveis de rendimento escolar. Tratam por um lado do modo como os professores ensinam as ciências e, por outro, do que fazem os alunos (I.V.S. Mullis *et al.*, 1988:99).

Quadro 6
Prática de ensino das ciências, 1986

Com que frequência (%) o vosso professor de ciências utiliza as seguintes estratégias no seu ensino?	Todos os dias	Várias vezes por semana	Uma vez por semana	Menos de uma vez por semana	Nunca
1º ano do secundário					
Expõe a matéria	31 (0,9)	21 (0,9)	18 (1,1)	13 (0,9)	18 (0,9)
Demonstra um princípio científico	15 (1,0)	22 (1,3)	22 (1,0)	22 (1,3)	19 (0,9)
Requer a explicação dos resultados de uma experiência	14 (1,1)	19 (1,3)	20 (0,8)	24 (1,1)	23 (1,0)
Solicita hipóteses	8 (0,8)	12 (0,8)	15 (1,0)	24 (1,0)	42 (1,5)
Solicita a interpretação de dados	9 (0,9)	12 (0,7)	15 (0,7)	15 (1,0)	48 (1,2)
5º ano do secundário					
Expõe a matéria	55 (1,0)	24 (0,8)	9 (0,8)	3 (0,3)	8 (0,7)
Demonstra um princípio científico	18 (1,0)	31 (1,6)	28 (1,2)	13 (1,0)	10 (0,7)
Requer a explicação dos resultados de uma experiência	16 (1,1)	29 (0,9)	28 (1,0)	15 (0,9)	12 (0,8)
Solicita hipóteses	11 (0,7)	19 (0,8)	23 (1,3)	22 (1,0)	25 (1,3)
Solicita a interpretação de dados	15 (0,8)	24 (0,9)	23 (0,9)	16 (0,9)	22 (1,0)

Erro-padrão entre parêntesis

Quadro 7
Ensino das ciências.
Frequência (%) de certas actividades didácticas, em 1986

Indique quantas vezes as seguintes actividades ocorrem na disciplina de ciências	Todos os dias	Por semana	Por mês	Por ano	Nunca
3º ano da primária					
Filme	—	23 (1,8)	16 (0,9)	18 (1,3)	31 (1,3)
Visita de estudo	—	3 (0,5)	7 (0,6)	24 (1,1)	55 (1,7)
Fazer experiências	11 (1,0)	30 (1,4)	—	26 (1,0)	19 (1,4)
Leitura do manual	33 (1,7)	27 (1,5)	—	11 (0,7)	15 (0,8)
Fazer um relatório oral ou escrito	20 (0,9)	16 (0,9)	—	17 (0,9)	33 (1,4)
1º ano do secundário					
Filme	4 (0,7)	11 (1,0)	23 (1,4)	40 (1,7)	21 (1,4)
Visita de estudo	1 (0,2)	1 (0,2)	1 (0,3)	16 (1,5)	82 (1,5)
Fazer experiências sozinho	3 (0,5)	5 (0,5)	13 (0,9)	29 (1,0)	50 (1,4)
Fazer experiências em grupo	4 (0,5)	9 (0,7)	22 (1,4)	34 (1,5)	30 (1,3)
Descrever uma experiência	4 (0,6)	5 (0,5)	10 (1,1)	21 (1,3)	60 (1,6)
Leitura do manual	43 (1,7)	24 (1,2)	15 (0,8)	8 (0,9)	9 (0,9)
Leitura de artigos científicos	12 (1,1)	14 (0,8)	23 (1,3)	25 (1,1)	26 (1,3)
Fazer um relatório oral ou escrito	3 (0,4)	5 (0,5)	11 (0,9)	35 (1,9)	46 (1,9)
5º ano do secundário					
Filme	2 (0,2)	7 (0,6)	22 (1,2)	42 (1,2)	27 (1,1)
Visita de estudo	0 (0,2)	0 (0,1)	1 (0,4)	13 (1,2)	86 (1,2)
Fazer experiências sozinho	2 (0,3)	7 (0,6)	20 (0,9)	25 (1,0)	46 (1,3)
Fazer experiências em grupo	3 (0,6)	14 (0,9)	36 (1,2)	28 (1,4)	18 (1,4)
Descrever uma experiência	2 (0,3)	10 (0,9)	22 (1,3)	24 (1,5)	41 (1,6)
Leitura do manual	28 (1,0)	26 (1,0)	16 (0,9)	12 (0,8)	18 (1,3)
Leitura de artigos científicos	7 (0,7)	10 (0,8)	18 (1,1)	26 (1,1)	39 (1,5)
Fazer um relatório oral ou escrito	2 (0,3)	3 (0,3)	10 (0,8)	34 (1,3)	52 (1,6)

Erro-padrão entre parêntesis

– *Repercussão a nível nacional*

Os resultados alarmantes dos estudos que acabamos de apresentar, não significam, porém, que o ensino nos Estados Unidos é sempre de pior qualidade do que noutra parte qualquer. É evidente que a explicação se encontra, em parte, nas condições muito desfavoráveis em que tantas escolas têm de trabalhar.

A manifestação mais palpável da inquietação que reina no país é, sem dúvida, a publicação pelo Presidente dos Estados Unidos, de um opúsculo intitulado *A Nation at Risk* (Uma Nação em Risco). No seu retumbante discurso de 18 de Abril de 1991, George Bush proclama que os Estados Unidos se deviam tornar “uma nação de estudantes”: que no ano 2000, 90% dos jovens deviam terminar, com sucesso, o ensino secundário de longa duração e que os alunos americanos deviam passar a ser os melhores do mundo em matemática e ciências. O futuro dirá se tal desejo se torna realidade...

1.2. Acervo de publicações

A NAEP tem centenas de publicações no seu activo. São do domínio público e podem ser reproduzidas sem autorização. Além das já referidas, há que mencionar:

- *Relatórios técnicos*, publicações de alto nível científico que descrevem, em particular, as técnicas estatísticas utilizadas. O NAEP 1983-84 *Technical Report* Nº15, da responsabilidade de A.E. Beaton (1987), é um modelo do género.
- *Fascículos de dados* objectivos, preparados por comissões de investigadores, especialistas de várias disciplinas, responsáveis pela administração, professores, especialistas em avaliação, representantes das empresas e cidadãos a quem o assunto diga respeito.
- *Conjuntos de itens de testes*. Depois de serem utilizados, cerca de metade dos *itens* de testes construídos para as avaliações são postos à disposição dos Estados e das comunidades locais. Os *itens* ficam a pertencer ao domínio público. Dizem respeito a várias disciplinas e apresentam um vasto leque de dificuldades. Vêm acompanhados da percentagem da amostra nacional de estudantes que responderam correctamente às questões.
- *Estudos*. Estes documentos de alto nível, elaborados por especialistas, tratam de aspectos particulares da avaliação e de questões com ela relacionadas.
- *Publicações metodológicas*. Trata-se, em certos casos, de verdadeiros tratados de várias centenas de páginas.

1.3. The National Education Goals Reports

Em 1989, o Presidente dos Estados Unidos reuniu os governadores de todos os Estados para definir as *principais finalidades educativas da nação* para o ano 2000. Com o fim de fazer a pilotagem das acções destinadas a alcançar estes objectivos, constituiu-se uma comissão chamada *National Educational Goals Panel*. Edita todos os anos um relatório público (*The National Education Goals Report*), com informações sobre os progressos alcançados pelo sistema educativo na prossecução destes objectivos.

As grandes rubricas deste relatório são as seguintes:

1. Competências e aptidões favoráveis ao estudo.
2. Percentagem do grupo etário que termina o ensino secundário de nível superior.
3. Rendimento escolar em matemática e ciências.
4. Competência na leitura.
5. Vida saudável, longe da droga.

O relatório de 1992 insiste especialmente no contexto internacional: os níveis de rendimento observados nos Estados Unidos são comparados aos de outros países.

1.4. Outras avaliações nacionais importantes: estudos longitudinais

Paralelamente às investigações nacionais realizadas pela NAEP, são levadas a cabo no país outras investigações sobre avaliação de grande envergadura. Merecem uma atenção especial os estudos longitudinais em grande escala.

Normalmente, os indicadores recolhidos ao longo do tempo não se referem sempre aos mesmos alunos, mas a amostras representativas de diversas populações de alunos, que se encontram em determinado nível dos seus estudos. Parte-se da hipótese de que essas sucessivas populações apresentam características equivalentes. O que nem sempre é correcto. Acompanhar as mesmas amostras durante vários anos pode, evidentemente, fornecer informações mais rigorosas e mais válidas sobre as alterações de comportamento e de níveis de rendimento e, de uma maneira geral, sobre a evolução dos indivíduos em causa.

Os estudos longitudinais são difíceis. Verifica-se, em especial, que a amostragem inicial diminui devido a mudanças de domicílio ou de escola, por desinteresse ou por razões mais graves ainda. Têm de se empregar meios consideráveis e dispendiosos para manter os contactos. Sabe-se também que os alunos que

recusam continuar a participar, podem apresentar certas características psicológicas específicas, o que tende a provocar desvios na amostragem. Finalmente, sobretudo quando se trata de determinar os efeitos de programas educativos, pode ser necessário recorrer a planos experimentais de grande complexidade.

Mas não é menos verdade que acções deste tipo, ainda que relativamente modestas, são altamente desejáveis.

Neste campo, e dada a crise educacional que atravessam neste momento, os Estados Unidos não regateiam meios. R.C. Atkinson e G.B. Jackson (1992:46-47) descrevem as acções em curso. São impressionantes. De facto, o Centro Nacional de Estatísticas da Educação (NCES) dos Estados Unidos financia, simultaneamente, três grandes estudos longitudinais: o *National Longitudinal Study* de 1972, o *High School and Beyond Study* de 1980, e o *National Educational Longitudinal Study* de 1988.

Projecto Talento

Antes de descrever as investigações acabadas de referir, importa recordar que foram precedidas de uma outra, que continua única na história da educação:

O Projecto Talento. Esta investigação foi financiada pelo *American Institute for Research* e realizada pela Universidade de Pittsburgh (Flanagan, J.C. et al., 1964; Flanagan, J.C., Ed., 1971).

Em paralelo com o recenseamento legal de 1960, os Estados Unidos decidiram proceder a um inventário das aptidões de todos os jovens de quinze anos de idade e acompanhar o seu desenvolvimento durante vinte e cinco anos, realizando nova avaliação de cinco em cinco anos. Uma das finalidades deste acompanhamento era, concretamente, descobrir quais as experiências educativas que poderiam ter contribuído mais para a realização das potencialidades detectadas.

Foi escolhida, ao acaso, uma amostra de um estrato nacionalmente representativo. Abarca 500 000 jovens, entre os quais se contam todos os alunos de quinze anos matriculados no último ano do ensino secundário inferior e nos três anos do secundário superior. Foram, igualmente, testados jovens que já não frequentavam a escola.

Não só se recolheram informações de carácter geral (especialmente com a ajuda de um inventário de interesses e de um questionário de personalidade), como também se aplicaram os seguintes testes: língua materna (expressão, correcta utilização, ortografia), matemática (conhecimento, compreensão e utilização de conceitos básicos), aptidão para a aprendizagem de línguas

estrangeiras. Os jovens foram ainda interrogados sobre os seus planos para o futuro. Paralelamente foram entrevistados os responsáveis pela gestão e os conselhos pedagógicos das escolas.

Não deixa de ter interesse histórico recordar que, como sucedeu com os estudos internacionais da IEA, estas investigações só foram possíveis com a chegada do computador: tratava-se da análise de mais de dois milhões de formulários e do tratamento de mais de mil milhões de dados...

Todos os jovens examinados com quem foi possível manter o contacto, foram de novo objecto de estudo um, cinco, dez e vinte anos após a idade normal de saída do ensino secundário. Foram interrogados, em especial, sobre os estudos realizados, as actividades de trabalho e de lazer e sobre os seus problemas.

As publicações relativas a este estudo continuam, ainda, a sair.

National Longitudinal Study, 1972: entrada no ensino superior e na vida activa

Este estudo feito sobre uma amostra nacional de 16 683 alunos prestes a terminar os seus estudos, refere-se à transição do ensino secundário de longa duração para o ensino superior e do ensino superior para a vida activa. Os dados recolhidos são de natureza demográfica e escolar: cursos seguidos, resultados escolares, projectos de estudos e de trabalho. No prosseguimento do processo, foram recolhidos dados em 1973, 1974, 1976, 1979 e 1986, a fim de identificar as alterações nos estudos, no trabalho e no estado civil.

Verificou-se que a maioria dos alunos testados em 1972 continuaram os seus estudos, mas sem seguirem o modelo tradicional de estudos superiores feitos em quatro anos, imediatamente após a saída do secundário. Geralmente os homens levam mais tempo a acabar os seus estudos do que as mulheres.

O estudo *High School and Beyond* de 1980, versou também sobre a transição da escola para a vida activa; a quebra de rendimento escolar foi a principal razão desta investigação. Em 1980 foram testadas amostras nacionais de 28 000 alunos do penúltimo ano e de 30 000 alunos do último ano de 1100 escolas secundárias.

Em 1982, 1984 e 1986 foram examinados subgrupos alargados destas duas secções de alunos; em 1982, os alunos que tinham sido testados no penúltimo ano do secundário em 1980, foram de novo observados. Os estudos versaram sobre as atitudes dos pais, as opiniões dos professores e os níveis alcançados pelos alunos em diferentes testes. Foi possível registar observações muito importantes sobre o abandono dos estudos e sobre os factores que determinam as expectativas no decurso dos dois últimos anos do secundário.

Comparações entre os dois estudos acabados de referir permitem determinar em que medida as mudanças sociais e culturais ocorridas modificaram as atitudes e os níveis de rendimento escolar dos estudantes observados.

National Educational Longitudinal Study, 1988: factores de sucesso escolar

Este estudo tem por objectivo determinar os factores que contribuem para os níveis de rendimento escolar, para a prossecução de estudos e para a entrada no ensino superior. Durante dez anos, far-se-á o acompanhamento de 25 000 alunos a partir do quarto ano do ensino secundário. Para além da recolha de informações de carácter geral, foram aplicados em 1988, 1990 e 1992, testes de leitura, matemática, ciências, geografia e história. Ficou, assim, a dispor-se de uma base de 36% de alunos brancos, 44% de alunos negros e 39% de hispânicos que apresentam, pelo menos, dois factores de risco. Estes factores são: famílias monoparentais, pobreza (rendimento anual inferior a 15 000 dólares), crianças deixadas sozinhas em casa durante mais de três horas por dia, pais que não tiveram sucesso nos estudos secundários, abandono dos estudos por parte de um irmão ou irmã, fraco conhecimento da língua inglesa.

1.5. Sistemas de indicadores dos próprios Estados

Quase todos os Estados possuem os seus testes de conhecimento e muitos elaboram o seu próprio sistema de indicadores e de pilotagem. Califórnia, Connecticut, Estado de Nova Iorque e Carolina do Sul estão especialmente avançados nesta matéria.

Exemplo da Califórnia

O sistema de indicadores PACE, actualmente em vigor na Califórnia, trata dos seguintes aspectos, dados a conhecer numa publicação anual: **Situação da Educação na Califórnia**.

Inputs

- Número de alunos e suas características: pertença a minorias étnicas ou linguísticas, composição e rendimentos da família,...

- Recursos humanos, incluindo aqui o perfil dos professores e dos responsáveis pela administração; habilitações e programas de formação contínua dos professores (inclusive durante o primeiro ano em que exercem funções) e do pessoal.
- Organização das turmas e sistema disciplinar; características da escola e do distrito.

Processos

- Currículo e programas especiais, ramos de ensino seguidos pelos alunos, exigências para a atribuição do diploma de conclusão do ensino secundário, exames, manuais escolares e outros materiais, financiamento de programas especiais.

Outputs

- Níveis de rendimento dos alunos, incluindo aqui a tendência dos seus resultados nos *California Achievement Tests* (resultados escolares globais, de acordo com a etnia, a língua materna, o nível sócio-económico da família), níveis obtidos no *Scholastic Aptitude Test* (teste de aptidão para os estudos superiores), informações sobre os trabalhos em casa, o número de horas diárias de televisão e, em geral, todas as outras variáveis que possam influenciar os resultados escolares.

Para a maior parte destes indicadores já se dispõem de tendências que abrangem um intervalo de dez anos.

Os **Performance Reports for California Schools** dizem respeito a cinco sectores principais. Os quatro primeiros são dedicados à qualidade da educação, enquanto que o quinto trata das variáveis contextuais relacionadas com as escolas e os alunos:

1. Cursos escolhidos pelos alunos e competências mínimas exigidas.
2. – Níveis de rendimento na leitura e na matemática dos alunos que terminam o ensino secundário de longa duração.
- Percentagem de alunos que se situam, respectivamente, no primeiro, segundo ou terceiro nível.
- Níveis alcançados pelas escolas.

3. Níveis obtidos pelos alunos que pretendem seguir estudos superiores:

– No *Scholastic Aptitude Test* (SAT). A utilização deste teste de aptidão ao ensino superior, em todo o país, há já muitos anos, permite detectar a evolução havida, e os resultados servem, muitas vezes, de ponto de referência em debates públicos. (O SAT trata sobretudo, de vocabulário, de raciocínio verbal, de conhecimento do programa de matemática do ensino secundário, de raciocínio quantitativo. Possui elevado poder discriminativo relativamente à selecção inicial e uma boa validade na predição do sucesso dos estudos universitários).

4. Taxas de abandono escolar.

Um questionário a que os alunos respondem, permite determinar o nível sócio-económico e sócio-cultural das escolas, assim como outras características do meio envolvente.

2. INGLATERRA

2.1. Unidade de Avaliação do Rendimento Escolar (Assessment of Performance Unit, APU)

A criação da APU pelo Ministério da Educação e Ciência foi decidida em 1975, mas os trabalhos efectivos só começaram em 1978; duraram até 1988. De início, a avaliação destinava-se não tanto ao grande público, mas mais aos responsáveis pela distribuição dos créditos para a educação. Passado pouco tempo, a publicação dos resultados alargou-se muito mais. Por princípio não se publicava nem o nome dos alunos testados nem o das escolas. Pretendia-se, deste modo, evitar que as escolas com melhor rendimento atraíssem os alunos mais dotados, o que enfraqueceria, ainda mais, as escolas com resultados menos bons. Não é verdade que o que importa, sobretudo, é oferecer a estas últimas meios para poderem melhorar?

Foram testadas amostras representativas de alunos de 11, 13 e 15 anos nas seguintes áreas:

- Linguagem: capacidade de comunicação em língua materna através da leitura, da escrita, da capacidade de ouvir e de falar.
- Matemática (insistência na comunicação).
- Ciências: modo de lidar com as provas, verificação de hipóteses, experimentação, capacidade para extrair conclusões com fundamento.

- Desenvolvimento pessoal e social: aprender a conhecer-se melhor, sentido das responsabilidades, comportamento moral em relação à sociedade e ao meio.
- Desenvolvimento estético.
- Desenvolvimento físico incluindo a psicomotricidade.

A partir de 1980 os resultados das escolas nos diversos exames foram tornados públicos. Muito rapidamente sistemas de moderação permitiram uma melhor interpretação destes resultados.

Em 1988 teve lugar uma reforma geral da educação que levou à adopção de normas de pilotagem muito avançadas. A APU foi suprimida, assim como a *School Evaluation and Assessment Council* (SEAT). Esta decisão coincide com a adopção de um currículo nacional.

2.2. Currículo Nacional e sua avaliação: uma pilotagem sistemática

Em Inglaterra, é tradição que as autoridades locais em matéria de ensino determinem, com muita liberdade, o currículo das escolas que dependem da sua jurisdição. A fim de se poder dispor de informações sobre os rendimentos a nível nacional, uma lei de reforma da educação (*Education Reform Act*) de 1988, prescreve uma avaliação nacional sobre um currículo nacional. Como acabámos de referir, esta avaliação vem substituir a da APU. A fim de verificar em que medida os objectivos definidos estão a ser alcançados, devem ser testadas amostras das populações de 7, 11, 14 e 16 anos.

A lei de 1988 marca uma viragem na história da educação na Inglaterra. É criado, pela primeira vez, um quadro legal destinado a elevar o nível de rendimento escolar. A lei cria também um *Conselho Nacional do Currículo*, organismo responsável pela reforma contínua dos programas; a este conselho cabe ainda a missão de dar pareceres ao governo e às escolas.

O Currículo Nacional comprehende três disciplinas “centrais” – língua materna, matemática e ciências – e sete disciplinas “fundamentais” – tecnologia, história, geografia, uma língua estrangeira, artes plásticas, música e educação física. Acrescenta-se a esta lista a religião, se os pais assim o desejarem.

Há quatro ciclos de estudos distintos:

1. dos cinco aos sete anos.
2. dos oito aos onze anos.
3. dos doze aos catorze anos.
4. dos quinze aos dezasseis anos.

Os alunos dos dois primeiros ciclos devem estudar todas as disciplinas, excepto língua estrangeira. Nos dois ciclos superiores há maior flexibilidade.

A lei de 1988, foi completada em 1992 com um livro branco intitulado *Escolha e diversidade: novo quadro para as escolas*. Trouxe consigo, concretamente, as indicações seguintes, relativas ao aumento de autonomia dos estabelecimentos de ensino e à pilotagem:

- A constituição de redes de escolas autónomas (*self-governing*) é estimulada, a fim de melhor responder às diversas necessidades e interesses dos alunos.
- É dada, ainda, maior importância à pilotagem rigorosa das escolas e à informação dos pais e da comunidade educativa.

Um aspecto particular da pilotagem prevista no documento, merece menção especial. Como actuar quando as avaliações e relatórios da inspecção revelam que uma escola, dependente das autoridades locais, não garante aos seus alunos uma educação de qualidade satisfatória, funciona de uma maneira inaceitável e deve, portanto, ser considerada como um estabelecimento de ensino "de risco"?

"Numa primeira fase, a autoridade local em matéria de ensino deve comprometer-se a elaborar um plano de reforma. Se não tiver êxito, o ministro encarregará uma associação educativa de assegurar o funcionamento da escola. Esta associação deve incluir pessoas que possuam experiência reconhecida em gestão e educação. A associação deverá, de facto, gerir a escola até que os resultados educativos melhorem e atinjam o nível desejado. Só nessa altura é que o Estado poderá conceder subsídios" (The White Paper, *School Update*, Outono de 1992:3).

São definidos alvos a atingir na aprendizagem (*attainment targets*) para cada ciclo e para cada disciplina, assim como os programas e meios de avaliação.

Ao falar de *alvo de aprendizagem* referimo-nos aos conhecimentos, às competências e à compreensão real que os alunos devem, normalmente, atingir no fim de cada ciclo. Para cada um destes objectivos a atingir distinguem-se dez níveis de rendimento; este processo deverá permitir avaliar o progresso realizado no decurso da escolaridade obrigatória.

Ao falar de *programa de estudos* queremos designar os conteúdos, as competências e os processos a adquirir por alunos com diferentes aptidões e maturidade, ao longo dos quatro ciclos.

Ao falar de *dispositivo de avaliação* estamos a referir-nos às avaliações efectuadas no fim, ou quase no fim, de cada ciclo, a fim de determinar em que medida é que os níveis de rendimento atingem os *alvos*. A avaliação resulta da combinação dos resultados obtidos em testes elaborados no exterior e da apreciação dos professores. Facto importante: os testes não são iguais para todas as turmas; os

professores podem elaborar provas, a partir de um banco de questões. Em qualquer dos estádios, os pais serão informados dos resultados dos filhos. Aos onze, catorze e dezasseis anos, os resultados conjuntos serão tornados públicos.

Os programas dos diversos ciclos, bem como os dez níveis dos alvos de aprendizagem a atingir, são muito pormenorizados. Como prova disso, apresenta-se a seguir um curto extracto dos alvos de aprendizagem de tecnologia da informação, ponto que faz parte do programa de tecnologia.

NÍVEL 1

- 1a *Ser capaz de trabalhar com um computador.*
- 1b *Ser capaz de explicar que brinquedos ou aparelhos de uso doméstico podem ser comandados por sinais ou instruções informatizadas.*

NÍVEL 2

- 2a *Ser capaz de utilizar imagens, símbolos, frases criadas no computador para comunicar o seu ponto de vista.*
 - 2b *Utilizar o computador para armazenar e encontrar informação.*
- ...

NÍVEL 9

- 9a *Ser capaz de avaliar um conjunto de material informático ou um tipo de computador, analisar a situação para a qual foram concebidos; avaliar a sua eficácia, facilidade de utilização, nível de adequação e sugerir aperfeiçoamentos.*
- 9b *Conceber e realizar um sistema informático, capaz de ser utilizado por outros, e redigir instruções relativas à sua utilização.*
- 9c *Compreender as consequências da introdução de dados falsos nos ficheiros que contêm informações pessoais.*

NÍVEL 10

- 10a *Ser capaz de construir um modelo de sistema, conceber e testar a sua realização; justificar os métodos utilizados e as opções tomadas.*
- 10b *Discutir problemas postos pela tecnologia da informação, nos domínios da economia, da ética, da moral e da sociedade.*

A organização e direcção da avaliação são da alçada do Schools Evaluation and Assessment Council (SEAC).

Esta avaliação nunca se traduz em termos de sucesso ou de insucesso. Qualquer aluno que não atinja o nível previsto é submetido a testes diagnósticos

complementares. São assim concebidos testes especiais para as crianças com necessidades específicas. Sob este ponto de vista o novo sistema pretende ser claramente formativo.

A avaliação está também concebida de modo a fornecer aos professores o *feedback* de que necessitam para planificar o ensino a um nível adaptado aos seus alunos. Em especial, a distinção em dez níveis para cada alvo de aprendizagem deverá ajudar os professores a tomarem como ponto de partida aquilo que os seus alunos já sabem, compreendem e são capazes de fazer. É deste modo que se espera obter uma subida sistemática dos níveis de rendimento.

Finalmente, as escolas têm liberdade para aplicar o currículo de acordo com as modalidades que mais lhes convenham.

Como já foi dito, o público tem acesso aos resultados das escolas e os pais ou os responsáveis pela escola podem, também, ter acesso aos resultados de determinada turma. Estes resultados não são expressos em termos de valor acrescentado, em função do nível sócio-económico dos alunos matriculados, mas é elaborado um relatório para cada escola, não só sobre o seu funcionamento mas também sobre dados contextuais – sócio-económicos e outros – que possam ter influência.

As escolas têm obrigação de enviar aos pais um relatório anual, indicando tanto os resultados obtidos pelos filhos nos testes do Currículo Nacional como também todas as outras informações sobre os progressos realizados.

Concretamente, e a título de exemplo, os pais de uma criança de sete anos, submetida a testes nesse ano, devem receber as seguintes informações:

- Qual a situação do filho em língua materna, matemática, ciências e tecnologia, relativamente aos alvos de aprendizagem a atingir.
- Qual a situação do filho em relação a certos aspectos particulares destas disciplinas (perfis). Por exemplo, os níveis obtidos em língua materna estão adequados à expressão oral, à compreensão de mensagens orais, à leitura e à expressão escrita.
- Qual a posição destes resultados face aos dos outros alunos de sete anos que frequentam a escola.
- Os progressos em cada uma das disciplinas do Currículo Nacional.
- Os progressos nas outras disciplinas e actividades.
- Uma apreciação global sobre o progresso do aluno.
- O levantamento das faltas.
- O nome de alguém da escola com quem o relatório possa ser discutido.

O programa de avaliação adoptado em 1992 era o seguinte:

- A partir de 1993, os alunos de sete anos serão submetidos a testes de língua materna, matemática, ciências, tecnologia, história e geografia.
- A partir de 1994, os alunos de onze anos serão submetidos a testes de língua materna, matemática, ciências, tecnologia e os alunos de catorze anos a testes de história e geografia.
- A partir de 1995, os alunos de onze anos serão submetidos a testes de história e geografia e os alunos de catorze anos a testes de línguas estrangeiras.

Contudo, em 1993 foi preconizado um aligeiramento deste programa de avaliação devido à oposição em massa dos professores.

Uma decisão importante foi resolver publicar, futuramente, as medidas estandardizadas dos progressos realizados pelos alunos de cada escola no Currículo Nacional. Mais precisamente:

- A partir de 1995, determinam-se os progressos realizados entre a primeira sessão de testes dos sete anos e a dos onze anos.
- A partir de 1997, entre a sessão dos onze e a dos catorze anos.
- A partir de 1999, entre a sessão dos onze e a dos dezasseis anos.

Além disso, todas as escolas deverão publicar anualmente um prospecto que, juntamente com várias informações, deve mencionar:

- os resultados obtidos pela escola nos testes sobre o Currículo Nacional.
- o que vieram a ser os alunos que terminaram a escolaridade obrigatória na escola.

O Currículo Nacional, que contraria uma longa tradição de liberdade local, foi objecto de grande oposição por parte dos professores e dos seus sindicatos, mas o governo foi em frente. Chama a atenção para o facto de se oferecerem grandes possibilidades de modelar o currículo a nível local. Teme-se, também, que a publicação dos resultados obtidos pelas escolas venha a prejudicar as que trabalham em contextos mais difíceis. A ideia liberal da livre concorrência, inclusive no domínio escolar, encontrou contudo, também aqui, novos defensores.

A oposição dos professores explica-se, sobretudo, pela alteração aos seus hábitos de trabalho. Temem a criação de uma grande máquina administrativa com carácter tecnocrático, capaz de monopolizar o seu tempo e energia, em especial

para tarefas de avaliação, em detrimento da função pedagógica. Por outro lado, a reforma não foi objecto de uma suficiente concertação prévia.

Temores e críticas que são justificados, pelo menos em parte, mas que não devem ocultar a grande qualidade global do currículo implantado, sobretudo das opções pedagógicas fundamentais, das finalidades e objectivos, da definição de competências básicas, do espírito e dos instrumentos de avaliação e da qualidade do acompanhamento dos alunos, tanto dos especialmente dotados como dos que revelam dificuldades de aprendizagem. De facto, é difícil determinar em que medida as inquietações e críticas não são causadas, em parte, pelos grandes esforços de adaptação exigidos pela considerável elevação de nível de "profissionalização" que as autoridades pretendem suscitar, tanto no que respeita ao ensino como à avaliação dos seus resultados.

2.3. Indicadores de rendimento

As Autoridades Escolares Locais (LEA) da Inglaterra preparam-se para publicar indicadores sobre os seguintes aspectos:

- Taxa de enquadramento (*ratio* professor-alunos).
- Habilidades dos professores (estudos feitos).
- Competências pedagógicas dos professores.
- Atitudes profissionais dos professores (dedicação às tarefas, satisfação profissional, ...).
- Qualidade na exploração do currículo.
- Utilização do tempo atribuído.
- Nível de investimento dos alunos no trabalho.
- Qualidade das experiências de aprendizagem.
- Deveres de casa.
- Pontualidade, absentismo.
- Atitudes e comportamentos dos alunos (incluindo aqui o vandalismo e o riscar as paredes).
- Custo por aluno.

Concluindo a análise da orientação geral da Inglaterra no domínio da pilotagem, Gipps pensa que a avaliação ligada a objectivos específicos do currículo, tal como é praticada actualmente no seu país, constituirá um poderoso instrumento de controlo e de avaliação das escolas, dos professores e dos alunos.

3. FRANÇA: UM DESENVOLVIMENTO EXEMPLAR

Se bem que a França tenha sido pioneira em matéria de avaliação de atitudes e conhecimentos – A. Binet e Th. Simon publicam, em 1905, o primeiro teste operacional de inteligência sobre os processos cognitivos superiores (*L'échelle métrique de l'intelligence*) e H. Piéron é o pai da docimologia (1922) – atravessou uma época de desconfiança em relação à avaliação objectiva no domínio da educação. O Maio de 1968 exacerbou ainda mais esta tendência. Após um período de transição, durante o qual se procurou e, em parte, se alcançou um justo equilíbrio entre o aspecto qualitativo e quantitativo, a França ocupa, actualmente, lugar entre os líderes em matéria de pilotagem de sistemas escolares. De facto, sobretudo sob a influência da *Direcção de Avaliação e Prospectiva* de que falaremos e de alguns grandes especialistas, surgiu nos últimos anos uma “cultura da avaliação”.

Por circular de 24 de Maio de 1974, o ministro J. Fontanet criou um serviço de avaliação dos níveis de rendimento em matéria de educação que servisse para a pilotagem. Este serviço transforma-se, em 1987, na *Direcção de Avaliação e Prospectiva* (DEP) que desempenha hoje um papel de primeiro plano na cena nacional e internacional.

A DEP não intervém na construção de currículos. Contudo, as informações que fornece constituem elementos importantes para a adaptação de programas já existentes ou para a concepção de novos programas.

Um documento redigido por M. Le Guen (1991) para a OCDE, apresenta de forma pormenorizada o funcionamento da DEP. O que dizemos a seguir é inspirado nesse trabalho.

Para fins de pilotagem, a DEP realiza amplos estudos normativos, a que se associam todos os membros da sociedade educativa, principalmente inspectores, professores, investigadores, sindicatos de professores e associações de pais. Os representantes de cada uma destas entidades constituem comités de acompanhamento. Participam na definição dos objectivos destes estudos, sobretudo a partir de programas e directivas oficiais.

Em seguida, são elaborados e experimentados testes e questionários. Recolhem-se, então, os comentários dos professores responsáveis pelas turmas submetidas aos pré-testes. A partir deste momento é rigorosamente respeitado o anonimato dos participantes, sendo a DEP a única a ter acesso aos dados.

As amostras de escolas são escolhidas ao acaso. Distribuem-se segundo dois critérios: o meio envolvente (urbano ou rural) e o tamanho das escolas. Nos estabelecimentos escolhidos testam-se todos os alunos que pertençam aos níveis

escolares em causa. Pode deste modo estudar-se a influência da dimensão dos estabelecimentos de ensino.

Facto importante: graças à colaboração estreita dos professores nas actividades, dá-se uma transferência de metodologias e de técnicas, o que suscita a esperança dos centros regionais poderem, de futuro, proceder a estudos particulares concebidos por equipas locais.

Os trabalhos desenrolam-se em três direcções principais: avaliação das aquisições, avaliação da política educativa, produção e difusão de instrumentos de avaliação.

3.1. Avaliação das aquisições durante a escolaridade

Desde o começo dos anos oitenta que a França estabeleceu a pilotagem das aquisições escolares. Esta abrange na totalidade os currículos do ensino primário e secundário, tendo-se igualmente em conta as aprendizagens extra-escolares. Além de outros aspectos, os processos cognitivos, as atitudes e opiniões são objecto de questionários.

Para além da leitura – que é objecto de particular atenção – são examinadas, desde a escola primária, a expressão escrita, a matemática e aspectos de cultura geral. No ensino secundário são abrangidas de forma recorrente todas as disciplinas.

É utilizada uma grande variedade de instrumentos: questionários de escolha múltipla, questões de resposta aberta breve ou longa, *checklists*, escalas, etc.

Por outro lado, sobretudo no ensino secundário, a “vida na escola” é também objecto de observação. Para tal, alunos, professores e responsáveis pela direcção dos estabelecimentos respondem a questionários que tratam, sobretudo, da adaptação à vida escolar, das relações interindividuais, das regras de funcionamento, da aceitação e percepção do ensino, da ajuda dada pelas famílias, dos trabalhos em casa. Analisa-se, ainda, a adaptação ao meio social exterior à escola e a capacidade de liderança. Há um terceiro aspecto relacionado com a preparação para a vida activa: informações recebidas pelos alunos, qualidade da orientação.

Faz-se uma descrição das escolas e no que respeita às turmas em particular, cria-se para cada aluno uma ficha com a idade, sexo, nacionalidade, antecedentes escolares, língua materna, línguas estudadas, categoria sócio-profissional e estudos dos pais.

Finalmente, os questionários destinados aos professores, abordam as disciplinas ensinadas, os objectivos julgados importantes, apreciação sobre a

dificuldade dos testes para os alunos, previsão do sucesso nas provas. Também é pedida uma breve descrição dos métodos de ensino.

Em qualquer dos casos é garantido o anonimato.

A análise dos dados assim recolhidos diz respeito, sobretudo:

- À percentagem de alunos de cada nível escolar que atingem o critério de sucesso das provas.

Figura 5 a
**Percentagem de sucesso nas questões de Francês e Matemática,
no fim do 5º ano, em 1991**

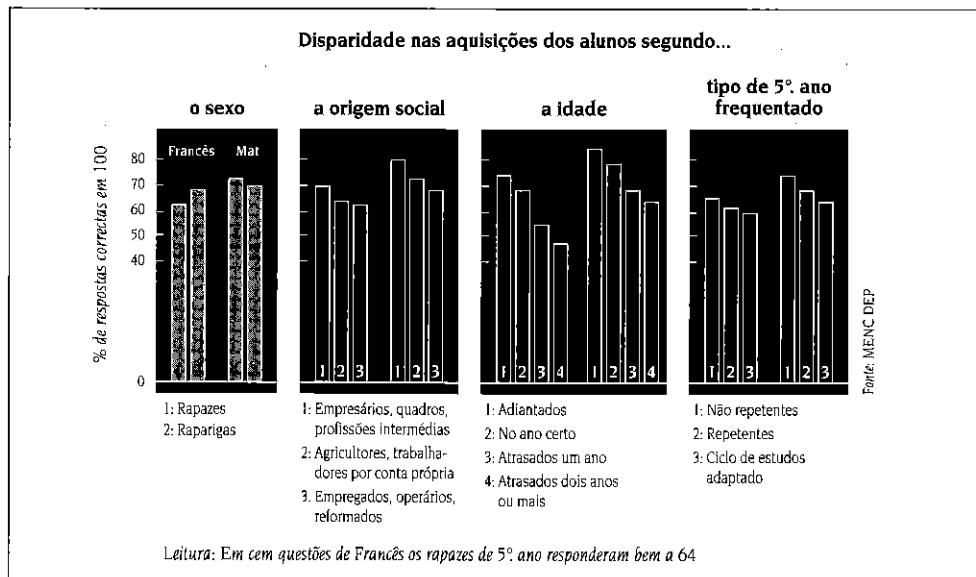
	Conjunto	10% dos mais fracos	10% dos melhores
Francês	66,7	36,7	89,7
Matemática	74,9	44,4	95,2

Leitura: em média os alunos responderam corretamente a 66,7% das questões de Francês.

- À relação entre o saber e o saber fazer, numa ou em várias disciplinas.
- Às correlações entre o rendimento escolar e outras variáveis (repetência, dimensão da escola, categoria sócio-profissional dos pais,...).

Avaliações deste tipo permitem, por exemplo, fazer uma estimativa do progresso nos conhecimentos ao longo da escolaridade, verificar se há diferenças entre rapazes e raparigas e porquê, e identificar os principais problemas que se apresentam.

Figura 5 b
Disparidade nas aquisições dos alunos segundo o sexo, a origem social, a idade e o tipo de 5º ano frequentado



O exemplo anterior, sobre a disparidade de aquisições, é extraído de *L'état de l'école*, 1992:39. O exemplo que a seguir se apresenta, trata da avaliação das aquisições; é retirado de *Education et formation*, 1993:34.

Quadro 8
Evolução das aquisições cognitivas e sócio-afectivas dos alunos,
entre Outubro de 1990 e Maio de 1992
(ou seja entre o começo do 6º ano e dois anos após)

Resultados	Outubro de 1990		Maio de 1992		Evolução (em proporção do desvio-padrão dos resultados de 1990)
	Média (M)	Desvio Padrão (s)	Média (M)	Desvio Padrão (s)	
Teste de abordagem ao francês	37,3%	19,1%	51,0%	20,0%	+ 0,717 σ
Teste de abordagem à matemática	50,4%	25,3%	62,6%	24,1%	+ 0,482 σ
Estratégias de trabalho	- 0,057	0,657	- 0,064	0,700	- 0,011 σ
Organização	0,908	0,538	0,872	0,640	- 0,067 σ
Ajuda da família no trabalho em casa	- 0,218	0,952	- 0,574	0,964	- 0,374 σ
Tempo dedicado ao trabalho em casa:					
dias por semana	4,5 dias	1,5 dias	4,7 dias	1,4 dias	+ 0,130 σ
duração diária	1h17 min	52 min	1h20 min	53 min	+ 0,058 σ
Atitudes de desmobilização	- 1,017	0,539	- 0,874	0,590	+ 0,263 σ
Motivação escolar	0,830	0,587	0,549	0,710	- 0,479 σ
Auto-imagem escolar	0,249	0,692	0,36	0,704	+ 0,160 σ
Auto-imagem social	0,583	0,918	0,794	0,922	+ 0,231 σ
Gosto pela cooperação	0,589	0,728	0,697	0,716	+ 0,150 σ
Segurança, sociabilidade	0,207	0,239	0,163	0,288	- 0,184 σ
Atitude defensiva, agressividade	- 0,047	0,617	0,072	0,634	+ 0,193 σ
Cultura cívica	13,66	3,99	14,41	4,14	+ 0,176 σ

Nota:

- os resultados cognitivos estão expressos em % de itens respondidos acertadamente;
- as escalas sócio-afectivas (à excepção da Cultura Cívica) foram medidas numa escala que vai de -2 a +2;
- os resultados de Cultura Cívica são a soma das respostas de tipo "democrático" aos vinte e quatro itens que compõem a escala.

O quadro apresenta a evolução das aquisições dos alunos, desde o começo do sexto e o fim do segundo ano de escola secundária (5º ou 6º ano conforme os casos). Para cada um dos critérios escolhidos, podemos observar a média (M) e o desvio-padrão dos resultados (s) considerados em cada momento, bem como um índice de evolução.

Este índice obtém-se, calculando a diferença entre a média dos resultados de "entrada" e de "saída" que, em seguida, se divide pelo desvio-padrão dos resultados de entrada.

Este processo destina-se a facilitar a leitura dos resultados da evolução: permite, de facto, situar os "progressos" ou os "retrocessos" observados numa mesma escala, seja qual for o critério. Um índice de evolução próximo de 0 significa que, em dois anos, não houve praticamente evolução. Um índice positivo (por exemplo: + 0,250), indica que o resultado médio considerado após dois anos de escola secundária, é superior (no exemplo: um quarto de desvio-padrão) ao observado no início do 6º ano. Um índice negativo (por exemplo: - 0,500), significa que o resultado médio baixou entre 90 e 92 (no exemplo: meio desvio-padrão).

3.2. Evolução da política educativa

As principais preocupações são a aceitação ou rejeição das medidas que são tomadas e a avaliação do nível a que são alcançados os objectivos julgados prioritários. É dada particular atenção às medidas tomadas em favor dos alunos que apresentam dificuldades nos estudos.

Um melhor conhecimento do funcionamento do sistema escolar deve conduzir a um progresso nas aprendizagens e permitir uma maior eficácia do sistema.

3.3. Produção e difusão de instrumentos de avaliação

Os professores associados à concepção de instrumentos de avaliação e/ou encarregados de testar os seus alunos com estes instrumentos, familiarizam-se com este tipo de técnicas. É boa política deixar, depois, os instrumentos de avaliação à disposição dos professores, a fim de que possam, se quiserem, avaliar as potencialidades e pontos fracos dos seus alunos, no início e no decurso do ano escolar. Familiarizar-se com os métodos e instrumentos de avaliação é importante, também, para os quadros pedagógicos, os pais e os interlocutores sociais.

3.4. Alguns esclarecimentos sobre a testagem em grande escala

No sistema educativo francês a *escola elementar* compreende três anos de pré-primária e quatro anos de *primária* (cursos elementares 1 e 2; cursos médios 1 e 2), a escola secundária (*collège*) abrange quatro anos (6º, 5º, 4º, 3º) e o liceu (*lycée*) três anos (segundo, terceiro, terminal).

Desde 1989 que se faz uma avaliação nacional, no início do curso elementar 2, no início do 6º ano da escola secundária e depois no início do segundo ano do liceu.

J. Colomb (1993:3-4) esclarece:

"Nestas três importantes fases de articulação do sistema educativo (...), o conjunto dos alunos de cada um destes níveis são submetidos a uma bateria de provas nacionais que pretendem avaliar as suas competências em francês e matemática, sendo estas provas escolhidas por um grupo de pilotagem nacional, a partir de propostas surgidas no 'terreno'.

Trata-se, pois, de uma operação de grande envergadura: dirige-se, todos os anos, a mais de dois milhões de alunos! No primeiro ano, centralizaram-se todos os dados para

uma análise a nível nacional; a partir de 1990 são centralizados e objecto de análises diversas, apenas os dados correspondentes a uma amostra representativa de alguns milhares de alunos por nível.

Três grandes finalidades estão na origem da implantação desta avaliação:

- Aperfeiçoar o conhecimento individual dos alunos a fim de detectar, desde o início do ano escolar, os alunos que apresentam lacunas na sua aprendizagem e poder assim recorrer rapidamente a processos de remediação, antes que se passe a novas aprendizagens sobre noções que tiveram um tratamento deficiente. A personalização da ajuda a dar aos alunos, embora difícil de realizar, constitui um meio essencial da gestão dos públicos escolares.*
- Pôr à disposição dos parceiros do sistema educativo referências nacionais que permitam situar melhor as acções levadas a cabo, do nível da turma ao nível nacional. A repetição anual destas operações permite um melhor acompanhamento do sistema educativo e, portanto, a definição e aplicação de processos de regulação.*
- Constituir instrumentos de "medida" da eficácia da escola que, mesmo dando conta apenas duma pequena parte da realidade, têm o mérito de permitir uma perspectiva mais objectiva da realidade do funcionamento da escola."*

Colomb observa, finalmente (p.12):

"Estes resultados são bem aceites pela opinião pública quando confirmam as ideias correntes e, estranhamente, são muitas vezes ignorados quando contradizem estas mesmas ideias... Pelo contrário, a nível de professores, os instrumentos de diagnóstico que por este meio lhes são fornecidos pelo ministério, parecem ter cada vez melhor acolhimento, apesar das pesadas tarefas materiais que lhes são impostas. O risco de constituição duma 'norma nacional' e, portanto, duma espécie de exame dissimulado acompanhado da correspondente 'preparação ad hoc', não parece estar a verificar-se."

O autor, contudo, na página treze contradiz esta última nota de optimismo. Ao observar que as avaliações nacionais fornecem aos responsáveis pelas decisões preciosos indicadores e instrumentos de análise do funcionamento do sistema educativo (e portanto da sua pilotagem), Colomb queixa-se da avaliação desempenhar, cada vez mais, este papel "e não o de controlar e regular o sistema, papel que deveria ser o seu." Não há, contudo, incompatibilidade entre as duas funções.

3.5. Notável conjunto de publicações

Várias publicações periódicas dão testemunho da considerável actividade da DEP:

– *L'état de l'école* (publicação anual)

A edição de 1992 apresenta a escolha de trinta e seis indicadores relacionados com a actividade, o custo e os resultados do sistema educativo. Esta publicação anual informa o grande público sobre o estado do sistema, a sua evolução e, em relação a alguns indicadores, sobre a situação da França em relação a outros países. A principal preocupação é dar resposta a duas questões: o sistema educativo é eficaz; é equitativo?

Na medida da eficácia, dá-se relevo à capacidade de escolarizar as crianças, e à capacidade de lhes fazer adquirir os conhecimentos e competências desejáveis. Os indicadores de "actividade" e de "custos" fornecem elementos de resposta a estas questões. Por seu lado a equidade é analisada sob três aspectos importantes: os estabelecimentos de ensino são equivalentes entre si, os benefícios da educação distribuem-se de forma igualitária seja qual for o sexo e a origem social? A edição de 1992 trouxe novos esclarecimentos sobre a situação das regiões na competição escolar ao publicar as taxas regionais de acesso ao diploma de conclusão do ensino secundário em 1991.

L'état de l'école tem a possibilidade de ser completada com publicações mais específicas. Assim, foi publicada em 1993 uma *Géographie de l'école*. Reúne trinta e seis indicadores que mostram as disparidades geográficas e a sua evolução.

– *Repères et références statistiques sur les enseignements et la formation*.

Esta publicação anual trata, sobretudo, dos aspectos estruturais do sistema educativo. Em cada ano são tratados aspectos particulares ligados à actualidade educativa. (Veio substituir os *Tableaux des enseignements et de la formation*, publicados de 1966 a 1982).

– *Vade-mecum statistique*.

Publicação anual (disponível em disquetes) relativa aos principais dados de referência sobre educação, na forma de séries cronológicas.

– *Éducation et formation*.

Revista trimestral que apresenta, em forma de artigos, dados ou estudos, temas que abrangem todos os aspectos do sistema educativo, assim como o seu meio económico e social. Trata também de aspectos metodológicos cujo conhecimento é necessário à compreensão dos resultados estatísticos. São

publicados além disso, números especiais para difundir conjuntos coerentes de resultados estatísticos ou estudos particulares.

Por exemplo, o nº 22 (Janeiro de 1990), trata dos métodos de análise dos estabelecimentos escolares e da sua utilidade para a avaliação, dos níveis obtidos na prova de acesso ao ensino superior, dos indicadores de eficácia dos estabelecimentos, da perspectiva sociológica do funcionamento dos liceus e da singularidade dos estabelecimentos escolares.

– *Notes d'information*

Pequenos documentos, publicados uma ou várias vezes por mês. Por exemplo, os três números de Outubro de 1992 tratam, respectivamente, da taxa de aprendizagem, do ensino da educação física e desportiva e da origem social dos estudantes de 1960 a 1990.

4. CANADÁ

Até ao presente e devido à autonomia ciosamente defendida pelas diferentes províncias, estas desenvolveram os seus próprios sistemas de pilotagem. O Quebec servirá de exemplo.

Foi criado em 1967 um Conselho de Ministros da Educação do Canadá (CMEC). Funciona como *forum* vocacionado para a cooperação nacional e pode lançar projectos capazes de aumentar a qualidade da educação, tanto a nível nacional como provincial. O projecto “*Indicateurs nationaux des rendements scolaires*” orientado para as preocupações comuns, é um deles.

4.1. Quebec: indicadores do ensino primário e secundário

Desde 1985 que o Ministério da Educação do Quebec publica um conjunto de indicadores, antecedidos de comentários gerais, que procuram pôr em evidência os pontos mais salientes.

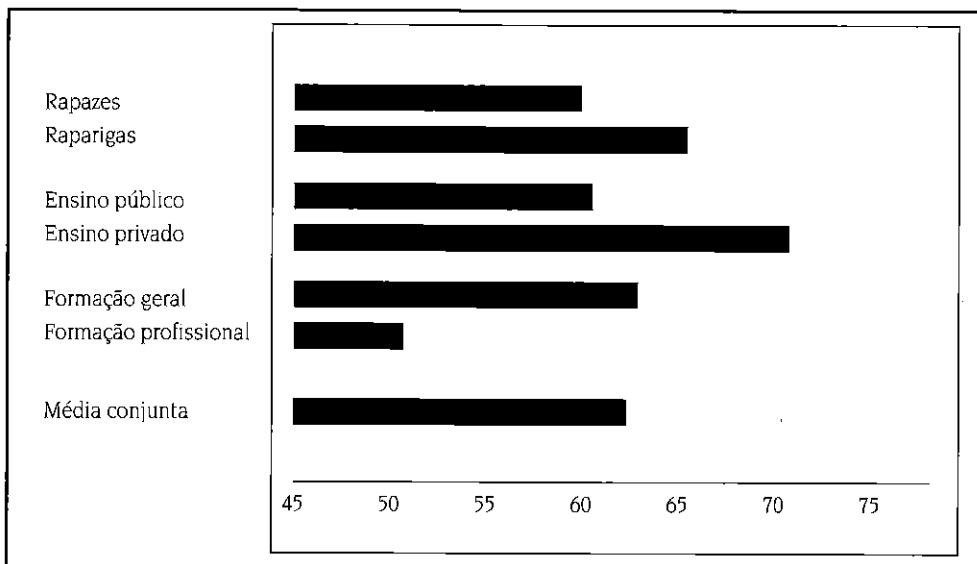
Uma secção trata das vias de ensino; outra dos resultados obtidos, quer nas provas próprias do Quebec quer nas provas internacionais, como as da Associação Internacional de Avaliação do Rendimento Escolar (IEA).

Eis um exemplo dos resultados publicados anualmente nos *Indicateurs sur la situation de l'enseignement primaire et secondaire* de 1989.

Quadro 9
Resultados na prova escrita de francês do 5º ano do secundário
segundo o sexo, a rede de ensino e o tipo de formação dos alunos

	Média (em %)			Percentagem de sucesso		
	1986	1987	1988	1986	1987	1988
Rapazes	55,4	57,5	60,2	41,4	43,2	49,9
Raparigas	61,7	63,6	65,6	57,4	59,0	63,9
Ensino público	57,3	58,9	61,2	46,3	47,2	52,8
Ensino privado	65,6	68,8	71,0	66,3	71,6	77,1
Formação geral	60,2	62,1	63,9	58,3	55,0	59,5
Formação profissional	49,1	51,0	51,2	26,6	26,4	25,9
Média conjunta	58,7	60,8	63,1	49,8	51,9	57,4

Figura 6
Média na prova escrita de francês do 5º ano do secundário
segundo certas variáveis, Maio de 1988



Estes resultados deixam transparecer um progresso regular e generalizado do rendimento. Este é sempre melhor nas escolas privadas do que nas públicas: as primeiras fazem uma selecção dos candidatos a admitir o que não se passa com as segundas.

Outros indicadores referem-se ao financiamento, ao número de diplomados e à integração dos jovens no mercado de trabalho.

4.2. Indicadores nacionais do Canadá

Este projecto, adoptado em 1988 pelo Conselho de Ministros da Educação do Canadá, tem como primeiro objectivo, “garantir uma base de informações que permita aos ministros provinciais avaliarem os seus programas escolares, comparando-os com as normas vigentes em todo o Canadá.” (CMEC, 1989:5).

Mais especificamente, os objectivos gerais do projecto são os seguintes:

- Permitir ao Conselho de Ministros da Educação a publicação dum relatório anual, que apresente à população canadiana os resultados obtidos, em cada província, relativamente a determinados indicadores da educação.
- Ajudar os ministros a avaliar os progressos realizados e a identificar as prioridades educativas.
- Ajudar os ministros a assegurar aos jovens canadianos um ensino de qualidade, preparando-os para as suas responsabilidades de cidadãos e de membros úteis da sociedade.

A ideia central do projecto é a comparação, quer entre os objectivos e as expectativas quer entre os resultados quantitativos e qualitativos. Distinguem-se três componentes:

1. *Indicadores de participação*: taxas de participação, de perseverança e de obtenção do diploma. Explicam até que ponto os jovens aproveitam a possibilidade de seguir estudos.

2. *Indicadores de expectativa*

- Em que medida são alcançados os objectivos comuns aos sistemas educativos?
- Em que medida os diplomados estão satisfeitos com os seus estudos?
- Que espera o público dos alunos que terminam a escolaridade obrigatória e que obtêm um diploma?
- Em que medida o sistema escolar dá resposta à expectativa do público em matéria de língua e de cálculo?
- Em língua e em conhecimentos matemáticos, qual o nível que os alunos pensam ter de atingir para terem sucesso na vida activa, após a obtenção do diploma?

3. Indicadores de resultados escolares

- Medida segundo a qual os alunos de treze e dezasseis anos atingem os objectivos expressos nos programas e expectativas do público quanto à compreensão da língua escrita e da matemática. Testes nacionais irão permitir que, pela primeira vez, as províncias verifiquem em que situação se encontram. Estas provas não vêm substituir as avaliações que os professores fazem dos alunos.
- Medida segundo a qual os resultados vão aumentando entre os treze e os dezasseis anos.

As sessões de testes deverão corresponder às seguintes exigências (Palmer, 1993:2):

- Medir os rendimentos dos sistemas educativos provinciais e não os dos alunos considerados individualmente (não se facultarão resultados, nem por escolas nem por turmas).
- Serem de referência criterial.
- Só será publicado o perfil de rendimento de cada província (nem médias, nem números particulares).
- A decisão de testar a população de dezasseis anos foi fácil de tomar, dado que a idade corresponde ao fim da escolaridade obrigatória. A opção de testar os alunos de treze anos, resulta de um compromisso: como a segunda sessão de testes tem lugar três anos depois, é ainda possível utilizar, pelo menos em parte, os mesmos instrumentos com fins comparativos.
- Os professores estão intimamente associados à avaliação. Participam na elaboração dos testes e experimentam-nos em turmas escolhidas por todo o país.
- A exemplo do que se faz nos Estados Unidos, distinguem-se três níveis de rendimento para a leitura, a expressão escrita e a matemática. Um teste preliminar permite determinar em qual destes cinco níveis o aluno deve ser testado; contudo, cada aluno é examinado em dois níveis consecutivos para determinar com mais segurança os seus limites.
- Em matemática são utilizados dois tipos de testes: um trata dos conhecimentos e outro da resolução de problemas. Além disso, os alunos devem responder a um questionário relacionado com a sua escola, com a língua falada e com o seu passado em matemática.
- Os testes de leitura compreendem quatro textos de dificuldade variável; a propósito destes textos, são feitas entre quarenta a cinquenta perguntas (questionários de escolha múltipla e questões abertas).

- Os testes são validados sob o ponto de vista do conteúdo em relação ao currículo e os professores são consultados quanto à qualidade das questões.

De acordo com os planos, os testes de rendimento escolar deverão também, posteriormente, abranger os conhecimentos científicos e culturais (conhecimento e consciencialização do património canadiano: geografia, história, literatura) e o sentido crítico. De acordo com a actual tendência é preconizado o recurso a testes sobre rendimentos/conhecimentos efectivamente alcançados.

5. SUÉCIA

5.1. Sistema precursor: harmonização liberal da atribuição de notas

No início

Há já muito tempo que a Suécia se decidiu pela pilotagem da maneira mais democrática possível. Desde os anos sessenta que o sistema de *moderação* foi adoptado para a avaliação no ensino secundário superior (S. Henrysson, 1964).

Por um lado, os professores continuaram a avaliar os seus alunos segundo os critérios habituais. Pedia-se, contudo, que fosse utilizada uma escala com sete níveis e que as notas se repartissem, no seu conjunto, segundo uma distribuição *gaussiana*. Era dada liberdade ao professor para ter em conta as características da sua turma (por exemplo, grandes competências ou quase todos os alunos médios, etc.).

Por outro lado, era posta à disposição dos professores uma bateria de testes de conhecimentos abrangendo as principais áreas do programa. Esta bateria era construída todos os anos por um serviço de investigação onde colaboravam especialistas em avaliação e professores com experiência. Actualmente, os testes de que trataremos a seguir são elaborados por quatro serviços de investigação universitários.

Os testes, administrados a amostras nacionais representativas, estão também divididos em sete categorias, o que possibilita a percepção, por parte dos professores, da situação dos seus alunos em relação à média nacional (os professores devem verificar se estão a ser mais ou menos severos do que os seus colegas doutras escolas ou se estão perante turmas excepcionalmente dotadas, ou excepcionalmente fracas).

Sem alterar a ordem da classificação inicialmente estabelecida, os professores tinham liberdade para rectificar as suas notas em função das comparações efectuadas. Não tinham de dar contas a ninguém das suas decisões.

Esta atitude particularmente liberal para com os professores contribuiu muito para lhes transmitir confiança nos testes; familiarizaram-se cada vez mais com eles. Quando, em 1993, a utilização dos novos testes nacionais, de que trataremos a seguir, se tornou obrigatória apenas para o ensino secundário superior, 95% dos professores do primário e do secundário inferior decidiram utilizá-los também (Wedman, 1993:7).

Actualmente

Actualmente o sistema de atribuição de notas é, globalmente, o seguinte (S. Marklund, 1991).

Durante os nove anos de escolaridade obrigatória, os professores dão notas aos alunos de uma forma inteiramente livre, não reprovando ninguém (não há repetências). No quarto ano e no sétimo dispõem de testes diagnósticos de língua materna e matemática (avaliação puramente formativa). No oitavo ano os alunos são submetidos a um teste estandardizado de inglês, língua estrangeira. No nono ano são administrados testes estandardizados de língua materna e matemática. Os resultados obtidos nestes testes têm apenas um valor indicativo para os professores: ficam a saber como estão os seus alunos em relação aos de outras escolas, em aprendizagens tidas como fundamentais.

Terminado o ensino secundário superior (12º ano, dezanove anos, para o ensino geral pré-universitário; 11º ano, para o ensino "geral técnico, profissional ou social"), a avaliação final decorre do seguinte modo:

1. Os professores submetem os seus alunos a exames do tipo tradicional. São atribuídas notas às provas, mas os resultados não são publicados.

Os professores utilizam uma escala com cinco níveis de notas, distribuindo-as *grosso modo* de acordo com a curva de Gauss:

Notas	1	2	3	4	5
Percentagem	7 %	24 %	38 %	24 %	7 %

2. Em Abril ou Maio, os alunos do ensino geral são submetidos a testes nacionais, escalonados também em cinco categorias, de língua materna, matemática, física, química e línguas estrangeiras (inglês e, à escolha,

alemão ou francês); os alunos do ensino técnico fazem teste de língua materna e de inglês.

3. Reúne-se, então, um conselho de professores (com autoridade legal) para examinar as notas propostas por cada um dos seus colegas, em função dos resultados obtidos em cada uma das duas provas acima mencionadas. Em princípio, as notas não podem apresentar um desvio superior a 0,2 pontos para mais ou para menos dos níveis obtidos nos testes nacionais. Se o professor pretender um desvio superior a 0,2 – o que é raro – deve apresentar justificação por escrito. Finalmente o conselho decide que notas atribuir nas disciplinas submetidas a testes nacionais. Nas restantes, o professor continua livre para actuar.

5.2. Avaliação nacional

Após dez anos de estudos preparatórios, também a Suécia instaurou uma avaliação nacional a partir de 1989. São avaliados os saberes e o saber fazer relativos à globalidade do currículo da escolaridade obrigatória, assim como o método de trabalho dos alunos. Pretende-se que esta avaliação seja formativa e que se destine, pois, a ajudar os professores no seu trabalho. São colocadas duas questões: "Trabalham as escolas suecas como deveriam? Os objectivos propostos às escolas são alcançados?".

A avaliação de 1989, abrangeu o 2º ano da primária (língua materna, matemática, música e outras artes), o 5º ano (as disciplinas anteriores, mais a educação cívica, as ciências e o inglês). Em 1992, fez-se a avaliação do 3º ano do ensino secundário. Estas avaliações repetem-se de três em três anos.

Comparado com os programas de testes de pouca duração – como os da NAEP – o programa de testes sueco leva muito tempo. Cada escola deve consagrar cerca de vinte horas à aplicação das provas e cada professor designado gasta quase o mesmo tempo para as corrigir. Segundo Wedman (1993:10), "tal investimento era tido como razoável dado que eram elevadas as vantagens retiradas da operação por todos os que nela participavam; valia a pena um esforço pedagógico desta natureza."

Todos os professores recebem os resultados dos seus alunos, bem como os resultados nacionais.

Para além desta avaliação quantitativa procede-se, todos os anos, a uma avaliação qualitativa das escolas.

O conjunto das avaliações realizadas traduz-se num relatório trienal dirigido ao governo sobre o estado do ensino. O governo dispõe de um ano para dar a conhecer as medidas que decide tomar, em função deste relatório.

Este rico programa de avaliação e de pilotagem trata, ao mesmo tempo, dos processos e dos resultados e alia, harmoniosamente, os aspectos quantitativos e qualitativos. Como conclui A. Ruby (1991:19): "É um programa bem integrado, em que as responsabilidades e o papel da autoridade estão perfeitamente definidos, e que permite uma grande liberdade; esta, ajuda a ter em conta as variações e diferenças locais, permitindo avaliar os progressos na obtenção dos objectivos definidos a nível nacional".

6. AUSTRÁLIA

A Austrália possui uma grande experiência no campo dos estudos avaliativos sobre educação. Foi um dos primeiros países do mundo a possuir um centro nacional de investigação pedagógica o *Conselho Australiano de Investigação em Educação* (ACER), fundado em 1930. Desde muito cedo que o ACER começou a organizar amplos *surveys* sobre o rendimento escolar. O primeiro (1931) foi sobre a leitura e a aritmética; foram testados cerca de 40 000 alunos de setecentas escolas diferentes.

Os Estados que constituem a Commonwealth Australiana gozam de uma ampla autonomia, concretamente no domínio da educação. Na prática, actualmente, todos desenvolveram o seu próprio sistema de pilotagem. Baseada na experiência assim adquirida, a *Conferência Australiana dos Directores Gerais da Educação* elaborou um projecto nacional capaz de avaliar os progressos na educação, sobretudo quanto aos resultados obtidos.

6.1. Pilotagem das escolas e dos sistemas escolares nos Estados

- O Estado de Nova Gales do Sul submete a testes de língua materna, matemática e ciências todos os alunos de 4º ano do secundário. Desde 1989 que uma amostra de alunos do 3º ano, e todos os alunos do 6º ano da primária, são submetidos a testes de leitura e aritmética.
- Desde 1988 que o Estado de Vitória submete a testes de leitura, expressão escrita e matemática, uma amostra de todos os alunos de dez e catorze anos. Os alunos respondem também a um questionário relacionado com a qualidade da vida na escola.

- Na Austrália Ocidental são elaborados testes de rendimento na leitura, na expressão escrita e na matemática e sobre o currículo do ensino secundário inferior, tendo em conta os currículos locais, a fim de permitir uma pilotagem formativa.

No mesmo Estado, são recolhidos indicadores destinados a dar informações sobre a saúde do sistema (T. Wyatt, 1988:15):

1. *Indicadores contextuais*

- Percentagem de tempo escolar dedicado às diversas componentes do currículo (língua materna, matemática, etc.), no 3º ano da primária e no 2º, 4º e 6º ano do ensino secundário.
- Percentagem de alunos matriculados nas escolas de acordo com o sexo e a situação geográfica.
- Taxa de mobilidade dos professores.

2. *Recursos*

- Qual a percentagem e qual o montante do orçamento de Estado consagrado à educação.
- Vencimento dos professores nos diversos momentos da sua carreira.
- Despesas correntes por aluno.

3. *Processos*

- Número de dias de formação contínua ao dispor dos professores e número de dias, efectivamente utilizados para este fim por escola, por tipo de escola, e de acordo com a situação geográfica.
- Número de escolas que organizaram uma estrutura participativa comunitária (por tipo de escola e de acordo com a situação geográfica).

4. *Resultados*

- Distribuição das notas do 2º e 4º ano do ensino secundário, no que toca a certos aspectos cruciais dos currículos de língua materna, matemática e outras disciplinas julgadas importantes.
- No 6º ano do secundário, percentagem de alunos capazes de serem admitidos no ensino superior.

- Taxas de absentismo nas escolas primárias e secundárias de acordo com o sexo, a origem étnica e a situação geográfica

A fim de se poder dispor de indicadores de rendimento, as actividades ou os projectos de que o ministério tomar a iniciativa, deverão vir acompanhados:

- De uma definição operacional dos objectivos.
- Da lista dos indicadores de rendimento a recolher.
- De uma definição de competência mínima.
- De uma especificação dos dados a recolher, das análises a fazer e das modalidades de relatório a fornecer.
- De uma descrição da utilização dos recursos disponíveis.
- De um calendário.

Este modelo geral deveria permitir abordagens quer formativas quer sumativas.

5.2. Projecto nacional

Este projecto começou por uma investigação sobre os indicadores que pudessem ser utilizados a nível nacional: a sua adequação, os problemas suscitados pela sua interpretação, os custos da operação e as vantagens que os participantes no projecto poderiam retirar dele, foram cuidadosamente analisados.

Estes diferentes aspectos foram temas de monografias amplamente divulgadas. Tratam, concretamente, dos conceitos, das definições e dos problemas que se põem sobre este assunto, das lições a retirar da experiência relacionada com os indicadores sociais. Foram discutidos diferentes indicadores tanto no que diz respeito à sua natureza como à sua potencial utilidade.

Desde 1990 que a Austrália publica anualmente um **Relatório Nacional sobre a Situação da Educação**. Deve servir para a pilotagem:

- do sistema educativo global;
- do percurso escolar dos alunos;
- da qualidade dos currículos;
- do rendimento dos alunos;
- da utilização dos recursos orçamentais.

6. OUTROS PAÍSES

Dinamarca, Nova Zelândia, Japão e Singapura contam-se entre os países em que a pilotagem vem conseguindo cada vez melhores resultados. E esta lista de países aumenta continuamente.

Por exemplo na sua publicação *Facts and figures. Education indicators* (1993), a Dinamarca apresenta trinta e sete indicadores (dezassete relativos aos recursos, onze à participação e nove aos resultados). Séries, abrangendo dez anos, indicam as grandes tendências. Os dados relativos à entrada dos jovens na vida activa são também de grande interesse.

• • • • •

INDICADORES INTERNACIONAIS

Considerar o mundo como um laboratório pedagógico onde, numa perspectiva experimental, as características dos sistemas educativos fossem as variáveis independentes e os rendimentos escolares as variáveis dependentes, é a ideia geral que preside às investigações comparativas sobre aquilo que produzem sistemas existentes em contextos sócio-culturais, por vezes muito diferentes. Tais estudos dependem quer da investigação fundamental quer de investigações destinadas à formulação das políticas educativas. Até agora, este segundo anseio redundou, muitas vezes, numa desilusão.

“A ligação entre as descrições do rendimento escolar e os problemas de política educativa é um dos aspectos menos satisfatórios de toda a literatura relacionada com estudos avaliativos internacionais. Muito embora respeitando as dificuldades inerentes à metodologia dos *surveys* em grande escala, os investigadores deviam esforçar-se por conceber os estudos e as análises, de modo a dar resposta às questões que os responsáveis pelas decisões em matéria de educação colocam, efectivamente. Por exemplo, se o interesse for por programas coroados de sucesso – quer dizer, pela política educativa e pelas condições que parecem andar associadas a rendimentos elevados – então devem ser introduzidos, na concepção da investigação, métodos apropriados e questões adequadas. Contudo, práticas coroadas de sucesso em determinado país não produzem, necessariamente, os mesmos resultados noutro país qualquer. As diferenças culturais e os modos de conceber a aprendizagem e o ensino surgem como determinantes poderosos. *Surveys* internacionais podem e devem ajudar a descobrir aspectos favoráveis do ensino e da aprendizagem que não sejam de

exclusivo uso nacional e que, por conseguinte, as políticas educativas possam ter em conta. Podem também existir aspectos que apenas interessem a subconjuntos de países." (U.S. Department of Education, 1992:45-46).

Em qualquer dos casos, sabe-se hoje em dia que os estudos internacionais e interculturais de rendimentos escolares podem produzir dados que constituem indicadores cada vez mais importantes sobre a capacidade das nações para conseguirem educar os indivíduos e prepará-los a desempenhar o seu papel no mundo do trabalho e da cidadania. Como sublinham N. Bottani e H. Walberg (1992:9), o objectivo não é classificar os países entre si mas facultar aos responsáveis pelas decisões políticas uma compreensão mais profunda dos factores que influenciam a qualidade da educação, de modo a poderem explorar uma gama maior de opções políticas.

Chegados a este ponto recordemos que recolha de indicadores e pilotagem não são sinónimos. Indicadores são informações úteis à pilotagem, isto é, em última análise, à tomada de decisões que se devem traduzir em acções. Nada garante que a existência de indicadores, ainda que da melhor qualidade, conduza à acção.

1. ASSOCIAÇÃO INTERNACIONAL DE AVALIAÇÃO DO RENDIMENTO ESCOLAR (IEA)

Surgida no final dos anos cinquenta por iniciativa de uma dúzia de especialistas de investigação experimental em educação, a IEA é hoje, com mais de cinquenta membros, uma organização mundial de investigação avaliativa.

O objectivo inicial era controlar objectivamente as asserções, se não mesmo os estereótipos relativos às virtudes e aos defeitos dos diversos sistemas educativos. Daí a ideia de uma avaliação comparativa dos rendimentos escolares recorrendo a estudos normativos (*surveys*). O desenvolvimento da educação comparada que, com a IEA, segue uma orientação experimental, e o aparecimento dos computadores, têm muito a ver com este empreendimento pedagógico que implica o tratamento de milhões de dados.

Desde 1961 que a IEA lança uma nova avaliação internacional, praticamente de três em três anos, de modo a completar um ciclo que, cada doze anos, abranja as disciplinas principais. Para a macropilotagem é fundamental o interesse destes dados comparativos.

Um estudo utiliza os seguintes instrumentos:

- Testes de conhecimento elaborados após uma análise comparativa dos países participantes. A matéria das provas internacionais é constituída pelo

que lhes for comum e for considerado importante. Os participantes ficam com liberdade para testarem, ao mesmo tempo, matérias específicas ou que, por alguma razão, privilegiem.

- Testes de rendimento efectivo: experiências laboratoriais, manipulações, observações, produções,...
- Um questionário internacional concebido para recolher grande número de informações sobre os alunos: sexo, idade, ano de estudos, meio familiar, número de horas de ensino na disciplina em questão, duração dos trabalhos em casa, tempo dedicado à televisão,...
- Testes de aptidão verbal, numérica, ...
- Escalas de atitudes que mostrem como os alunos encaram a disciplina e o modo como ela é ensinada e, em sua opinião, a importância que ela tem para a sociedade.
- Os professores respondem a um questionário sobre a sua idade, formação inicial e contínua. Relativamente a cada questão dos testes são convidados a referir se a matéria sobre a qual ela versa foi ensinada (recentemente ou não, de forma intensiva ou não, ...)
- Os responsáveis pela direcção dos estabelecimentos de ensino respondem a um questionário sobre a população de cada uma das suas escolas, sobre o equipamento, o financiamento, as relações com a comunidade educativa,...
- Questionário dirigido às autoridades organizadoras.

1.1. Primeiros tempos da IEA: Um estudo de exequibilidade

Em 1959, foi lançado um projecto piloto que durou três anos. Nele participaram a Inglaterra, Bélgica, Escócia, Estados Unidos, Finlândia, Israel, Polónia, República Federal Alemã, Suécia, Suiça e Jugoslávia. Abrangeu as crianças de treze anos, idade ainda de escolaridade obrigatória em todos estes países. Deste modo, 9918 alunos falando oito línguas diferentes foram submetidos a testes de compreensão da leitura, de matemática, de ciências, de geografia e de inteligência não verbal. Os resultados desta investigação foram suficientemente ricos para justificarem projectos mais ambiciosos ainda.

1.1.1. Estudo sobre o rendimento do ensino da matemática (1961-1967)

Doze países, entre os quais a França e a Bélgica, participaram neste primeiro estudo normativo abrangendo três populações de alunos: alunos com treze anos

de idade, alunos a estudar no nível escolar normalmente previsto para as crianças de treze anos, alunos a frequentar o ano que precede o ensino superior.

Numa altura em que o ensino da matemática se modificava profundamente, devido à introdução da teoria dos conjuntos, havia grande interesse em dispor de dados que permitissem posteriores comparações.

A colheita revelou-se rica em conclusões relativas:

- aos sistemas de ensino: duração da escolaridade, número de alunos por turma, programas de estudos, efeitos da selecção, avanço de certos países em relação a outros;
- à formação de professores;
- aos factores sociais do sucesso escolar;
- às diferenças entre sexos.

Foi por esta altura que o Japão se revelou, de longe, o país com melhor rendimento escolar, sobretudo no ensino secundário de longa duração, mesmo nos cursos em que a matemática não era a disciplina principal.

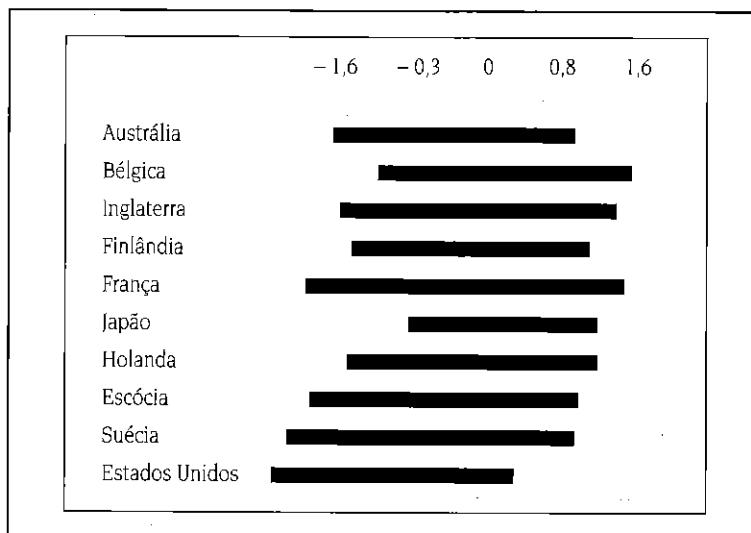
O quadro seguinte, que mostra a situação dos países em relação uns aos outros, é bem eloquente (G. de Landsheere e T.N. Postlethwaite, 1969:36). Observa-se uma grande homogeneidade nos resultados obtidos pelas quatro populações de alunos, em todos os países.

Quadro 10
Médias nacionais situadas em relação à média geral
(em unidades de desvio-padrão)

	Alunos com 13 anos de idade	Ano em que se encontra a maioria dos alunos com 13 anos de idade	Fim do secundário superior	
			Matemática como disciplina principal	Matemática
1,25	• Inglaterra	• Inglaterra		
	• Japão	• Japão	• Japão	• Japão
	• Holanda			
	• Bélgica	• Escócia		
1,00	• Escócia	• Israel	• Inglaterra	
	• Austrália	• Bélgica	• Bélgica	
	• Estados Unidos	• Estados Unidos	• Estados Unidos	
	• França	• França	• Suécia	• Inglaterra
0,75		• Austrália	• França	
		• Holanda	• Austrália	
		• Alemanha	• Alemanha	
			• Escócia	• Escócia-Bélgica
0,50	• Suécia	• Suécia	• Alemanha	• Estados Unidos
	• Finlândia	• Finlândia	• Finlândia	• Finlândia
			• Israel	• Alemanha
			• Holanda	• Suécia

A observação do aumento do rendimento do ensino da matemática nos cursos em que esta é a disciplina principal, entre os treze anos e o fim do ensino secundário, é também muito interessante.

Figura 7
Matemática como disciplina principal



O Japão parece progredir menos do que os outros países, mas é por partir de um nível mais elevado (mais rapidamente atinge o tecto máximo). A Bélgica apresenta um progresso espectacular devido ao elevado grau de selecção espontânea que caracteriza a escolha dos cursos em que a matemática é disciplina principal.

Um estudo realizado no Japão para tentar explicar os seus brilhantes resultados, apesar do número geralmente muito elevado de alunos por turma (muitas vezes mais de quarenta), revelou duas características principais:

- Um ensino impositivo muito exigente.
- Uma pressão contínua exercida sobre os alunos pelos pais, professores e sociedade civil.

Foram consideráveis as consequências deste estudo: aperfeiçoamento das técnicas de amostragem e de análise estatística, emergência de uma educação comparada experimental, constituição de uma vasta rede cooperativa internacional

dinamizada pelos mais avançados investigadores em educação na altura, incremento da investigação experimental em educação nos países participantes.

Sobre esta investigação, ver T. Husén (1967), G. de Landsheere (1967), G. de Landsheere e T.N. Postlethwaite (1969), G. Henry (1970).

1.1.2. Survey de seis disciplinas (1966-1973)

Este estudo, no qual participaram vinte e dois países, teve como objecto de estudo a compreensão da leitura, a literatura, as ciências, o inglês ou o francês ensinados como línguas estrangeiras, a educação cívica. A Bélgica francófona tomou parte nos trabalhos sobre ciências, leitura, literatura, inglês como língua estrangeira, educação cívica.

Foram testadas amostras de alunos de dez e catorze anos de idade e de alunos a frequentar o ano pré-universitário. Foram estudadas duzentas a quinhentas variáveis para cada disciplina e para cada um dos universos escolhidos, e a quantidade de dados ultrapassou os cento e cinquenta milhões. Serviram para constituir, em seguida, um banco de dados que foi posto ao dispor do mundo da investigação.

A repercussão desta investigação foi, e continua a ser, considerável. Em muitos países participantes a opinião pública fez suas as conclusões dos trabalhos e, à luz dessas mesmas conclusões, foram desencadeadas reformas dos programas escolares, especialmente na Hungria e na Suécia.

— *Investigação sobre o ensino das ciências*
Exemplo de uma observação grave para a Bélgica

Por ocasião desta investigação, verificou-se um fenómeno preocupante: no domínio dos conhecimentos os alunos belgas revelaram-se fracos em termos comparativos e, sobretudo, o seu interesse pelas ciências e pelas actividades científicas diminuía à medida que avançavam nos estudos.

Na escala concebida para avaliar o interesse, um resultado de +15 corresponde a um interesse muito elevado, um resultado -15 a um interesse muito fraco. A consulta do quadro que a seguir se apresenta parece revelar, à primeira vista, que o interesse é tendencialmente fraco em todos os países. A comparação entre eles revela, contudo, grandes diferenças.

Quadro 11
Interesse pelas ciências

País	População I		População II		População III	
	Média	Desvio-Padrão (sigma)	Média	Desvio-Padrão (sigma)	Média	Desvio-Padrão (sigma)
Alemanha	-1,5	-4,7	-3,7	-4,4	-4,5	-4,4
Inglaterra	-3,6	-4,2	-5,0	-4,5	-5,1	-4,6
Austrália	—	—	-4,1	-4,5	-4,6	-4,5
Bélgica (Fl.)	-3,0	-4,3	-5,2	-4,2	-5,3	-4,1
Bélgica (Fr.)	-3,8	-4,1	-4,7	-4,2	-6,1	-4,5
Escócia	-3,4	-4,1	-5,9	-4,7	-5,7	-4,7
Estados Unidos	-1,6	-4,7	-3,9	-4,6	-4,7	-4,7
Finlândia	-4,2	-3,8	-6,2	-3,9	-5,5	-4,1
França	—	—	—	—	-3,6	-4,4
Holanda	-3,1	-4,1	-5,6	-4,1	-5,6	-4,3
Hungria	-2,2	-4,2	-1,9	-4,2	-3,0	-4,3
Itália	-3,8	-4,0	-3,5	-4,2	-4,2	-4,3
Japão	-1,2	-4,7	-4,2	-4,7	—	—
Nova Zelândia	—	—	-4,3	-4,2	-4,0	-4,7
Suécia	-3,9	-3,4	-4,9	-3,7	-5,2	-4,1

G. Henry (1976:155-157) conclui: também aqui, a Bélgica francófona se apresenta em situação pouco invejável: nona em doze na População I, oitava em catorze na População II, e décima quarta em catorze na População III.

— *Investigação sobre literatura*

Esta investigação internacional comparativa sobre o estudo da literatura constitui uma estreia. Participaram dez países. As populações examinadas eram constituídas por alunos de catorze anos e pelos que frequentavam o ano pré-universitário.

Ambos os grupos foram submetidos a um teste de escolha múltipla, sobre duas pequenas novelas (uma comum a todos, e a outra escolhida ao acaso entre três). Os alunos responderam, ainda, a um questionário sobre o seu interesse pela literatura, as suas preferências e, de uma maneira geral, sobre o seu interesse pela leitura.

A diversidade de culturas e de patrimónios literários, não permitiu a elaboração de um teste comum. Daí a actuação ter incidido, essencialmente, sobre os

aspectos cognitivos e afectivos das reacções dos alunos perante a literatura e sobre os seus interesses e expectativas.

Ver A.C. Purves, 1973.

– *Investigação sobre educação cívica*

Neste momento em que as sociedades ocidentais atravessam um período grave no que se refere a valores tradicionais, e em que estes começam a ser postos em causa, pensamos que é importante avaliar os conhecimentos e atitudes sócio-políticas dos jovens cidadãos do futuro. Foram submetidas a testes duas populações de jovens: alunos do 3º e do 6º ano do ensino secundário.

Foram utilizados testes de escolha múltipla sobre os temas seguintes:

- Fundamento e natureza da cidadania.
- Instituições políticas – temas nacionais.
- Processos políticos – temas nacionais.
- Processos e instituições políticas – temas internacionais.
- Processos e instituições económicas.
- Instituições e processos sociais.

Foi concebida uma versão diferente dos questionários para cada uma das populações. Alguns *itens*, porém, eram comuns a fim de dispor de uma abordagem que permitisse fazer comparações.

Foram também apresentadas escalas afectivas. Tratavam de aspectos como: Popular/Impopular; Digno de confiança/Não digno de confiança; Forte/Fraco; Simpático/Antipático.

Por fim, os alunos responderam a um questionário de *background* pormenorizado.

Entre as observações levadas a cabo, há a notar:

- Pouca diferença entre sexos.
- Uma correlação positiva geral e elevada, entre os resultados dos testes e o nível sócio-cultural de origem. Existe também uma relação do mesmo tipo com o nível de aspiração dos alunos (profissão que esperam vir a exercer; duração dos estudos que contam fazer). É também determinante a frequência das discussões que os jovens têm em família, relativas aos problemas da vida social e política.

Esta investigação confirma a importância do papel que a escola pode desempenhar na preparação do futuro cidadão. "A informação e formação cívicas só se conseguirão realizar, verdadeiramente, se se permitir a abordagem franca de temas sociais e políticos, com a condição, é claro, de se garantir um verdadeiro

pluralismo ideológico. (...) Para se atingir este objectivo não bastam apenas algumas orientações. Há que elaborar um currículo diversificado, tanto no que respeita a conteúdos como a métodos. E, sobretudo, há que ter com as atitudes um cuidado pelo menos tão grande como o que se tem com os conhecimentos" (Massoz e Henry, 1980: 202).

Sobre os *surveys* das seis disciplinas ver ainda J.B. Carroll (1975), L.C. Comber e J. Keeves (1973), A. Grisay e G. Henry (1973), A. Grisay (1974), G. Henry (1974), G. Henry (1975), G. Henry (1976), A. Grisay (1976), M.T. Loret (1980), A.C. Purves (1973), R.L. Thorndike (1973), J.V. Torney *et al.* (1976), D.A. Walker (1976).

1.1.3. Segundo estudo sobre o ensino da matemática (1980-1986)

Os objectivos deste estudo eram:

- Descrever os programas de matemática aplicada nos sistemas escolares e verificar as modificações ocorridas desde 1960.
- Verificar em que medida o currículo previsto estava a ser, efectivamente, implantado.
- Avaliar o rendimento escolar em matemática e estudar a influência relativa de diferentes determinantes.
- Avaliar o aumento de conhecimentos durante um ano escolar e investigar as razões das diferenças deste aumento nos alunos, considerados individualmente e nas turmas.
- Recolher, de forma intensiva e extensiva, informações sobre a didáctica aplicada.

Foram analisados os currículos dos países participantes, em especial para identificar os objectivos a atingir e avaliar se tinham sido alcançados. Examinou-se, também, o modo como a matemática era ensinada.

A Bélgica participou nesta investigação ficando assim na posse de uma referência comparativa preciosa.

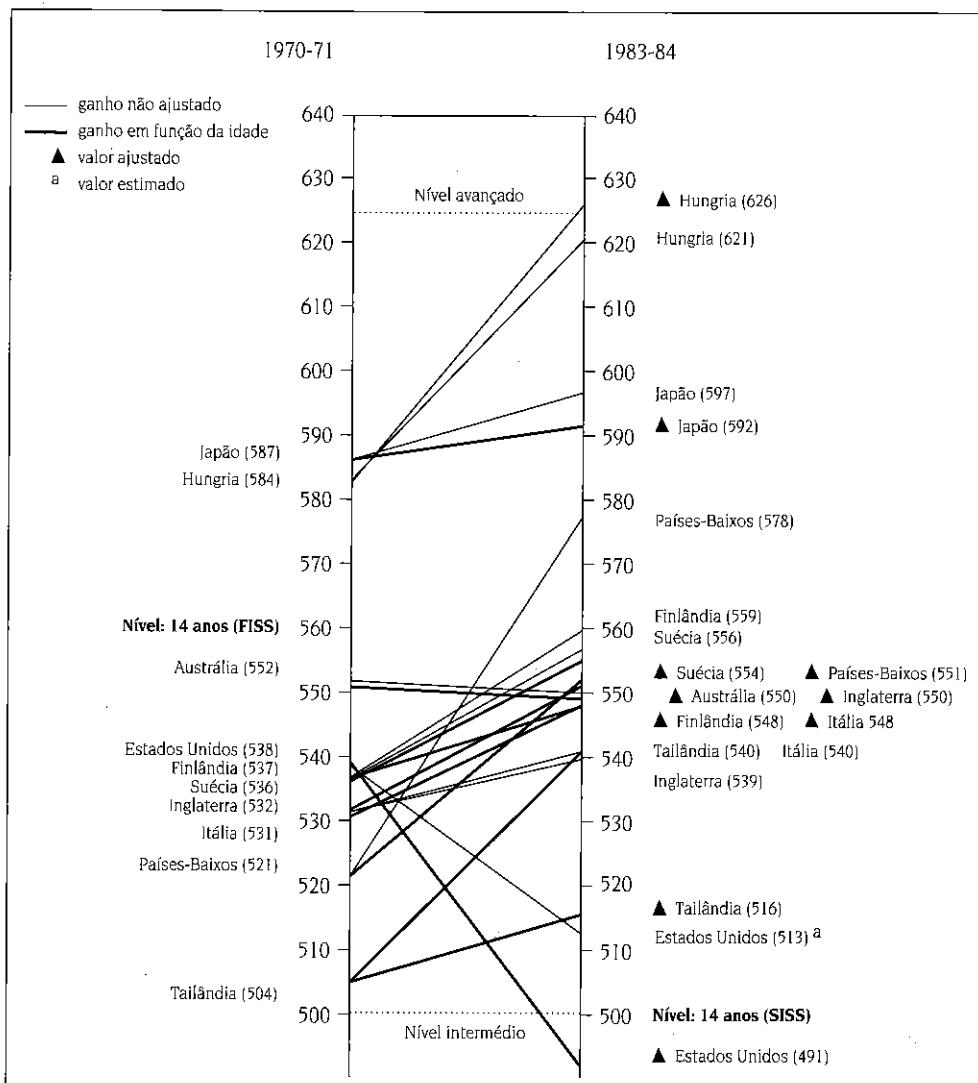
Publicação: N. Deltour e G. Henry (1988).

1.1.4. Segundo estudo sobre ciências (1984)

Os principais objectivos deste estudo eram observar, com atenção, a evolução do ensino das ciências no mundo (sobretudo a partir de 1970, ano da primeira avaliação) e determinar os factores que explicam as diferenças de rendimento e de outros produtos do ensino das ciências.

Por falta de verba, a Bélgica não pode participar neste estudo (Ver M. Rosier e J. Keeves, 1991). Facto mais lamentável ainda quando, como mostra o gráfico seguinte (J. Keeves, 1992:14), surgem fenómenos particularmente importantes numa altura em que as ciências se expandem e têm um peso cada vez maior no futuro das nações. O gráfico permite, de facto, a comparação do rendimento do ensino das ciências a alunos de catorze anos, entre 1970-71 e 1983-84.

Figura 8
Resultados médios em ciências
Alunos de 14 anos: 1970-71 e 1983-84



Quer com resultados ajustados, quer com resultados não ajustados em função da idade, verifica-se sempre um grande progresso na Hungria. Este país reformulou, profundamente, o seu programa de ensino das ciências em função dos resultados observados em 1970-71; concretamente foi muito aumentado o tempo dedicado a trabalhos práticos de laboratório.

Os resultados do Japão continuam muito elevados, mas os progressos são menos espectaculares do que na Hungria. Talvez tenha começado a esboçar-se, nesta altura, o fenómeno de alcance do tecto máximo que se observa, actualmente, no Japão. Talvez tenha sofrido os inconvenientes de um ensino demasiado impositivo, num momento em que o indefinido, o mais ou menos indeterminado, desempenham uma crescente função interventora, e em que, portanto, a divergência adquire uma importância cada vez maior.

Por outro lado, e contrariando a tendência geral, os Estados Unidos estão em regressão acentuada. A verificação deste facto despertou grande inquietação no país e influenciou, profundamente, as tomadas de posição espectaculares do Presidente Bush, ficaram expressas numa publicação que deu brado – *A Nation at Risk* – a que já fizemos alusão.

Este exemplo mostra bem o proveito e os ensinamentos que podem retirar-se de uma pilotagem apoiada em dados comparativos internacionais.

1.1.5. Estudo sobre a utilização pedagógica dos computadores (1987-1990)

OBJECTIVOS:

- Determinar como e em que medida os computadores são usados e com que resultados.
- Observar a evolução dos materiais e dos programas.
- Observar como são formados os professores para a utilização pedagógica do computador.

Dezanove países, entre os quais a Bélgica, participaram nesta avaliação. É evidente que, até agora e de uma maneira geral, se continua a recorrer pouco ao computador na aprendizagem escolar (Ver H. Pelgrum e T. Plomp, 1991).

1.1.6. Estudo sobre a educação pré-escolar: qualidade de vida das crianças de quatro anos (1986-1996)

Esta investigação assinala uma viragem nas opções metodológicas da IEA. De facto neste caso é a perspectiva qualitativa que domina. Além disso, enquanto os

estudos efectuados sobre as escolas primárias ou secundárias tratavam de meios institucionalmente homogéneos, a investigação sobre a educação pré-escolar deve ter em conta, não só a grande heterogeneidade das soluções encontradas para acolher as crianças num meio extra-familiar, como também o facto de esse local de acolhimento poder variar várias vezes num mesmo dia. Será possível determinar o impacto diferencial destes diversos meios no desenvolvimento da criança? Foi escolhida uma população de crianças de 3/6 a 4/6 anos de idade.

OBJECTIVOS:

- Determinar as necessidades em modalidades de acolhimento das crianças de quatro anos e observar a utilização que é feita das possibilidades existentes.
- Determinar quais os factores (culturais, sócio-económicos,...) que influenciam a escolha dos pais.
- Quais as características dos meios de acolhimento? Qual a qualidade de vida das crianças nestes locais?
- Em que medida os valores e expectativas dos adultos influenciam a organização dos locais de acolhimento e as actividades que aí decorrem?
- Que diferenças se observam entre países?

Verifica-se, por exemplo, que o tempo diário que o pai passa sozinho com o filho é de 0,30 horas na Bélgica, 0,42 horas nos Estados Unidos, 0,48 horas na Finlândia e 0,54 horas na China. Outra observação impressionante: na Bélgica, 40% das crianças de quatro anos passam diariamente por dois locais de acolhimento diferentes, para além da sua própria casa (4,4% na Tailândia, 16% nos Estados Unidos,...). Estamos a falar, evidentemente, de médias.

- Quais as consequências sobre o desenvolvimento da criança?
- Em que medida as crianças são preparadas para a escolaridade?

A investigação decorreu em três fases:

- Fase 1: Procurar saber, em especial, o modo como as famílias utilizam as diferentes possibilidades de acolhimento dos filhos.
- Fase 2: Será estudada a fundo uma amostra dos diferentes meios de acolhimento, identificados durante a Fase 1, a fim de avaliar a qualidade de vida das crianças nesses locais.
- Fase 3: Estudo contínuo das crianças de sete anos: como se desenvolveram?

Ver P. Olmsted e D. Weikart, 1989.

1.1.7. Investigação sobre compreensão da leitura (1990-1993)

Esta investigação abrangeu duas populações: nove e catorze anos; foram submetidos a testes 210 000 alunos em trinta e dois países. A idade de nove anos surgiu como particularmente interessante pois será legítimo pensar que quatro anos de escolaridade bastem, normalmente, para assegurar o domínio elementar da leitura. Há que insistir na palavra *elementar*, pois o ensino sistemático da leitura deveria continuar durante os estudos secundários. Foi escolhida a idade de catorze anos por ser uma idade em que o ensino é ainda obrigatório, em todos os países participantes.

Os objectivos gerais eram os seguintes:

- Produzir instrumentos válidos para a avaliação da capacidade de leitura (nível de literacia).
- Fazer uma estimativa da percentagem de cada população que atinge determinado nível em leitura: elementar – intermédio – elevado.
- Determinar a importância das actividades de leitura nas duas populações.
- Fazer uma estimativa da qualidade dos programas de ensino da leitura nos países participantes.
- Determinar a relação entre o nível de capacidade de leitura e os indicadores económicos.
- Estabelecer, em 1990, uma base para a pilotagem nos países participantes.
- Comparar as práticas de ensino.

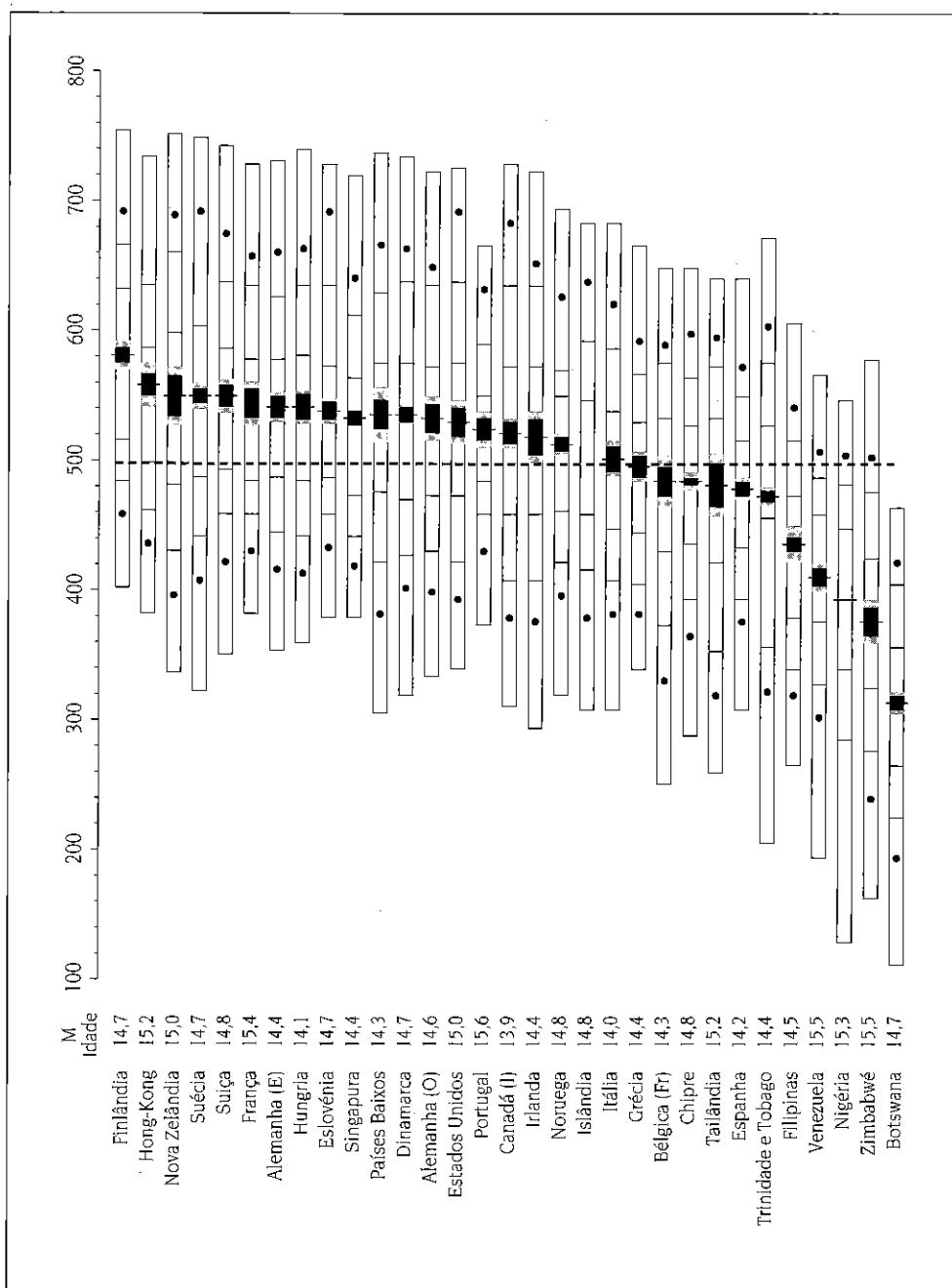
Os testes versaram sobre três tipos de textos:

- Narrações: curtas histórias em que intervinham animais.
- Exposições. Por exemplo, descrição da areia movediça, seguida de perguntas sobre o modo de a reconhecer.
- Documentos. Leitura de informações estruturadas em forma de gráficos, de quadros, de mapas, modos de utilização.

Os resultados obtidos pela população de alunos de nove anos da Bélgica francófona, situam-se dentro da boa média; pelo contrário, os resultados da população de catorze anos são inquietantes, no que respeita aos três tipos de textos (Ver W.B. Elly, 1992).

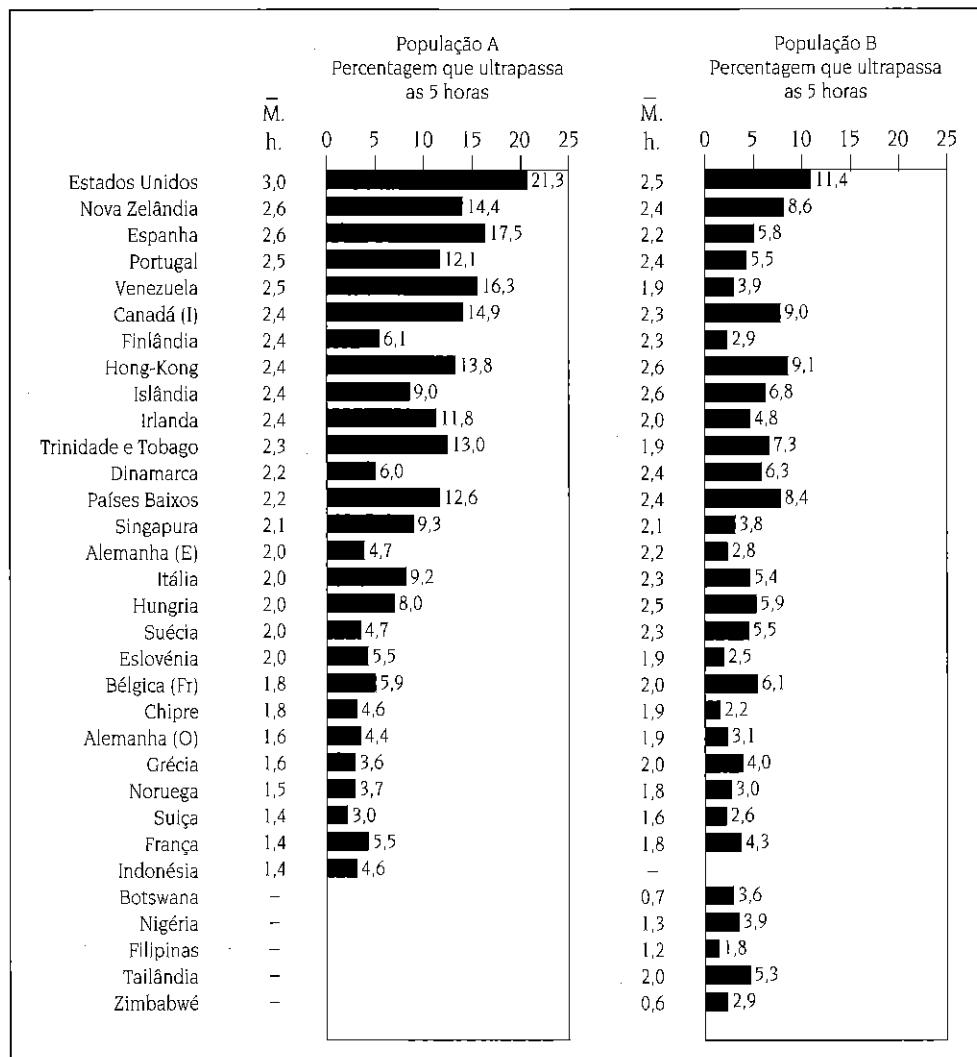
O gráfico seguinte mostra a distribuição dos resultados da leitura de textos narrativos aos catorze anos de idade.

Figura 9
Distribuição dos resultados da leitura de documentos informativos
(quadros, imagens, mapas,...) – Alunos de catorze anos



Há que haver uma explicação para observações desta natureza. Onde é que, eventualmente, falham os métodos de ensino? Entre outras explicações possíveis, deparamos com a questão do número de horas dedicado diariamente à televisão. O quadro seguinte indica a percentagem de jovens que vêem mais de cinco horas por dia de televisão.

Figura 10
Média do número de horas consagradas diariamente à televisão
e percentagem de alunos que declaram ver televisão
mais de cinco horas por dia



As análises estatísticas revelam a existência de uma correlação negativa entre a capacidade de leitura e uma frequência exagerada de televisão. Pelo contrário, parece ser vantajoso ver televisão menos de três horas e meia por dia. É o que se verifica, concretamente, na Finlândia.

Numa obra que é o prolongamento desta investigação (*Effective Schools in Reading*), T.N. Postlethwaite e K. Ross (1992) descreveram as escolas que ensinam a ler da forma mais eficaz e formularam um conjunto de propostas para os responsáveis por planos educativos.

1.1.8. Terceiro estudo sobre o ensino da matemática e das ciências (1991-1995)

Entre trinta a quarenta países participaram nesta investigação, cujos instrumentos estão em vias de elaboração.

1.2. Papel histórico e contribuições metodológicas da IEA

A IEA fez obra de precursor. Fruto da sua época, acumulou grande quantidade de dados comparativos e de estudos, de que não souberam aproveitar-se plenamente, longe disso, os poderes organizadores e o mundo da educação em geral. Fenómeno que se explica, em boa parte, pela tecnicidade dos relatórios e pela insuficiente divulgação na maior parte dos países. Além disso, na altura em que a IEA iniciou os seus estudos, poucos centros nacionais dominavam os conhecimentos e as técnicas estatísticas, ou doutro tipo, necessárias à execução dos trabalhos. Sobre este aspecto a IEA levou a cabo uma acção de formação e de estímulo de valor incalculável.

A rede científica e social desenvolvida pela IEA à escala mundial, contribuiu muito para o progresso da investigação avaliativa em educação e da pilotagem dos sistemas educativos. É aliás sintomático verificar que foi a dois antigos presidentes da IEA que foi confiada a direcção da redacção das duas edições sucessivas da *International Encyclopedia of Education: Research and Studies* (Oxford, Pergamon), monumental síntese, em doze volumes, da investigação científica em educação à escala mundial (2^a edição em 1994).

Entre as contribuições mais pertinentes da IEA para a metodologia da investigação avaliativa, podem citar-se como exemplos:

- Ter em conta de forma sistemática a *oportunidade de aprender* (*Opportunity to learn*, OTL) na avaliação das aquisições. Num processo educativo é, de facto, contestável submeter os alunos a exames sobre matérias que não tiveram

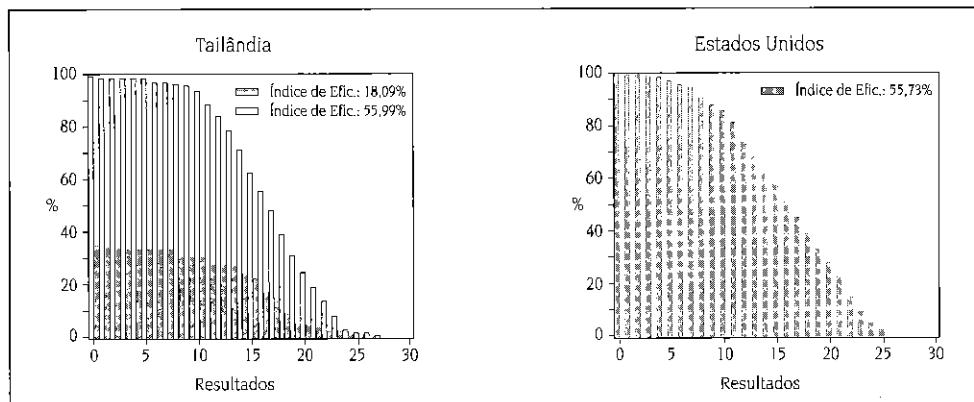
oportunidade de aprender (a não ser que se esteja interessado nas aprendizagens ditas accidentais).

A propósito de cada questãoposta, é esclarecedor saber se ela trata de noções que, na opinião dos professores, foram ensinadas ou revistas durante o ano lectivo em curso, durante os anos anteriores, ou nunca.

- A noção de *produtividade cognitiva* (*productivity yield*). O índice de produtividade corresponde à relação entre a taxa de respostas correctas e a percentagem escolarizada do grupo em idade escolar. Por exemplo, se um teste de trinta *itens* for administrado a uma amostra representativa de um grupo etário totalmente escolarizado, e se todos os alunos responderem correctamente, a produtividade cognitiva será de 100%.

T.N. Postlethwaite e D. Wiley (1991) analisaram os resultados em ciências, obtidos na Tailândia e nos Estados Unidos, por ocasião da "Investigação das seis disciplinas" (1966-1973). A percentagem de respostas correctas aos *itens* foi a mesma em ambos os países. Porém, na Tailândia, apenas um terço dos jovens de treze anos de idade estavam escolarizados, quando nos Estados Unidos a escolarização era de 100%. O gráfico seguinte revela bem a diferença de produtividade. Aparentemente, o índice de eficácia da Tailândia (55,99%) é praticamente o mesmo que o dos Estados Unidos (55,73%), mas na Tailândia a produtividade real é apenas de 18,09%.

Figura 11
Produtividade cognitiva



- Ter em consideração a *duração da implicação efectiva na tarefa* (*time on task*). Foi durante a investigação sobre o ensino de línguas estrangeiras (1966-1973), que J.B. Carroll fez uso sistemático deste índice. De facto, não é o número de horas de ensino previstas no programa de uma disciplina que possibilita

uma previsão do que irá ser a sua aprendizagem, mas antes o tempo que o aluno dedica, realmente, ao trabalho.

1.2.1. Um processo inevitavelmente longo

As investigações da IEA, acabadas de expor, são penosas e demoradas.

São múltiplas as causas deste fenómeno:

- Trata-se da execução de um projecto cooperativo, financiado (às vezes com muitas dificuldades e incertezas) pelos Estados participantes, que dispõem todos do mesmo poder de decisão aquando das assembleias gerais anuais.
- A dificuldade da análise comparativa dos currículos, destinada a detectar os objectivos e conteúdos de ensino comuns em conformidade com os quais serão elaborados os primeiros esboços dos instrumentos de avaliação. Cada participante é convidado a propor *itens*. Há sempre a possibilidade de acrescentar opções locais aos instrumentos internacionais; é também possível ter em conta os objectivos específicos de determinado país.
- As repetidas experiências e sucessivas revisões dos instrumentos de avaliação.
- Rapidamente começam a surgir as deficiências científicas de certos participantes, que não possuem nem formação aprofundada em pedagogia experimental ou em odometria, nem experiência de avaliações objectivas em grande escala. É fornecida assistência técnica e as aprendizagens que esta suscita revestem-se, por vezes, de mais importância para os países, do que os resultados da investigação em curso.
- Dificuldade em fazer respeitar as mesmas regras de amostragem das populações a investigar, tal como foram definidas. Apesar da existência de peritos em amostragem, há países – mesmo entre os mais avançados – que cometem erros que alteram os resultados.
- A dificuldade da experimentação geral dos instrumentos, da recolha de dados e, em geral, de todas as operações a executar. Devem retirar-se lições desta experiência.
- Como os trabalhos decorrem à escala mundial, pode haver desfasamentos de seis meses entre os anos escolares dos vários países. Facto que a organização deve ter em conta.
- Quando os dados definitivos chegam ao centro internacional, há que proceder à sua “limpeza”, o que, à escala de milhões de dados, não é tarefa fácil. Para facilitar esta operação foram concebidos recentemente programas informáticos sofisticados.

- Segue-se o tratamento estatístico que permite detectar as dominantes que não oferecem dúvidas e relacionar as variáveis entre si. A identificação das variáveis a ter em consideração é, também só por si, um acto importante e, muitas vezes, complexo.
- Os resultados são, em seguida, enviados a cada país participante para que faça deles uma interpretação tão clara quanto possível.
- Depois de muito ir e vir são finalmente redigidos os relatórios. São apresentados aos participantes antes de serem publicados.
- Estas publicações assumem várias formas, desde grandes volumes destinados ao mundo científico e aos especialistas em administração superior de sistemas educativos, até edições simples concebidas para grande divulgação

Assim se explica que uma investigação da IEA leve sempre vários anos.

2. INTERNACIONAL ASSESSMENT OF EDUCATIONAL PROGRESS (IAEP)

O *Educational Testing Service* (ETS), que já tinha a seu cargo a *National Assessment of Educational Progress* (NAEP) dos Estados Unidos, tomou a iniciativa de fazer investigações internacionais rápidas. Se bem que, metodologicamente, não sejam comparáveis às que a IEA realiza, há quem tivesse visto nisto uma concorrência, uma guerra de influências. De facto, é no momento em que a OCDE decide lançar o seu projecto sobre indicadores do ensino (INES) e vai, deste modo, necessitar de dados internacionais sobre rendimentos escolares, que surge um potencial mercado destes produtos. Esperança vã: o projecto INES é lançado, mas a OCDE, pelo menos provisoriamente, encarregará cada país de fazer a recolha dos dados necessários.

2.1. Estudo IAEP I (1987-1988)

A IAEP procura, explicitamente, alcançar dois objectivos:

- Reduzir a duração e o custo dos estudos internacionais comparativos, recorrendo aos instrumentos e métodos criados pela NAEP, dos Estados Unidos.
- Fazer com que os países interessados se familiarizem com as técnicas utilizadas pela NAEP e, assim, possam ver se elas lhes convêm para os seus próprios projectos de avaliação locais.

Foram feitas propostas neste sentido em 1987, e em 1988 foi decidido e executado um primeiro estudo sobre o rendimento dos alunos de treze anos em matemática e ciências. Participaram neste projecto cinco países: Coreia, Irlanda, Canadá, Espanha, Estados Unidos e Reino Unido. Os gastos com a coordenação e tratamento dos dados deste estudo foram suportados pelo *Department of Education* e pela *National Science Foundation*, dos Estados Unidos.

As questões de avaliação foram retiradas dos 281 *itens* de matemática e dos 188 *itens* de ciências utilizados na investigação da NAEP, em 1986. Peritos provenientes dos países participantes, avaliaram em que medida estas questões correspondiam aos seus currículos nacionais, se podiam ser traduzidas, ou se deviam ser culturalmente adaptadas (alteração das unidades de avaliação, dos nomes das crianças, das plantas e dos animais,...).

As técnicas odometrícias utilizadas no tratamento dos dados são de vanguarda e permitiram ganhar muito tempo.

Em Janeiro de 1989, o ETS publicou os resultados do seu primeiro *survey* (A.E. Lapointe et al., 1989). Entre os dados comparativos, chama-se a atenção, em particular, para uma interessante tentativa de pôr em evidência a relação dos deveres de casa de matemática e o rendimento observado nesta disciplina (ver Fig. 12 na página seguinte).

2.2. Estudo IAEP II (1990-1991)

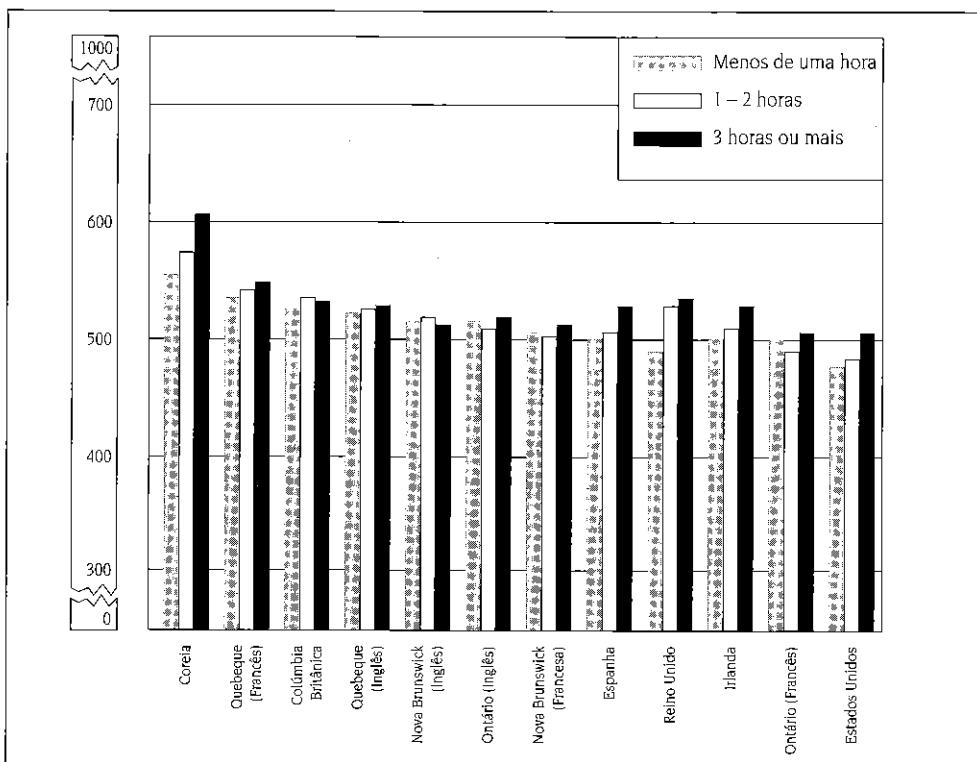
Este trabalho, realizado com o mesmo espírito do primeiro, diz respeito a populações de alunos de nove e treze anos de idade, e abrange as disciplinas de matemática, ciências e geografia. Participaram nele vinte e oito sistemas educativos (nacionais ou provinciais).

Como conclusão deste segundo estudo, o Ministério da Educação dos Estados Unidos espera conseguir alguns progressos nos processos de amostragem e de execução do estudo, informações mais precisas sobre a capacidade de comparação das amostras pré-selecionadas nos diferentes sistemas educativos, e a clarificação de questões que ainda subsistem quanto à interpretação de dados.

2.3. Crítica fundamental

Há uma crítica fundamental que continua a ser feita à IAEP: mesmo procedendo a diversas consultas, não deixa por isso de exportar em menor número, questões cuja concepção é feita em função dos programas dos Estados

Figura 12
Competência média em matemática, em relação com a duração semanal dos deveres de casa nesta disciplina



Unidos. Por ocasião da segunda investigação, em 1991, o ETS modificou, em parte, este estado de coisas, ao convidar os países participantes a proporem *itens* de testes que reflectissem, directamente, os seus respectivos currículos. A abordagem continua, porém, do tipo "centro-periferia". Contrasta, fundamentalmente, com o processo de trabalho "periferia-centro" da IEA e com a intensidade das relações humanas e das acções de formação multidirecionais que o acompanham.

Ao contrário da IEA, a IAEP não procede a uma análise pormenorizada dos factores associados às diferenças observadas entre os países. Até agora, não recolheu em larga escala informações contextuais sobre os sistemas educativos avaliados, o que torna difícil a explicação das diferenças de rendimento observadas.

A IAEP tem trabalhado, até agora, a curto prazo, enquanto que a IEA tem perspectivas a longo prazo e pensa que os efeitos secundários do seu trabalho são, pelo menos, tão importantes ou mais do que as vantagens imediatas.

3. PROJECTO DA OCDE: INTERNATIONAL INDICATORS AND EVALUATION OF EDUCATIONAL SYSTEMS (INES) – INDICADORES INTERNACIONAIS DO ENSINO.

3.1. Primeira experiência: indicadores sociais

Desde os anos sessenta que a Organização de Cooperação e Desenvolvimento Económico (OCDE), manifestou desejo de produzir estatísticas susceptíveis de ajudar à planificação da educação. As atenções estavam sobretudo concentradas nos fluxos das populações escolares, mas não estava ausente a preocupação de elaborar balanços dos sistemas educativos, à semelhança do que se fazia em economia.

Em 1970, a OCDE propôs-se produzir um conjunto – que se pretendia exaustivo – de indicadores sobre o rendimento do ensino, relativamente aos objectivos propostos. Em 1973, um relatório conceptual de base, redigido por R. Carr-Hill e O. Magnussen, faz um apanhado geral dos problemas em causa e analisa os tipos de indicadores a recolher. O interesse histórico desta publicação é grande; a qualidade dos seus colaboradores é garantia de que reflecte, correctamente, o pensamento mais avançado da época. De facto, entre os colaboradores encontram-se: K. Harnqvist (que desempenhará, também, um papel importante na concepção do projecto INES), C. Arnold Anderson, M. Blaug e J. Coleman.

Partiu-se da ideia inicial de que era importante completar os índices económicos gerais, como o PNB (*input*), que não dão uma ideia exacta “do bem-estar dos povos”, com indicadores sociais relativos ao *output*. Um dos objectivos prioritários a atingir é a sugestão de “um quadro geral para um sistema de informação estatístico, capaz de controlar a política de ensino” (p.27).

Um indicador social define-se da seguinte maneira: “Estatística de interesse normativo directo que facilita a formulação de pareceres concisos, exaustivos e equilibrados, sobre o estado dos principais sectores da sociedade. Trata-se, em todo o caso, de uma medida de bem-estar social, capaz de suscitar a ideia de que, se evoluir em ‘bom’ sentido, e se todos as outras coisas permanecerem iguais, a situação melhorará, isto é, as pessoas beneficiarão de uma situação melhor. É por isso que o número de médicos ou de polícias não pode ser um indicador social, mas sim, as estatísticas sobre saúde, ou a taxa de criminalidade.”

Quando se julga existir uma relação entre uma medida estatística e “este vago conceito de bem-estar social”, dá-se a essa medida o nome de “indicador”, no sentido em que mede o *output* ou o rendimento (p.12). Tais indicadores

"deveriam ser elaborados, sempre que possível, no seio dum *sistema geral de compatibilidade social*" (p.12).

Distinguem-se cinco áreas de objectivos respeitantes às relações entre o ensino e a sociedade. Justifica-se uma apresentação um pouco mais pormenorizada destas áreas por duas razões. Primeiro, pelo interesse intrínseco que têm e porque são um bom testemunho do espírito do tempo. Depois, porque esta apresentação ajuda a consciencializar-nos melhor da distância que separa o actual projecto INES da OCDE, das sugestões de 1973 dificilmente operacionáveis.

1. *Transmissão de conhecimentos teóricos e práticos*

Estes conhecimentos devem fazer com que os indivíduos possam agir, mais ou menos independentemente, no seio das principais instituições da sociedade.

Há a distinguir:

- *Alfabetismo funcional*. Os conhecimentos, por si mesmos, não bastam. Para viver de maneira satisfatória numa sociedade complexa, o indivíduo deve aprender a compreender o que pode e o que deve fazer em determinada situação. Os autores do relatório sublinham, de imediato, a dificuldade de recolher sobre este tema um ou vários indicadores que permitam, concretamente, a comparação entre países.
- *Participação na vida política*. Os autores reconhecem que é praticamente impossível determinar a parte que cabe ao ensino, na aptidão para exercer esta actividade.
- *Extensão do poder real e aparente dos alunos*. Trata-se, concretamente, da parte que lhes cabe na tomada de decisões.
- *Eficiência dos consumidores*. Aptidão para escolher bem o que se compra. O número de horas de frequência de cursos de economia doméstica e a inscrição em sociedades de defesa do consumidor, surgem entre os indicadores tidos em consideração.
- *Relações sociais*. Reservar a frequência da escola apenas aos jovens pode privá-los da experiência do trabalho com adultos. Daí a proposta de admitir adultos entre os alunos do ensino secundário. A igualdade entre rapazes e raparigas é outro aspecto tido em conta.
- *Cidadãos de amanhã*. Educação sexual, iniciação ao controlo da natalidade, ecologia das sociedades humanas.

2. *Ensino e economia*

- *Contribuição para o crescimento económico*. Conhecimento das funções de produção e de gestão financeira dos países. O indicador a ter em conta

aqui, é a diferença entre o rendimento por habitante do país p e o de um país tomado como referência, que pode ter explicação nas diferenças entre o capital humano de cada um deles.

- Qualidade da mão-de-obra com instrução.
- Diferença no que se ganha em função do nível de instrução e da profissão.
- Desemprego em função do nível de instrução.
- Peso da formação em exercício.
- Diferenças entre o nível de instrução dos novos diplomados que entram no mercado de trabalho e o nível médio de instrução da população activa.

3. Igualdade de oportunidades

- *Taxas de escolaridade* a todos os níveis e em todos os tipos de ensino, em função do sexo, da raça, do Q.I., da idade, do nível social de origem.
- *Nível de instrução dos pais*.
- Proporção dos *recursos de ensino* consagrados especialmente aos grupos considerados como desfavorecidos.
- *Situação e proventos* em função dos diferentes níveis de sucesso escolar, por idade, raça e nível social dos pais.

4. Sistema de ensino ao serviço das pessoas

Satisfação da procura privada (ou procura social) e vantagens do ensino para o indivíduo: saúde (contributo da educação), tempos livres (proporção de pessoas que continuam a exercer uma actividade de tempos livres aprendida na escola), oferta de educação, criatividade (importância dada ao trabalho pessoal aquando da avaliação do rendimento escolar), disciplina (actos de vandalismo, absentismo), aptidão para a educação (percentagem de adultos que frequentam voluntariamente cursos não relacionados com a formação profissional).

A diversidade dos objectivos a atingir, as grandes diferenças na importância que lhes era atribuída, a dificuldade em traduzir as propostas em especificações operacionais e a prioridade dada ao *output* – deixando os processos de lado – explicam, em grande parte, o fracasso desta primeira tentativa da OCDE.

3.2. Projecto INES

Em 1988, o *Centro de Investigação e Inovação no Ensino* (Centre for Educational Research and Innovation, CERI) da OCDE, lançou um novo projecto cooperativo para pôr a funcionar e fazer levantamentos periódicos de indicadores internacionais do

ensino. Um dos objectivos, em especial, era informar acerca da eficácia da educação e ajudar a definir e a adaptar as políticas educativas nos países membros. O ideal era criar laços entre o conjunto dos indicadores internacionais e os sistemas de indicadores nacionais, realçando os indicadores semelhantes nos dois conjuntos. O quadro teórico deste projecto, sobretudo no que respeita à avaliação, é mais sólido e mais realista do que o utilizado na primeira tentativa.

O facto da OCDE se ter empenhado no projecto INES é significativo do ponto de vista da ligação cada vez mais estreita que existe entre educação e economia. A publicação nos Estados Unidos do célebre relatório *A Nation at Risk* (Uma Nação em Risco), dramatizou esta relação e impressionou profundamente o mundo político e a opinião pública. Verificou-se que eram muito preocupantes os dados fornecidos pela IEA e pela NAEP sobre os conhecimentos dos alunos americanos em matemática e sobre a sua capacidade de leitura. O relatório de MacKnight *et al.* (1987) concluía que uma boa formação em matemática "era essencial à saúde e competitividade do país no mercado tecnológico mundial".

É neste contexto que o governo dos Estados Unidos se empenhou na defesa da ideia de um estudo internacional, sobre a situação da educação. Deviam ser recolhidos indicadores relativos a três aspectos: resultados, recursos e contextos.

Em Novembro de 1987, o Departamento de Educação dos Estados Unidos organizou em Washington em colaboração com a OCDE, uma conferência internacional encarregada de analisar esta proposta. A conferência reuniu vinte e sete países membros da OCDE, bem como peritos do Banco Mundial, da IEA e de diversos organismos americanos.

Foi nesta ocasião que surgiu, de novo, a oposição entre adeptos da avaliação quantitativa e adeptos da avaliação qualitativa. "Contudo, o sentimento geral da maioria é que, quaisquer que sejam as limitações dos indicadores actuais do rendimento escolar, eles fornecem sempre informações que fazem com que os responsáveis pela educação possam dar resposta à procura de dados internacionais que ajudem a situar, comparativamente, os dados nacionais relativos aos rendimentos dos sistemas escolares. Os delegados referiram ainda as frequentes e insistentes exigências dos Chefes de Estado, Ministros e Secretários de Estado sobre orçamentos, resultados escolares e eficácia dos sistemas educativos, tanto nos seus próprios países como no estrangeiro. Estas exigências são postas, com particular insistência, pelos países preocupados com a competitividade económica, como os Estados Unidos, ou que estejam a proceder a importantes modificações estruturais dos seus sistemas educativos, como a França e o Reino Unido" (A. Ruby, 1988:26).

Em Março de 1988, a Conferência de Poitiers completou esta reflexão geral. Foi reconhecida, por grande maioria, a urgência de lançar o projecto sobre

indicadores. Como consequência, o CERI reuniu imediatamente um grupo de peritos encarregados de operacionalizar o projecto, e a partir do mês de Maio as propostas deste grupo foram aceites, com algumas modificações, pelo Comité Director do CERI. O projecto estava pronto para ser lançado. Dedicou-se uma primeira fase exploratória, de dezoito meses, à definição e à experimentação da recolha de dados. Foi uma fase coroada de sucesso e pôde dar-se início à realização propriamente dita.

Constituiram-se quatro redes internacionais de investigação; desempenharam um papel fundamental neste projecto:

- Rede A: Aquisições dos alunos.
- Rede B: Educação e inserção no mercado de trabalho.
- Rede C: Características dos estabelecimentos e dos sistemas de ensino.
- Rede D: Atitudes, planos e expectativas.

Estas redes têm uma grande actividade. Ajudam, em especial, na evolução do sistema de indicadores internacionais, de modo a unir cada vez mais a realidade e a evolução dos sistemas educativos.

A seriedade da reflexão própria deste início dos trabalhos, está bem patente no número impressionante de estudos teóricos e de relatórios redigidos pelo CERI. Um momento decisivo foi a publicação, em 1992, de *Regards sur l'éducation. Les indicateurs de l' OCDE*, apresentação de um primeiro conjunto de indicadores internacionais do ensino. A repercussão deste documento na imprensa e nos organismos oficiais mostra bem o interesse despertado.

Foram escolhidos provisoriamente trinta e seis indicadores, sendo dez deles de natureza claramente experimental (aparecem em itálico na lista que se segue). A ambição é dispor de um *quadro de comandos, de visores*, relativos aos aspectos mais significativos e mais sensíveis da instituição educativa e que forneçam, em comparação com outros sistemas de ensino, uma visão sintética dos pontos fortes e fracos de cada país.

São os seguintes os indicadores que fizeram parte do documento publicado em 1992.

A. *Contexto demográfico, económico e social*

- Nível geral de formação da população adulta (dos 25 aos 64 anos).
- Percentagens de mulheres na população adulta dos 25 aos 64 anos, segundo o nível de formação.
- Taxa de emprego e de desemprego.

- Percentagem da população dos 25 aos 29 anos, relativamente à população total.
- Produto interno bruto (PIB) por habitante.

B. Custos, recursos e processos escolares

- Despesas públicas e privadas com a educação, expressas em percentagem do PIB.
- Despesas públicas com a educação, em relação ao total das despesas públicas.
- *Fontes das despesas públicas com a educação.*
- Despesas de funcionamento.
- Despesas por aluno.
- Investimento nacional a favor do ensino.
- Despesas relativas por aluno e por nível de ensino.
- *Percentagem da população activa a trabalhar no ensino público.*
- Taxa de escolarização no sistema de ensino.
- Taxa de pré-escolarização.
- Taxa de escolarização no ensino secundário superior.
- Taxa de acesso a estudos que confirmaram um diploma de ensino superior.
- Taxa de escolarização no ensino superior não universitário.
- Ratio alunos/professor

C. Resultados do ensino

- Taxa de sucesso no final do ensino secundário do segundo ciclo.
- Taxa de obtenção de um diploma de tipo universitário.
- *Taxa de permanência no ensino superior de tipo universitário.*
- Diplomas por área científica.
- *Percentagem da população jovem que obtém um diploma científico.*
- Desemprego e nível de formação.
- Salários e nível de formação.
- *Resultados a matemática.*
- Comparação inter-escolas dos resultados a matemática.
- Comparação inter-alunos dos resultados a matemática.

Para esta rede de investigação, certos aspectos deverão ser objecto de particular atenção, devido ao seu poder explicativo, nomeadamente: o número de horas de ensino dedicadas a cada disciplina, o grau de individualização do ensino, o número de dias com aulas por ano, o currículo efectivamente seguido, a taxa de

investimento nas tarefas (*time on task*), o trabalho em casa (natureza e duração média).

Por seu lado, a rede de investigação C deseja que, no futuro, o trabalho se concentre mais nos estabelecimentos de ensino do que nos sistemas globalmente considerados. Em nota datada de 1992, J. Scheerens responsável pela rede, adverte que para identificar os indicadores de processo se recorreu, sobretudo, a modelos empíricos da eficiência de um estabelecimento de ensino. Isto quer dizer que, de uma maneira geral, se seleccionaram aqueles indicadores que revelaram uma correlação positiva com os resultados da educação. Scheerens aponta três outras vias possíveis para a selecção de novos indicadores:

1. Elaborar, partindo do que já se publicou até agora, um inventário dos modelos-tipo da "escola ideal". Deste modo, haveria a possibilidade de identificar aspectos que a investigação empírica não consegue detectar, especialmente certos valores como a satisfação que os professores retiram do seu trabalho, assim como a aptidão dos estabelecimentos de ensino para gerarem pressões exteriores, para manterem um bom nível de recrutamento, etc.
2. Analisar os principais programas da política da educação relativamente a um determinado aspecto da problemática educativa (por exemplo, os programas especiais em favor de grupos minoritários), ou a uma determinada categoria de escolas.
3. Analisar o funcionamento dos estabelecimentos de ensino adoptando o ponto de vista do consumidor de educação: pais, alunos,... Poder-se-á, deste modo, identificar as "características de processo" desejadas por certo tipo de escolas.

Eis, a título de exemplo, dois extractos de *Regards sur l'éducation* (1992). O primeiro, refere-se ao rendimento do ensino da matemática de alunos de treze anos, em 1982. Os dados foram cedidos pela IEA. A apresentação, particularmente expressiva, facilita as comparações.

Figura 13
Resultados de matemática

Comparação dos resultados de aritmética Segundo estudo internacional sobre o ensino da matemática, alunos de 13 anos, 1982.

	Média	Erro-tipo	Países Baixos	Canadá (CB)	Bélgica (Fl)	Fráncia	Bélgica (Fr)	Canadá (Ont)	Estados Unidos	Escócia	Inglaterra/Gales	Nova Zelândia	Finlândia	Luxemburgo	Suécia
Japão (3)	60,3	60,3	●	●	●	●	●	●	▲	▲	▲	▲	▲	▲	▲
Países Baixos (2,3)	59,3	59,3	●	●	●	●	●	●	▲	▲	▲	▲	▲	▲	▲
Canadá (3) (Colúmbia Britânica)	58,0	58,0	●	●	●	●	●	●	▲	▲	▲	▲	▲	▲	▲
Bélgica (3) (Comunidade Flamenga)	58,0	58,0	●	●	●	●	●	●	▲	▲	▲	▲	▲	▲	▲
Fráncia (3)	57,7	57,7	●	●	●	●	●	●	▲	▲	▲	▲	▲	▲	▲
Bélgica (3) (Comunidade Francesa)	57,0	57,0	●	●	●	●	●	●	●	▲	▲	▲	▲	▲	▲
Canadá (3) (Ontário)	54,5	54,5	●	●	●	●	●	●	●	▲	▲	▲	▲	▲	▲
Estados Unidos (1)	51,4	51,4	▼	▼	▼	▼	▼	●	●	●	●	●	●	●	▲
Escócia (1)	50,2	50,2	▼	▼	▼	▼	▼	▼	●	●	●	●	●	●	▲
Inglaterra/ País de Gales (1)	48,2	48,2	▼	▼	▼	▼	▼	▼	●	●	●	●	●	●	●
Nova Zelândia (3)	45,6	45,6	▼	▼	▼	▼	▼	▼	●	●	●	●	●	●	●
Finlândia	45,6	45,6	▼	▼	▼	▼	▼	▼	●	●	●	●	●	●	●
Luxemburgo	45,4	45,4	▼	▼	▼	▼	▼	▼	▼	●	●	●	●	●	▲
Suécia	40,6	40,6	▼	▼	▼	▼	▼	▼	▼	●	●	●	●	●	●

(1) Taxa de resposta dos alunos ou das escolas abaixo do nível de 85%

(2) Amostra igual ou inferior a 90% da população em causa

(3) Taxa de resposta dos alunos ou das escolas com que se não pôde contar

 Média inferior à do
país de referência

 Sem diferença estatisticamente significativa do país de referência

 Média superior à do país de referência

Fonte: Dados do segundo estudo internacional sobre o ensino da matemática (SIMS), realizado pela Associação Internacional de Avaliação do Ensino (IEA), entre 1980 e 1982.

O segundo exemplo diz respeito ao grau de autonomia das escolas:

Quadro 12
Decisões tomadas no âmbito da escola, em percentagem do total
das decisões tomadas nos sistemas públicos de ensino,
por nível escolar (1991)

	Nível de Ensino		
	Primário	Secundário 1º ciclo	Secundário 2º ciclo
Espanha	56	56	40
Países Baixos	52	46	59
Portugal	40	39	44
Suécia	37	37	37
Finlândia	35	45	59
Bélgica	35	35	35
Noruega	30	30	23
Itália	29	29	29
Estados Unidos	22	22	22
França	17	34	32
Suiça	10	10	23

DECISÕES TOMADAS PELAS ESCOLAS PRIMÁRIAS E SECUNDÁRIAS

Principais resultados

Na maior parte dos países, o peso das decisões tomadas pela escola é o mesmo no primário e no secundário.

A Finlândia e a França distinguem-se dos outros países: as escolas primárias têm menos "poder".

Definição

Só se tomaram em consideração as decisões tomadas pelas escolas. As comparações referem-se às diferenças entre níveis de ensino público.

Notas explicativas

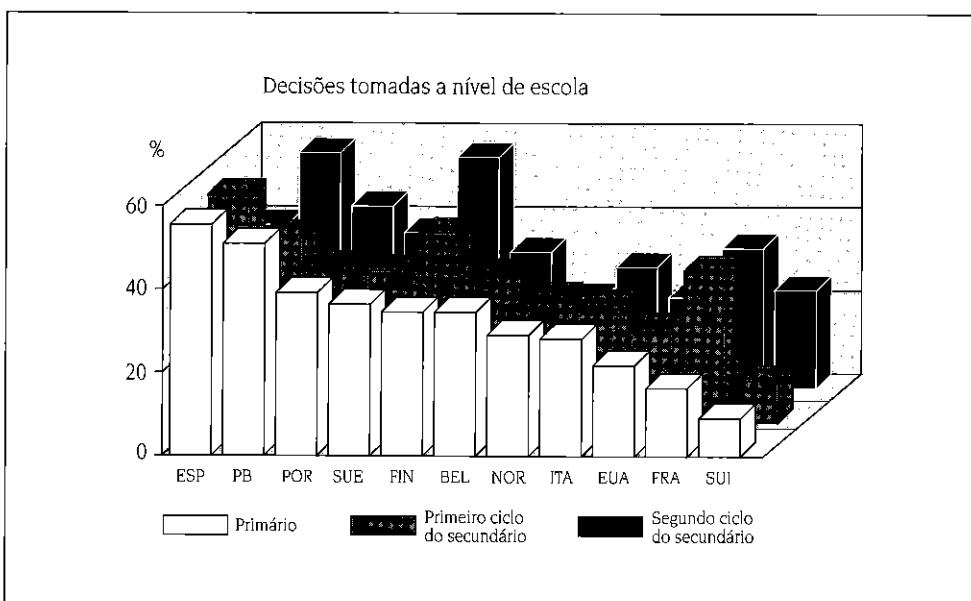
Os países surgem por ordem decrescente segundo o peso das decisões tomadas pela escola primária (CITE 1).

Na maior parte dos países, a legislação é a mesma para o conjunto da escolaridade obrigatória (níveis CITE 1 e 2).

Em quatro países (Bélgica, Estados Unidos, Itália e Suécia), verifica-se uma situação idêntica nos três níveis de ensino.

Na Finlândia, pode observar-se uma evolução no poder de decisão da escola: aumenta com os níveis: do primário para o secundário inferior e deste para o secundário superior. Estamos em crer que é esta a situação que se verifica com mais frequência.

Figura 14
Decisões tomadas pelas escolas primárias e secundárias



No futuro o número de indicadores aumentará concretamente e a natureza de alguns deles modificar-se-á. Pensa-se, concretamente, que é conveniente insistir mais nos rendimentos escolares, aquilo que os alunos aprendem e o seu desenvolvimento pessoal e social é, de facto, o que mais prende a atenção do mundo da educação.

Para países como a Bélgica, onde não existem exames nacionais organizados a nível central, não é fácil obter indicadores de rendimento válidos. Apenas as grandes investigações normativas da IEA ou da IAEP fornecem alguns pontos de referência, mas não dão resposta completa às necessidades anuais da OCDE. A questão torna-se ainda mais difícil, dado que, nas estruturas federais, os currículos de ensino tendem para a diversificação.

Casos como o da Austrália, que congrega os seus diversos Estados no seio da Commonwealth continuando a publicar estatísticas nacionais, pode ser fonte de inspiração para se encontrar uma solução. O mesmo acontece nos Estados Unidos e no Canadá. Quanto à Inglaterra, onde a independência das autoridades locais em matéria de ensino era enorme até há pouco tempo, acaba de adoptar, como vimos, um currículo nacional e os correspondentes mecanismos de avaliação.

3.2.1. Problemas particulares relativos às atitudes, concepções e expectativas

Na primeira edição de *Regards sur l'éducation*, não aparecem indicadores relativos às atitudes e expectativas (Rede D). Esta ausência explica-se quer pela complexidade da avaliação destas características quer pela insuficiência de dados existentes nos países. Ora esta é uma questão de grande importância que foi objecto de atenção por parte da OCDE logo desde o seu primeiro projecto sobre indicadores do ensino (ver *supra*): muitas das avaliações pretendidas já abrangiam as expectativas (Carr-Hill e Magnussen, 1973). É por isso que, de acordo com J. De Groot *et al.* (1992) em particular, damos a este tema um desenvolvimento especial.

Os indicadores relativos às atitudes e às expectativas, referem-se ao modo como os actores directos do sistema de ensino o apreendem, o vivem, dele esperam benefícios e se situam em relação a ele. Em particular as atitudes influenciam directamente o rendimento escolar e, juntamente com as expectativas, determinam em larga medida, o contexto em que se desenvolve o processo educativo. O fenómeno anda associado a uma enorme carga cultural, o que explica o carácter delicado da escolha de indicadores internacionais neste campo.

Os trabalhos da Rede D foram, até agora, apenas de natureza exploratória. Teve em vista dois objectivos principais:

- Definir atitudes e expectativas das comunidades educativas em matéria escolar que fossem comparáveis a nível internacional. Devem ser escolhidas cinco atitudes e cinco expectativas que exerçam uma influência duradoura sobre aspectos permanentes do sistema de ensino. Os dados serão avaliados segundo a idade, o sexo e a classe social.
- Definir cinco atitudes e cinco expectativas dos professores e dos alunos relacionadas com o bom funcionamento da escola.

Actualmente quase não se dispõe de resultados de estudos fiáveis e regulares sobre este tema na maior parte dos países participantes. Deste modo, a fim de poder dispor de um quadro de referência capaz de ajudar à escolha de indicadores, a Rede D elaborou, partindo dos estudos já publicados, um inventário das preocupações comuns, repartindo-as em sete grupos:

1. Professores
 - Qualidade
 - Estatuto e moral
 - Nível de profissionalismo

- Formação
- Atitude perante inovações e reformas

2. Gestão da escola

- Objectivos a atingir pela escola
- Clima, estado de espírito, cultura da escola
- Modelo de gestão
- Relação com pais e empregadores.

3. Currículo

- Qualidade do currículo
- Currículo previsto
- Currículo efectivamente praticado (currículo implantado)
- Níveis de rendimento dos alunos
- Relação entre currículo, vida e mundo do trabalho

4. Relações professores/alunos

- Expectativas dos professores em relação aos alunos
- Expectativas dos alunos em relação aos professores
- Estado de espírito dos professores
- Estado de espírito dos alunos
- Avaliação e auto-avaliação dos professores
- Avaliação e auto-avaliação dos alunos

5. Comunicação

- Circulação da informação sobre objectivos e prioridades
- Comunicação com outros grupos profissionais
- Modalidades de consulta dos pais
- Comunicação dos resultados escolares aos pais.

6. Equidade

- Discriminação racial, social, sexual
- Igualdade de oportunidades
- Imparcialidade dos professores
- Medidas tomadas a favor dos alunos portadores de deficiências ou com necessidades especiais
- Respeito pela língua materna

7. Política de ensino

- Objectivos a atingir no sistema educativo
- Prioridades orçamentais
- Escolas públicas – Escolas privadas
- Liberdade de escolha de ensino

3.2.2. O Futuro

É um dado adquirido que, a partir de agora, os indicadores da educação da OCDE terão o seu lugar ao lado dos indicadores da economia. J. Magy (1992:10) pensa, com razão, que “lhes cabe desempenhar a função de pano de fundo permanente das análises fundamentais das políticas nacionais de educação”.

Há que sublinhar, por outro lado, o efeito estimulante deste projecto no desenvolvimento dos serviços estatísticos e nos sistemas de pilotagem, nos países que ainda não se tinham dotado, de uma maneira satisfatória, destes instrumentos de trabalho. Notou-se também este efeito, a nível das grandes organizações internacionais (OCDE, UNESCO, UE), que se esforçaram por coordenar as suas estatísticas. A *Classificação Internacional Estandardizada do Ensino (International Standard Classification of Education - ISCED)*, inicialmente publicada pela UNESCO, vai ser revista, após ampla consulta aos utentes, a fim de evitar várias das ambiguidades actuais.

A Conclusão nº 3 da Assembleia Geral dos países participantes no projecto INES, que teve lugar em Lugano em 1992, é a seguinte: “Como os indicadores (internacionais) não foram muito elaborados, devem ser completados e ponderados com dados suplementares. Para analisar as informações que eles fornecem, e retirar daí conclusões, é sobretudo essencial ter em conta os contextos nacionais isto é, ter em conta muito em especial, as estruturas governamentais e os contextos nos quais se elaboram as decisões em matéria de ensino (importância da descentralização, da autonomia das escolas nas decisões fundamentais, etc.), as diferenças de condições económicas e de recursos dedicados ao ensino (patentes no peso que o ensino tem no PIB), as diferenças de estruturas da sociedade (grau de homogeneidade, percentagem de grupos minoritários...), as diferenças nas principais estatísticas sócio-económicas (demografia, tendências da população...), assim como certas componentes relacionadas, entre outras coisas, com a cultura, os valores, a história, que contribuem para construir um quadro especial, no qual se efectuam as tomadas de decisão e a concretização das políticas de ensino. (...) Daí que, na medida do possível, devam ser utilizadas outras fontes para confirmar, corrigir ou compreender melhor as informações fornecidas pelos indicadores.” Não se podia dizer melhor. Esta necessidade de esclarecer o significado dos indicadores repete-se, sob outro aspecto, na Conclusão nº 17: “É certo que os indicadores do ensino podem beneficiar muito se forem confrontados com os indicadores dos sectores das ciências, da tecnologia, da economia”.

A importância do projecto INES que adquire actualmente toda a sua pujança, é considerável não só pelo seu interesse intrínseco mas também pelo efeito de

treino que cria entre os países membros. Assim, desejosa de dar a conhecer os seus próprios dados (já não misturados com os das outras comunidades belgas), a comunidade flamenga da Bélgica publicou, em Outubro de 1993, os seus próprios *Regards sur l'éducation*, com o título sugestivo de *L'école fait rapport. L'enseignement flamand dans un contexte international*. A comunidade francesa da Bélgica vai fazer o mesmo.

Porém, o projecto INES está cheio de obstáculos. Tenha-se por exemplo em conta certas dificuldades relacionadas com a recolha de dados sobre resultados escolares e aspectos contextuais. Pensou-se, a certa altura, que a IEA ou a IAEP, ou estes dois organismos em conjunto, poderiam encarregar-se deste trabalho. Em seguida, a OCDE resolveu deixar a cada país o cuidado de recolher os respectivos dados. Ao terminar a redacção deste livro as hesitações subsistem ainda.

CAPÍTULO IV- MICROPILOTAGEM

Generalidades

As considerações anteriores dizem respeito à pilotagem em grande escala: a nível nacional e regional. Porém, as avaliações em pequena escala têm também muito interesse. Assim, uma autarquia ou uma escola que tenham elaborado um projecto educativo com objectivos particulares podem também tentar verificar em que medida as suas intenções se realizam. A observação de rendimentos diferentes, cuja explicação se deve ao contexto local, tem muitas vezes importantes implicações pedagógicas e políticas.

Há a distinguir a *auditoria da escola*, muitas vezes levada a cabo por inspectores, a *auditoria participativa da escola*, em que aqueles que aí trabalham cooperam na avaliação com os observadores externos, e a *auto-avaliação*.

Só haverá autoregulação fecunda se se resolverem estes dois problemas: o da amplitude de manobra que o poder organizador concede ao estabelecimento de ensino e a qualidade das relações que este estabelece com o meio envolvente.

A propósito do primeiro problema, a conclusão que T. Wyatt e A. Ruby (1990) tiraram da sua rica experiência australiana não apresenta ambiguidades: "Para assumir as suas responsabilidades em relação ao sistema escolar global, o poder organizado deve estar preparado para usar se necessário a sua autoridade, para que sejam cumpridas as decisões tomadas sobre os objectivos principais a atingir e para assegurar o bom funcionamento do sistema. Mas, ao mesmo tempo, é essencial que as escolas gozem de uma liberdade que lhes permita agir em função das necessidades particulares dos seus alunos. O exercício desta liberdade de acção, respeitadora das opções essenciais tomadas pelo poder, nem sempre é fácil. E não é possível com autoridades autocráticas, hipercentralizadoras. É pois necessário desenvolver um sistema complexo de relações, baseadas numa compreensão comum dos objectivos e das responsabilidades, o que só é possível se existir um mecanismo eficaz de consulta e de comunicação".

Quanto ao segundo problema, uma escola não tem melhor maneira de cumprir a sua missão do que cultivar a interacção com todas as forças vivas e activas do meio envolvente. Trata-se, evidentemente, mais de uma questão de espírito, de informação mútua sobre o essencial, do que da gestão técnica do ensino e da avaliação que, obviamente, continuam a ser uma competência directa dos professores.

Para a avaliação dos níveis de rendimento, os serviços nacionais ou comunitários podem pôr os instrumentos de macropilotagem construídos por

especialistas, à disposição de quem quiser servir-se deles, a nível local e, eventualmente, ajudar no tratamento dos dados. Acontece cada vez mais. Contudo, estes instrumentos não poderão dar resposta a todas as questões colocadas por determinada escola, em particular, a propósito das principais facetas do seu projecto, da sua acção e das interacções com o meio em que está inserida. A construção de instrumentos de medida, deve ser um dos objectos da formação (inicial e contínua) de todos os professores.

N. Bottani e H. Walberg (1992:73) elaboraram um inventário das conclusões dos principais autores (Scheerens, 1989; Windham, 1988; Beneviste, 1987; Teuber, 1987; Oakes, 1987) que propuseram listas de variáveis, cuja importância foi demonstrada pelas investigações experimentais sobre eficiência escolar.

Eis a síntese desse trabalho:

- Diferentes maneiras de estimular o rendimento escolar.
- Política educativa que estimule o rendimento escolar.
- Elevado nível de expectativa (percentagem de alunos que os professores esperam que venham a ter sucesso; esperança dos alunos em prosseguir os seus estudos).
- Qualidade da gestão pedagógica (tempo que o responsável pela direcção da escola dedica à animação pedagógica, tempo dedicado aos problemas de educação nas reuniões de professores,...).
- Clareza dos objectivos prioritários que a escola se impôs.
- Cooperação dos professores na planificação do ensino e noutras actividades.
- Qualidade do currículo.
- Qualidade da avaliação.
- Qualidade dos professores (formação disciplinar e pedagógica inicial e contínua, flexibilidade, antiguidade, etc.).
- Qualidade estrutural do ensino (definição clara dos objectivos, divisão das aprendizagens em unidades coerentes, verificação regular das aquisições, *feedback* frequente, individualização ocasional,...).
- Oportunidade de aprender.
- Pilotagem do progresso das aprendizagens (utilização de testes de conhecimentos; processos informatizados dos resultados escolares individuais, etc.).
- Organização pedagógica (trabalho de grupo, etc.).
- Taxa de enquadramento (ratio professor/alunos).
- Utilização do tempo por parte dos professores (tempo de ensino – tempo de não-ensino).

- Distribuição diária das actividades escolares.
- Tempo efectivamente investido no trabalho (*time on task*).
- Tempo de trabalho escolar fora da escola (deveres de casa).
- Duração do ano escolar.
- Clima que se vive na escola e modo de exercer a disciplina.
- Formas de indisciplina: gazeta, absentismo, vandalismo, perturbações no trabalho.
- Influência dos colegas

1. REALIZAÇÕES

1.1. Inglaterra: um exemplo de *checklists* de auto-avaliação

D.L. Nuttal (1993) descreve as acções de auto-avaliação (*School Self-Evaluation, SSE*) desenvolvidas no decurso destes últimos anos no Reino Unido. Verifica-se que as *checklists* são os instrumentos mais acessíveis às escolas.

O primeiro grande passo foi dado com a publicação em 1976 de *Keeping the School under Review* (Manter a escola sob análise crítica permanente) (G. Rogers, 1977). Este instrumento de trabalho, elaborado pelos inspectores da *Inner London Education Authority*, contém centenas de perguntas que ajudam as escolas a questionar-se sobre a sua organização, os seus recursos, os níveis de rendimento escolar e as relações mantidas com o exterior. É evidente que qualquer escola pode construir um instrumento deste tipo. Em vez de pretender avaliar tudo de uma só vez, é possível também tratar sucessivos aspectos particulares ciclicamente.

Eis os temas que constam da lista de *Keeping the School under Review* para o ensino secundário (Nuttall, 1993:43):

- Alunos. Órgão directivo da escola, pais, comunidade, pessoal da escola.
- Currículo. modo de elaboração. Currículo dos diferentes anos. Documentos de ensino. Opções metodológicas fundamentais. Avaliação. Igualdade de oportunidades rapazes/raparigas. Normas favoráveis aos alunos com necessidades especiais.
- Auto-avaliação e formação contínua dos professores.
- Processos de tomada de decisão e comunicação no interior da escola.
- Meio envolvente.
- Recursos.
- Estatísticas.

Todos os professores são convidados a interrogar-se sobre o modo de trabalhar na aula e sobre o seu conhecimento dos alunos. Assim como os

responsáveis pela direcção da escola se devem questionar sobre o modo como utilizam o tempo, os objectivos que pretendem atingir e sobre a organização da escola.

- Questões-teste: (a) Recomendaria eu a um colega que procurasse emprego na minha escola? (b) Recomendaria eu aos meus amigos que matriculassem o filho na minha escola?

São sugeridas questões a propósito de cada uma das rubricas apresentadas anteriormente. Eis, por exemplo, as que se referem à elaboração do currículo:

- Como são definidos os objectivos da escola?
- Como é definido o programa? Está coerente com os objectivos?
- Como se situa o programa da escola, em relação ao programa oficial e às directivas da inspecção?
- O processo de consulta engloba todos os que trabalham na escola, bem como os pais, os responsáveis pela direcção da escola, os serviços de inspecção?
- Por que importantes alterações devia passar o estabelecimento de ensino no decurso dos próximos anos e qual a estratégia a adoptar para isso?
- Qual a política da escola sobre a linguagem enquanto comunicação? Que disposições tomar para melhorar a capacidade de expressão dos alunos?
- No decurso dos últimos anos, discutiram-se muito na escola publicações importantes relacionadas com o currículo?

Nuttall observa que, desde o final dos anos setenta, foram publicadas quarenta e cinco versões deste tipo de listas pelas autoridades locais da Inglaterra. Nota ainda que poucas delas fazem sugestões sobre o modo de responder às questões, sobre o tratamento das respostas e a sua utilização. O modo de trabalhar e o proveito retirado desta auto-avaliação também variaram muito de escola para escola.

O Projecto GRIDS (grelhas) revelou-se mais fecundo. Tratou-se de um processo em cinco etapas que, rapidamente, se difundiu pelo país.

1. Fase preparatória: plano de trabalho, estimativa das despesas, escolha de um eventual coordenador. O processo de auto-avaliação é explicado ao pessoal e é feito um esforço para conseguir a sua cooperação.

2. Avaliação inicial. Um questionário de resposta anónima aborda as potencialidades e os pontos fracos da escola, na perspectiva do pessoal. Quem trabalha na escola é convidado a sugerir aspectos que mereçam especial avaliação. Os aspectos a avaliar resultam de uma discussão geral das respostas.
3. Avaliação propriamente dita. Há um esforço para a concentrar em poucos aspectos. A equipa de trabalho procura identificar os pontos fortes e os pontos fracos das práticas pedagógicas e sugere alterações. Só são tomadas em consideração as propostas aceites por unanimidade pelas pessoas que trabalham na escola.
4. Acção. São definidas as necessidades de formação contínua das pessoas que trabalham na escola e são tomadas medidas para as satisfazer, geralmente em colaboração com o poder organizador. É uma fase cuidadosamente acompanhada: verifica-se se as alterações ocorridas conduzem às desejadas melhorias e se não surgem dificuldades inesperadas.
5. Avaliação e adopção. O novo estado de coisas é comparado ao anterior, de modo tão objectivo quanto possível.

É importante referir que a participação das escolas no projecto é voluntária e que os professores se sentem responsáveis pelo trabalho. Não é imposta a redacção de um relatório de avaliação, nem é exigida a comunicação às autoridades de qualquer documento, eventualmente redigido. Pretende-se evitar assim que alguém se sinta alvo de acusação.

Nuttall refere, finalmente, que na experiência GRIDS se verificou que as escolas primárias se interessavam, sobretudo, por problemas pedagógicos (por exemplo, o ensino da leitura), enquanto que as escolas secundárias se voltavam mais para problemas de organização e de gestão. Foram apresentadas sugestões no sentido de se conseguir uma abordagem mais equilibrada.

(Foram publicadas numa colecção "GRIDS", por Longman Resources Unit, York, 1988, grelhas para o ensino primário e secundário, assim como manuais metodológicos sobre o tema).

1.2. França: Proposta de Peretti

Mais ou menos na altura em que o projecto GRIDS se desenvolvia, surge uma proposta de A. de Peretti (1986, comunicação pessoal) – desconhecedor na ocasião dos trabalhos britânicos – que vem no mesmo sentido: proceder a uma avaliação *positiva* da *qualidade* do ensino em cada escola. É proposto um conjunto

de cento e vinte e cinco indicadores. Deveriam "possibilitar aos órgãos directivos e às pessoas que trabalham nas escolas a escolha, em conjunto com as famílias e os alunos, dos indicadores mais adaptados às características da sua escola e aos esforços de aperfeiçoamento mais oportunos" (p.357). Esta escolha deve ser móvel, para evitar fenómenos de saturação e desgaste.

Peretti sugere que pessoas com expediente acompanhem as escolas na sua avaliação. Poderiam ser enviadas sínteses das observações efectuadas para os órgãos regionais ou nacionais.

Distinguem-se cinco categorias de indicadores e para cada uma delas são mencionados cinco aspectos:

A. *Indicadores de organização*

- Qualidade da definição dos objectivos (objectivos gerais, específicos; projecto global; características da escola; conhecimento dos objectivos por parte dos diversos parceiros).
- Organização diferenciada dos grupos de alunos (princípios que presidem à distribuição dos alunos por turmas, regras de formação dos grupos de nível, critérios de organização dos grupos de métodos, possibilidade de formação de grupos interdisciplinares).
- Flexibilidade na utilização do tempo (variações segundo os objectivos e as disciplinas, segundo a pedagogia seguida).
- Medidas destinadas a quebrar a monotonia ou a evitar perdas de tempo, facilidade na substituição dos professores, consideração das pretensões metodológicas dos professores).
- Eficácia da concertação entre professores (medidas destinadas a facilitar os encontros, funcionamento dos órgãos de gestão, coerência dos conselhos de turma,...).
- Ordenamento dos meios (agendamento e distribuição dos locais, centros de documentação, laboratórios e oficinas, locais postos à disposição dos alunos, adequação dos meios, dos horários, das reuniões de escola).

B. *Indicadores de funcionamento*

- Modo como decorre o ensino (coerência metodológica, peso das inovações metodológicas, grau de utilização do centro de documentação, grau de diferenciação da pedagogia, opções disciplinares oferecidas aos alunos).
- Engenharia metodológica e tecnologia educativa (importância e qualidade dos materiais, instrumentos de observação e de avaliação utilizados, audiovisuais, EAO⁽¹⁾, qualidade do boletim informativo enviado às famílias);
- Avaliação das aprendizagens (frequência da avaliação formativa, coerência das avaliações sumativas, unidades capitalizáveis, auto-avaliação).

(1) Projecto francês de instalação de computadores nos estabelecimentos de ensino secundário (N.do T.)

- Práticas de orientação e de permanência no mesmo ano.
- Medidas de formação (formação contínua dos professores dentro e fora da escola, professores a estagiar em empresas,...).

C. Indicadores de actividades

- Modo de atender e receber as pessoas (reuniões de pais, acolhimento das pessoas que visitam a escola,...).
- Auxílio aos alunos (acompanhamento nos estudos, tutoria, apoio a alunos com dificuldades,...).
- Iniciação cultural.
- Produções e publicações.
- Competições (relacionadas com os programas escolares, artísticas, desportivas,...).

D. Indicadores de relação com:

- O mundo escolar
 - O mundo escolar e universitário (relações com outras escolas, etc.).
 - Natureza das relações de inspecção.
 - Percentagem de conselheiros pedagógicos entre os professores.
- O meio familiar (frequência das mensagens dirigidas aos pais, abertura de certos serviços às famílias, etc.)
- O meio envolvente (geminiação com empresas, estudos no terreno,...)
- A opinião pública (artigos na imprensa local, participações em emissões de rádio ou de televisão, dias de abertura ao meio,...)
- O mundo intelectual (bibliotecas, escolha de manuais, frequência de museus e de laboratórios de investigação; conferências de alto nível na escola,...).

E. Indicadores de coerência

- Resultados (acompanhamento do desenvolvimento dos alunos, domínio da expressão oral e escrita, trabalho independente, sucesso nos exames,...).
- Localização dos potenciais progressos.
- Regularidade (imagem da escola, taxa de substituição das pessoas no seu emprego, percentagem de professores em formação,...).
- Projecto de escola.
- Ligação à escola.

Desde 1986 que este quadro de indicadores vem sendo utilizado em jornadas de formação para responsáveis, pela direcção das escolas e em cursos universitários de verão. Cada um dos participantes é convidado a escolher da lista trinta indicadores que possam corresponder ao seu próprio projecto pedagógico. Esta escolha serve de base para a discussão entre os participantes. Pede-se-lhes que

expliquem as razões da sua escolha, tendo assim oportunidade de descobrir por que razão certos colegas, embora trabalhando em situações muito próximas das suas, escolhem indicadores diferentes.

1.3 Dinamarca

Para além de programas semelhantes aos que acabam de ser referidos, as escolas dinamarquesas podem também requerer uma peritagem externa. Neste caso, seis meses após a auto-avaliação, um inspector passa uma semana inteira na escola; tem encontros com alunos, professores, pais e representantes da comunidade local. É publicado em seguida um relatório provisório: só podem ser modificadas as informações factuais erradas e nunca os pareceres emitidos.

1.4. Nova Zelândia

Na Nova Zelândia a autopilotagem das escolas tende a generalizar-se (paralelamente à instauração de uma pilotagem nacional).

As escolas são convidadas a fazer um balanço do seu rendimento de três em três anos, antes de receberem a visita do delegado do *Education Review Office*, que vem verificar se os recursos postos à disposição da escola foram correctamente utilizados e se os objectivos que constam do projecto de escola foram atingidos.

O processo de auto-avaliação varia conforme a idade dos alunos e a dimensão da escola; em todo o caso deve começar sempre pela elaboração de um plano de avaliação, em que se definam os indicadores de qualidade da educação. A questão que se põe sempre é esta: que efeitos teve a política educativa adoptada pela escola nas aprendizagens dos alunos? Os indicadores são escolhidos pelas autoridades locais em matéria de ensino, em colaboração com os professores e outros agentes educativos.

De acordo com a sua natureza, os indicadores podem ser recolhidos em diversos momentos e de diferentes formas. Podem, inclusive, ir buscar-se aos indicadores usados a nível nacional.

Como todos os que participaram na formulação do projecto estão associados à avaliação, geralmente é fácil encontrar pessoas capazes de recolher as informações e os dados necessários. Os professores ficam, em parte, libertos das tarefas pedagógicas para colaborarem neste trabalho. Tal como as outras pessoas que intervêm no projecto, também eles devem estar preparados para a sua missão de avaliação.

Finalmente é elaborado um relatório. Nele se descreve o método utilizado, os indicadores de qualidade escolhidos e os resultados da avaliação. Deve terminar com a proposta de medidas a tomar, precisando quais os novos recursos eventualmente, necessários para ultrapassar os obstáculos identificados e permitir alterações de orientação nos trabalhos.

Numa primeira fase este relatório só é dado a conhecer aos implicados. Em seguida, é enviado ao *Education Review Office* que analisa o modo como as informações foram recolhidas, as propostas formuladas e faz eventualmente outras sugestões.

Os processos acabados de referir, a título de exemplo, têm tendência a privilegiar o aspecto descriptivo. Podem prestar grandes serviços à educação. Comentando a experiência inglesa na matéria, Gipps (1989) lastima, contudo, que os relatórios, muitas vezes abundantes, que as escolas redigem sobre o seu próprio funcionamento, raramente sejam lidos pelos pais e muitas vezes sejam negligenciados pelos políticos.

2. RECURSO A PERITOS EXTERNOS

A autopilotagem não exclui o recurso a peritagens externas quer seja de investigadores e inspectores, de pais possuidores de particulares competências ou de responsáveis pela administração. Esta ajuda externa, que pode verificar-se em qualquer dos estádios do processo, é recomendada no projecto GRIDS, e Ruby e Wyatt (1988) referem um processo semelhante, experimentado pelo Departamento de Educação do Estado de Queensland (Austrália): qualquer escola que sinta necessidade de informações que possibilitem a resolução de um problema, pode entrar em contacto com os especialistas do Serviço de Investigação do Estado. Os investigadores combinam com a escola as informações a recolher bem como o modo de as analisar e apresentar. Os investigadores não se substituem à escola na resolução dos problemas; limitam-se a uma função de apoio. A tomada de decisões e a condução dos trabalhos continuam a ser da responsabilidade da escola.

3. CONCLUSÃO

A micropilotagem deriva de uma atitude geral, de uma filosofia da acção: não se contentar com boas intenções mas obrigar-se a verificar sempre, tão

objectiva e sistematicamente quanto possível, se essas intenções foram realmente levadas à prática e em que medida. Há que desenvolver, de facto, uma nova “cultura de escola”.

Uma das condições para o sucesso da autopilotagem está na preparação dos professores para esta nova missão. O conceito de indicadores é relativamente novo nas escolas. A sua utilização exige, não só a aquisição de conhecimentos e capacidades novas, mas também uma transformação global do estado de espírito em que as escolas têm funcionado. Os professores nem sempre foram habituados a controlar, sistematicamente, a eficácia do seu ensino.

É certo que, de uma maneira geral, os professores reconhecem a importância da recolha sistemática de indicadores, de que acabámos de falar, mas na prática, este aspecto é muitas vezes considerado mais como uma sobrecarga, do que como uma componente essencial e integral do processo de ensino.

As hesitações em avaliar, rigorosamente, os programas e as práticas pedagógicas devem-se a vários factores. O peso de um cansativo dia de aulas deixa pouca oportunidade para planificar e efectuar o que é exigido e para cada um se interrogar sobre as observações realizadas. Além disso, não se deve subestimar a influência do estado de espírito do pós-Maio de 68, maioritariamente adverso a toda a medida quantitativa do rendimento escolar.

Finalmente, no ensino, as inovações sucedem-se e as pessoas esperam, muitas vezes, efeitos quase imediatos. Ora, a pilotagem é uma acção complexa, cujos efeitos se manifestam de forma mais ou menos lenta; exige atenção e esforços prolongados.

Não se deve também ignorar o medo de ser considerado como responsável. Se os indicadores revelarem situações pouco favoráveis, não serão os professores os primeiros a ser apontados como responsáveis? Será, evidentemente, um erro grave. Um dos meios mais seguros e legítimos para o evitar, consiste em prestar mais atenção à explicação dos resultados observados, do que aos resultados em si mesmos. A experiência mostra que a grande maioria dos professores se sente bem quando tem ocasião de exercer, ainda melhor, a sua profissão.

Resumindo, em relação à macropilotagem, a micropilotagem permite concentrar muito mais a atenção sobre aspectos qualitativos e contextuais. Não deriva da auditoria, mas pretende ser, essencialmente, uma análise e reflexão serenas sobre o que se passa e passou na escola, sem deixar de ser rigorosa.

Realizada num espírito democrático, a micropilotagem insiste muito menos nas carências do que nos modos de fazer cada vez melhor.

CAPÍTULO V – OS PRÓS E OS CONTRAS

1. BENEFÍCIOS AUFERIDOS

São de seis categorias:

- Melhor racionalidade política.
- Acção educativa geralmente mais eficaz.
- Maior equidade. Para a garantir é necessário começar por detectar as desigualdades e, depois, combatê-las, tendo sempre cuidado em respeitar as diferenças.
- Informação democrática de todos os membros da comunidade educativa.
- Desenvolvimento da investigação avaliativa em educação.
- Informações preciosas para a prática do ensino.

Analizando os efeitos secundários benéficos da pilotagem praticada na Inglaterra pela *Assessment of Performance Unit*, Gipps (1989) é de opinião que o período (muitas vezes de vários anos) que decorre entre duas avaliações pode ser utilizado para aprofundar os resultados. Em especial a análise dos erros pode fornecer preciosas informações aos professores. E Gipps conclui: "Não deixa de haver ironia no facto de a avaliação nacional, acolhida com imensas reservas pelos professores, acabar por se revelar uma preciosa fonte de *feedback* directo para os responsáveis pela prática pedagógica quotidiana."

1.1. Ruptura do isolamento

Num estudo em que situa a problemática dos indicadores e da pilotagem no contexto da política de educação da comunidade francesa da Bélgica, J. Magy (1992: 152-153) dá importância a um aspecto habitualmente esquecido. Refere, com razão, que em muitos aspectos os pais, os alunos, os educadores e os quadros pedagógicos são inteiramente abandonados a si mesmos, quando têm de tomar decisões mais importantes. Estão, de facto, muito longe de dispor sempre de informações suficientes que os ajudem a tomar uma opção esclarecida.

Esquematicamente, J. Magy refere os casos seguintes:

- Pais e alunos que têm de escolher a frequência de uma escola, sem estarem informados das reais qualidades dos diversos estabelecimentos de ensino.

- Responsáveis pelas escolas de forma alguma informados acerca do rendimento e outras características pedagógicas e de gestão de outros estabelecimentos de ensino.
- Poder organizador e inspecção com falta de dados de referência sobre conjuntos de escolas, ou sobre o que se passa, efectivamente, noutras sistemas educativos.
- Professores e formadores sem instrumentos de referência para definirem os objectivos e os níveis de mestria e, muitas vezes, mal informados sobre os instrumentos de avaliação existentes e sobre o modo de os utilizar adequadamente. Ora, tais instrumentos podem contribuir para o progresso da "profissionalidade" dos professores e, por esta via, da própria educação.

É evidente que se as diversas formas de pilotagem fossem aplicadas com base, por exemplo, em conjuntos de indicadores significativos, recolhidos com rigor científico, com a colaboração de todos os membros da comunidade educativa, em particular dos professores e dos quadros pedagógicos e administrativos, estaríamos em condições de esperar, não só um aumento directo da produtividade (no sentido nobre do termo) do sistema educativo, mas também, indirectamente, um importante aumento da qualificação dos actores implicados no processo.

A experiência adquirida por ocasião dos grande projectos avaliativos, realizados no decurso dos últimos decénios, mostra que a reflexão fundamental que exigem, e o domínio progressivo dos instrumentos e das técnicas que suscitam, trazem consigo benefícios muitas vezes superiores ao proveito imediato que se obtém com as avaliações.

2. PERIGOS

Embora se justifique plenamente que se verifique em que medida estão a ser atingidos os objectivos atribuídos à educação (e esta evidência é ainda maior, se a comunidade educativa, a começar pela classe docente, tiver oportunidade de participar nesta tarefa), não é menos certo que esta verificação é sentida por muitos professores como uma ameaça. Não irão ser tidos como os únicos responsáveis das falhas observadas, não se arriscarão a ser objecto de sanções (como acontece nos Estados Unidos)? São questões de que nos devemos acautelar seriamente. Não se instauram processos de reparação de danos contra outros membros de profissões intelectuais (inclusive contra médicos) que, de algum modo, prejudicaram os seus clientes? Por que razão gozariam os professores de imunidade? Nada justifica esta situação de excepção desde que

uma eventual falta profissional grave possa ser provada de forma irrefutável. Será um caso raro, sabendo-se como estão repartidas as responsabilidades do insucesso escolar e, *a fortiori*, de uma simples falta de rendimento.

Parece de longe preferível adoptar uma atitude positiva, optando sempre por uma avaliação formativa: as informações fornecidas pelos indicadores deveriam servir, essencialmente, para encaminhar no sentido de uma perfeição a todos os níveis: política educativa, modificação dos programas, aperfeiçoamento da didáctica, apoio aos alunos com dificuldades, incitamento à cooperação dos pais,...

Dito isto, continua válida a ideia inicial: sim, é preciso verificar se os objectivos escolhidos são de facto atingidos.

Há, contudo, que ter em conta algumas visões deformadas da questão, de que Madaus (1988) nos apresenta um inventário:

- Se alunos, professores e quadros pedagógicos acreditarem, embora erradamente, que os aspectos abordados pelos indicadores são importantes, dar-lhes-ão importância. Em especial, se os professores pensarem que dos resultados dos testes de rendimento dependem graves decisões, insistirão nas aulas nas matérias que sairão nessas provas.
- À medida que se conhecem as matérias abrangidas pelos testes, e o modo como são avaliadas as aquisições sobre essas matérias, também os programas ensinados nas aulas se modificam. Tal modificação não é, necessariamente, reprovável: alguns responsáveis pelo ensino viram, neste fenómeno, um meio de levar os professores a concentrarem os seus esforços nos objectivos realmente importantes.
- Os professores prestam particular atenção ao modo como os seus alunos são interrogados (respostas abertas, questionários de respostas múltiplas, testes mais centrados no saber do que no saber-fazer,...) e adaptam a sua forma de ensinar de modo a fazer face a estas exigências, mesmo discorrendo delas.
- Corre-se o risco de considerar os indicadores como informações infalíveis sobre o que a escola produz.

São mais fundamentais ainda as críticas dirigidas à avaliação feita através de testes de conhecimentos. Estes testes tratam, a maior parte das vezes, de capacidades intelectuais básicas e de disciplinas privilegiadas: ciências, geografia, etc..

Estas aprendizagens são certeza importante mas correspondem, apenas, a alguns dos objectivos da *educação*. Não se devem nunca perder de vista as preocupações de ordem moral, estética ou de cultura geral. Aprender a utilizar o

espírito crítico, perante tudo o que faz parte da nossa civilização, continua a ser um dos objectivos principais.

Ora, o que se torna mais difícil de avaliar são, precisamente, aquelas qualidades humanas a que mais nos devemos dedicar: capacidade de resistência e coragem perante a agressão, real investimento na vida profissional, sede de justiça e sentido social, solidariedade, dedicação ao bem comum,...

Há outro aspecto que, geralmente, não é tido em conta, embora seja muito referido nos discursos: o contexto social em que a educação opera. Há que conhecer muito melhor a condição familiar do aluno e saber em que medida está preparado para aprender quando chega à escola. De uma maneira geral, temos de nos interrogar sobre o apoio social prestado aos jovens. Se pretendemos que o público compreenda bem o significado dos resultados escolares, e tome consciência das condições do meio em que a escola leva a cabo a sua missão, é essencial que os responsáveis sejam os primeiros a compreender melhor a natureza desse mesmo meio.

Por fim, há uma objecção de ordem prática que acaba de ser suscitada pelos professores ingleses, actualmente, como vimos, a braços com a instauração de um sistema de pilotagem muito desenvolvido. Não irão as avaliações ocupar tempo em prejuízo da acção educativa? Não irão os professores ser confrontados com técnicas que ignoram?

Salvo casos excepcionais, as avaliações sumativas apenas ocupam algumas horas. Quanto às avaliações formativas fazem parte integrante do processo de ensino e, desde que bem conduzidas, favorecem o sucesso escolar; fazem pois, pelo contrário, ganhar tempo.

Quanto aos aspectos técnicos inerentes às tarefas avaliativas pedidas aos professores, são fundamentais e devem fazer parte da sua formação docimológica básica. Na medida do possível, os professores devem poder construir os instrumentos de avaliação de que necessitam e, quando necessário, ser apoiados nesta tarefa. O que não quer dizer que a construção de certos instrumentos sofisticados, a sua aferição e validação, assim como certas análises estatísticas complexas, não sejam da competência de especialistas, a quem compete apresentar os documentos e interpretações de forma comprehensível, a todas a partes interessadas.

3. DESDE QUE SAIBAMOS SERVIR-NOS...

Na auditoria do sistema de ensino belga pedida à OCDE, R. Haby, R. in't Veld e J. Tschoomy escrevem (1993:114): "Não basta ter o mínimo de instrumentos à

disposição. É também necessário que se desenvolva entre os actores implicados um estado de espírito favorável à circulação dos dados, às permutas sobre o seu significado e eventual utilização."

Distinguem-se três níveis fundamentais de tomada de decisão: o processo racional, o processo organizacional e o processo político-burocrático.

É bem conhecido *o problema da racionalidade do comportamento*. Aplicada ao nosso caso, a questão pode formular-se do seguinte modo: a informação fornecida pelos indicadores basta, por si só, para levar à tomada das decisões mais acertadas e à sua execução? Esta é uma questão discutida, com muita clarividência, por A. Ruby (1991:15-16).

A informação, sublinha sucintamente Ruby, devia conduzir, quase necessariamente, a um raciocínio de tipo "profissional" e suscitar assim acções mais adequadas. Ora, trata-se de um acto de fé: acredita-se que as pessoas agirão de um modo racional, se dispuserem de informações suficientes para sustentarem as suas decisões.

Não se pode ignorar a importância dos dados qualitativos e o aspecto sempre parcialmente subjectivo da escolha dos objectivos da educação e da adesão a determinados valores. Estas opções desempenham um papel fundamental na definição de uma política da educação, na gestão das escolas e na condução das aprendizagens. Também já não se pode abstrair do papel desempenhado, na resolução dos problemas, pelos conhecimentos accidentais, não científicos, que têm a sua origem na vida do dia a dia.

Além disso existem, quase sempre, várias maneiras de resolver o mesmo problema. Depois, uma informação racionalmente elaborada é, apenas, um elemento num processo de decisão. Não deve ser tomada como prova ou como verdade, mas antes, como uma hipótese a experimentar, a testar e, finalmente, a rejeitar ou aceitar conforme abra ou não uma via possível ou desejável.

Ruby conclui: quem fizer uma análise desapaixonada, sabe que uma situação relacionada com a escola implica:

- Múltiplos decisores.
- Múltiplas e discutíveis definições de sucesso.
- Elevado grau de tolerância quanto aos modos de traduzir em factos as definições de sucesso.
- Uso de informações que estão longe de ser completas e perfeitas.
- Homogeneidade aparente, apesar da possibilidades de variação dos serviços oferecidos pelas escolas.

No *processo organizacional*, a avaliação torna-se uma prática rotineira que ajuda a atingir os resultados desejados. Há que recolher os dados adequados e recorrer aos métodos mais eficazes para obter as informações necessárias à pilotagem.

Quanto ao *modelo burocrático* transforma a avaliação numa arma política que se utiliza ou não, segundo a oportunidade. A avaliação é, assim, posta ao serviço de interesses particulares. Os "factos" só são tidos em conta se houver conveniência nisso.

Os três modelos coexistem a maior parte das vezes, mas com dominantes que vão variando.

Será que estas observações e estes aspectos, justificam a rejeição de uma pilotagem sistemática? Muito longe disso. O facto de reconhecermos as suas limitações não equivale a uma negação da necessidade de indicadores capazes de fornecerem sólidas informações. O que esta discussão deixa sobretudo transparente é a necessidade de um *pluralismo metodológico* na gestão e na acção.

As grandes decisões cabem aos poderes organizadores que se esforçam, mais ou menos, por ouvir a opinião do conjunto dos membros da comunidade educativa. Mas, sabemos muito bem que uma decisão de *cúpula* só terá efeito se as bases lhe compreenderem a lógica e o sentido, e estiverem realmente motivadas para passarem à acção pretendida. É por isso que é importante promover uma vertente interactiva da pilotagem.

Uma primeira medida consiste em associar as principais partes interessadas à escolha dos aspectos que irão ser objecto de observação. Os interesses podem, evidentemente, divergir. Assim, por exemplo, os poderes organizadores podem, em certos casos, considerar prioritários os custos e rendimentos globais, enquanto que os professores e os pais privilegiarão, sem dúvida, a qualidade do ensino e os meios concretos de a garantir na prática.

Associar os professores à organização da pilotagem é um ponto essencial. Não apenas porque conhecem melhor o terreno mas também porque, se tiverem participado na definição, recolha e interpretação dos indicadores, mais facilmente serão convencidos da sua utilidade. Como qualquer outra tecnologia também a pilotagem deve ser agradável para quem a utiliza.

M. Crahay propõe que à recolha e primeira análise dos dados comunitários ou regionais suceda, imediatamente, a criação de um grupo de reflexão composto por inspectores, responsáveis pela administração, professores, representantes dos pais. Juntamente com os avaliadores, este grupo analisaria os resultados do estudo, a fim de detectar os contributos mais importantes e definir linhas de aperfeiçoamento do ensino. O grupo poderia ainda sugerir novas prioridades a investigar. Os relatórios de avaliação, redigidos pelos especialistas, seriam

aperfeiçoados ou completados, em função das reacções dos diversos parceiros implicados no processo.

A esta primeira análise comunitária ou regional sucederia uma segunda fase de pilotagem a nível local. Teria por objectivo reforçar a participação no processo dos professores, dos responsáveis pela direcção das escolas e dos inspectores. Nas diferentes circunscrições dos serviços de inspecção haveria sempre escolas convidadas a participar na operação. Seriam organizados grupos de *pilotagem local*. Volta-se assim à problemática da micropilotagem, de que já nos ocupámos, e às disposições práticas anteriormente propostas. Estas são repetidas em parte.

Nos primeiros encontros seriam explicados os princípios da pilotagem. Seriam comentados os resultados das avaliações comunitárias ou regionais e analisados os instrumentos.

Estes instrumentos, eventualmente adaptados ou completados, seriam postos à disposição das escolas ou dos professores desejosos de conhecerem melhor a sua própria situação. Deverá estimular-se a adaptação às condições locais. Esta participa do *princípio de hibridação*, cuja importância para o desenvolvimento é, actualmente, reconhecida. Trata-se de identificar os elementos mais válidos em cada cultura e de os fazer interagir, a fim de desenvolver métodos e soluções originais. O que se aplica também ao encontro entre a cultura escolar dos agentes no terreno e a dos peritos. Pensa-se que, deixando livre curso à criatividade de cada um, se poderão aumentar as oportunidades de maximizar as vantagens da pilotagem.

Além disso, não se deve perder de vista que a pilotagem, em grande escala, se baseia em amostras de escolas ou de populações. Mesmo quando a representatividade é garantida, os dados obtidos têm, apenas, um valor médio. Daí, novamente, o interesse das acções locais utilizando eventualmente instrumentos especialmente adaptados ou concebidos para ajudar a dar resposta a interrogações específicas.

Só quando, à luz da observação e da reflexão, se começar a produzir nas aulas acções com um objectivo bem preciso, é que se poderá esperar um aumento da qualidade do ensino no sentido desejado.

4. ...RETENDO O QUE NOS CONVIER

A conclusão fica a cargo de J. Nisbet que escreve, resumidamente, sobre a matéria, num recente relatório da OCDE-CERI (1993:10).

Actualmente, insiste-se muito na solidez dos dados tanto no que respeita a indicadores de resultados como no que respeita aos métodos utilizados para

realizar as avaliações nacionais. Mas não se insiste, suficientemente, na melhoria da compreensão da natureza da aprendizagem conseguida nos últimos anos. "À semelhança do que se passa com a gestão de quaisquer interesses, a educação deve preocupar-se ao mesmo tempo com o controlo e com o crescimento. O Educational Testing Service dos Estados Unidos prevê que a avaliação em matéria de educação mude mais durante os próximos dez anos do que durante os anteriores cinquenta anos. É *importante que esta evolução não seja considerada isoladamente, mas antes como um elemento da reforma geral dos currículos, e que o objectivo de uma melhor aprendizagem se sobreponha sempre à preocupação de controlar*" (o itálico é nosso).

ANEXO: ASPECTOS TÉCNICOS

Introdução

O que os alunos aprendem, o que são capazes de fazer ou de ser, é ou deveria ser a preocupação principal entre os diferentes aspectos da pilotagem. É por essa razão que este anexo técnico é essencialmente dedicado a esses aspectos.

Avaliar o conjunto dos produtos do ensino é missão impossível, não só porque são inumeráveis mas também porque não se pode pretender que haja unanimidade de acordo sobre a sua importância respectiva. Só uma atitude pragmática possibilita a ultrapassagem deste grande obstáculo: entre as aprendizagens geralmente consideradas como fundamentais, tanto para a resolução dos problemas da vida quotidiana como para um desenvolvimento desejável da pessoa humana, devemos atribuir valor ao que o pode ter de uma maneira tão válida quanto possível. É o caso, concretamente, dos conhecimentos e das competências intelectuais básicas.

Quando falamos de aspectos técnicos da construção dos instrumentos de avaliação (testes, escalas de atitudes, listas de opiniões,...), de montagem de inquéritos ou investigações, de métodos de amostragem, de análises estatísticas, é toda a ciência da avaliação que está a ser implicada. Neste campo, só nos resta recomendar a consulta de obras da especialidade. Apenas nos referimos, aqui, a alguns aspectos particulares, alguns dos quais só começaram a ter mais importância num passado recente. O ponto de partida é *a avaliação dos programas*.

A maior preocupação continua a ser *a validade das medidas* e a sua *fidelidade*. A este respeito, o impacto dos recentes progressos da psicologia cognitiva e o movimento actual favorável à *medição dos resultados efectivos*, constituem importantes fontes de progresso.

O conceito de *validade pedagógica* surge em paralelo com o de *validade de conteúdos*: não é verdade que é importante, antes mesmo de avaliar as aprendizagens, verificar se o aluno teve ocasião de as realizar, se lhe foram ensinadas as matérias em questão ou se teve outras possibilidades de as estudar? Em resumo, terá tido o aluno *oportunidade de aprender*?

A administração de provas destinadas a fazer o balanço das aprendizagens de uma matéria vasta, por parte de um considerável número de alunos, não é possível sem o recurso à *amostragem matricial*. Além disso, a observação da evolução da mestria de determinada noção ao longo do tempo, exige a construção de *testes paralelos*. Neste ponto, como em muitos outros, a *teoria dos aspectos latentes* permite importantes progressos.

A *interpretação dos resultados* levanta também numerosos problemas, quer se trate de definir *resultados-padrão* ou *competências mínimas*, da *análise da percentagem de sucesso*, da *probabilidade de respostas correctas* ou da *abordagem das escalas*.

Porém, antes de abordar estes aspectos, é importante analisar com rigor a questão fundamental e geral das possíveis repercussões da pilotagem, em especial da avaliação dos riscos que pode implicar e das modalidades da indispensável busca de consenso, sobre este ponto. Voltamos assim à *avaliação social das opções tecnológicas* a que se fez breve alusão no capítulo sobre a natureza da pilotagem.

I. A AVALIAÇÃO DAS OPÇÕES TECNOLÓGICAS NA PERSPECTIVA DA SOCIEDADE

Este tipo de avaliação – idealmente com carácter antecipativo – parte da conclusão de que a tecnologia é determinada socialmente: não se desenvolve nem como uma força autónoma da sociedade, nem como uma espécie de via de resolução de problemas racionalmente seguida em função dos objectivos definidos. Produzem-se múltiplas interacções entre os grupos sociais, e entre estes e a tecnologia. Assim se explica que os diferentes grupos possam dar um sentido diferente aos fenómenos. É a partir daqui que se pode elaborar uma sociologia da tecnologia.

Uma das melhores salvaguardas contra os efeitos nefastos das tecnologias reside na possibilidade de as aperfeiçoar, em qualquer altura do seu desenvolvimento. Esta possibilidade é de uma importância evidente para a pilotagem.

Como outros efeitos da tecnologia, também os efeitos da pilotagem poderiam ser de tal modo subtils que corressemos o risco de nos passarem despercebidos. Por exemplo, não irá a pilotagem dar azo ao exercício de um poder oculto sobre os professores?

A avaliação das tecnologias na perspectiva da sociedade procura detectar os problemas que se põem ou que se podem pôr, avalia os diversos modos de actuar possíveis e apresenta conclusões.

Neste contexto, os investigadores os especialistas em avaliação podem prestar uma colaboração preciosa aos responsáveis políticos que, muitas vezes, optam por soluções sem estarem suficientemente informados das suas principais consequências ou utilizando apenas uma parte dos dados seguros fornecidos pela investigação.

Informar os decisores das diversas soluções possíveis é um dos importantes papéis que os cientistas podem desempenhar. É o que Vinck (1988:13) chama

apresentação das alternativas. É preciso, contudo, não perder de vista que, mesmo que os políticos estejam muito longe de poderem ajuizar sempre da qualidade real das soluções técnicas que lhes são propostas, dispõem de informações sobre a sociedade que muitos investigadores ignoram (aceitação psicológica por parte do público ou no seio dos partidos, limites orçamentais, compromissos políticos e filiosficos,...).

Considerada sob o aspecto negativo, esta avaliação ajuda a que se esteja precavido contra consequências indesejáveis. Considerada sob o aspecto positivo, ela pode confirmar o interesse das tecnologias, levar o público a aceitá-las e chamar a atenção dos cientistas para as suas responsabilidades em relação à sociedade.

"A avaliação na perspectiva da sociedade das opções tecnológicas, implica a participação dos diferentes actores implicados no processo. É pois necessário criar espaço para um debate permanente" (Vinck, 1988:15).

A ética *etnográfica* estuda a co-emergência dos valores e das práticas nas relações de trabalho. Observa, em especial, quais os códigos e valores que surgem aquando da utilização das tecnologias.

A ética *estratégica* está orientada para as estratégias desenvolvidas pelos actores implicados para definirem as controvérsias como controvérsias éticas e para procurarem um consenso neste debate.

Esta procura de consenso deve continuar a ser uma preocupação permanente.

Referindo-se a um relatório entregue ao Parlamento Europeu em 1991 pela Direcção Geral XII, Programa FAST, da CEE, J. Berleur (1991) menciona métodos de avaliação que, no seu conjunto, devem ajudar a perceber as vantagens e os perigos das tecnologias e possibilitar que toda a sociedade, e não apenas os especialistas e decisores, possam dar a sua opinião sobre este assunto:

- Análises da literatura.
- Estudo retrospectivo e prospectivo de séries estatísticas.
- Investigações junto de pessoas expostas às novas tecnologias.
- Estudo de casos.
- Pareceres de peritos.
- Desenvolvimento de cenários possíveis.
- *Grupos de trabalho do futuro (future shops):* reflexão em grupos de trabalho de vinte cidadãos, não peritos, sobre o que poderá vir a ser o futuro. Esta reflexão comporta a avaliação da possibilidade de realização e propostas de acções.
- *Conferências de consenso:* não especialistas, vindos de horizontes diversos, começam por escutar a breve apresentação de uma problemática feita por

peritos. Os "não peritos" analisam as suas teses e propostas e têm de pôr-se de acordo quanto a um relatório final. Os peritos tomam conhecimento deste relatório, podendo apenas corrigir os erros factuais. O texto é, em seguida, apresentado aos meios de comunicação social.

– Experimentação social em situações reais.

Num estudo metodológico publicado, N.T. Nguyen, C. Lobet, J. Berleur e B. Kusters observam que "a ideia de uma avaliação das tecnologias na perspectiva da sociedade neutra e utilizando apenas um método, parece definitivamente abandonada, em favor de uma abordagem mais orientada para a constituição de parcerias, para o *diálogo* e a *concertação*". Os mesmos autores sublinham ainda (p.7) que um dos maiores desafios postos a quem avalia tecnologias, é o "de se ser capaz de informar, correctamente e *a tempo*, todas as partes implicadas". Este é um aspecto muito importante porque, sobretudo durante a fase de desenvolvimento da tecnologia, são possíveis várias vias, e como se não conhece ainda qual delas prevalecerá, não se pode dizer grande coisa sobre os impactos na sociedade. Mais tarde, quando a tecnologia já estiver bem instalada, "as implicações sobre a sociedade surgem com clareza mas, infelizmente, é já demasiado tarde para influenciar o seu desenvolvimento."

A Célula *Interfaculdades de Avaliação Tecnológica* (CITA) das Faculdades Notre-Dame de la Paix, em Namur, que dirige cada vez mais os seus esforços para a explicação das controvérsias sociais que surgem por ocasião de uma inovação tecnológica, distingue cinco tarefas principais (p. 9-11):

1. Identificação dos conflitos entre as *diversas percepções sociais* e as *concepções* da tecnologia.
2. *Descrição minuciosa* das controvérsias referentes à escolha da tecnologia e da sua evolução.
3. Elaboração de *mapas sociais* mostrando as diversas categorias de actores, directa ou indirectamente relacionados com a inovação tecnológica.
4. Proposta, a partir das hipóteses formuladas aquando das tarefas anteriores, de *cenários alternativos* e promoção do processo de *participação pública*.
5. De forma mais empírica, a realização aprofundada de *estudos de casos* ou de *experimentações sociais*, em colaboração com os actores directamente implicados no desenvolvimento tecnológico.

A palavra-chave deste processo é a *participação* e, metodologicamente, os processos qualitativos e quantitativos acabam por se combinar.

2. A AVALIAÇÃO DE PROGRAMAS

A técnica da pilotagem deriva da avaliação de programas.

Programa define-se, de uma forma geral, como um conjunto integrado de actividades concebidas para se atingir um objectivo comum. O objecto da avaliação é estabelecer em que medida os efeitos observados correspondem aos objectivos a atingir e, eventualmente, se esses mesmos objectivos devem ou não ser modificados.

Em matéria de avaliação de programas encontramos todas as tendências da investigação avaliativa contemporânea: perspectivas quantitativas ou qualitativas, nomotéticas, históricas ou ecléticas. A análise "custos-benefícios" constitui uma categoria particular.

Aplicando planos experimentais ou quase experimentais, a *avaliação quantitativa* estuda as relações entre as variáveis de entrada (*input*) e as variáveis de saída (*output*) antecipadamente definidas, a fim de determinar a produtividade global, bem como a dos diferentes "investimentos". O produto é, a maior parte das vezes, expresso em termos de rendimento escolar, de continuação dos estudos ou de empregos obtidos logo no final do programa.

Muitos indicadores têm a sua origem em simples dados numéricos: valores orçamentais, número de alunos ou de professores, duração do ano escolar, taxa de absentismo, percentagem de determinado grupo etário que obtém diploma de final de estudos secundários de longa duração, etc. É importante, evidentemente, que saibamos lidar com estes dados correctamente e não queiramos que os números digam coisas que eles de facto não contêm. Deve, por exemplo, determinar-se sistematicamente o significado estatístico das diferenças de percentagem. Há que interpretá-las correctamente, mesmo quando essas diferenças forem grandes.

Na *perspectiva qualitativa* o objectivo a atingir é, antes de mais, compreender a que se devem realmente os efeitos observados. Considera-se que os objectivos da avaliação muitas vezes só se definem com precisão no decorrer do processo. Dá-se muita importância à descrição, o mais minuciosa possível, das circunstâncias em que se inicia o processo e do que acontece na fase de concretização. No arsenal do avaliador a etnometodologia ocupa um lugar cada vez maior. Há um esforço por descobrir como é que em determinado meio particular cada um dos actores implicados viveu o programa, o que é que, efectivamente, retirou dele, e as questões que coloca a seu respeito. Daí o recurso frequente ao estudo de casos e à *observação participante*.

Uma das principais críticas feitas a este processo, é a de não colocar no primeiro plano das preocupações as razões que levaram ao lançamento do programa. Ora, em verdadeira democracia, a decisão resulta de um acordo de natureza política e social feito pelo conjunto de representantes da comunidade educativa. Só dificilmente se admite que, aquando da avaliação do programa, se possa abstrair das intenções iniciais.

Na investigação educativa, a oposição entre avaliação qualitativa e quantitativa, alimentou um debate apaixonado durante mais de um século. Embora subsistindo ainda, a controvérsia amainou, pelo menos. Não é verdade que em toda a qualidade existe sempre uma certa quantidade?...

Como o seu nome indica, a *perspectiva eclética* procura tirar proveito, simultaneamente, das duas perspectivas anteriores: a procura das relações causais (análise quantitativa) é completada por análises de processo e de contexto (análise qualitativa). São tidos em conta não só os objectivos assinalados inicialmente, mas também os que se lhes vieram juntar ou os vieram substituir, no decurso dos trabalhos, quando se viu interesse nisso.

A *análise custos-benefícios* derivada da economia, especialmente das teorias de gestão, é predominantemente quantitativa. Em geral este tipo de avaliação não tem grande sucesso na área da educação devido, exactamente, à forte componente afectiva que esta comporta. Por exemplo, para além do benefício material imediato que resulta da supressão do insucesso escolar, este facto tem outras consequências importantes: melhor auto-imagem, aumento da motivação para prosseguir os estudos, maior sucesso na vida adulta, etc. Ora estes "benefícios", provavelmente os mais importantes, são difíceis de quantificar.

Actualmente, a realização de trabalhos de pilotagem é confiada, a maior parte das vezes, a organismos centrais. É concretamente o caso dos Estados Unidos, da Inglaterra e da França. Estes organismos são quer serviços da administração central (por exemplo, a Direcção de Avaliação e Prospectiva em França depende do Ministério de Educação Nacional), quer centros de investigação e de desenvolvimento (por exemplo, o *Educational Testing Service*, ETS, nos Estados Unidos). Pode, contudo, conceber-se uma descentralização desde que se garanta uma rigorosa coordenação.

Praticamente nenhum organismo de pilotagem funciona sem recorrer a peritos exteriores e a serviços de investigação em educação (que intervêm, a maior parte das vezes, em resposta a apelos que lhes são dirigidos).

Sobre a avaliação qualitativa deverá consultar-se a obra fundamental de M.B. Miles e A.M. Huberman, *Analyse des données qualitatives*, Bruxelas, De Boeck, 1990.

3. PSICOLOGIA COGNITIVA E NOVAS PERSPECTIVAS

Os recentes progressos da psicologia cognitiva abrem novas perspectivas tanto para a avaliação prognóstica como para as provas de rendimento ou de diagnóstico. Embora as noções teóricas a seguir referidas – de forma muito simplificada – não possam, ainda, traduzir-se facilmente em práticas quotidianas, ajudam pelo menos a tomar consciência da problemática fundamental.

Em vez de se limitar a verificar que determinado aluno não é capaz de encontrar a solução de um problema, de levar a cabo determinada tarefa, a psicologia quer antes saber qual o processo mental que explica esse fenómeno. Graças sobretudo aos trabalhos de Piaget, qualquer educador sabe hoje que as possibilidades do aluno estão relacionadas com o estádio de desenvolvimento cognitivo, afectivo e psicomotor que ele atingiu. Porém, não se pode explicar tudo apenas através do conhecimento do nível de desenvolvimento cognitivo.

Comparando com o que se passa com os computadores (metáfora que a psicologia cognitiva privilegiou, durante algum tempo), o tratamento da informação pelo cérebro pode ser descrito da seguinte maneira:

- As estruturas do sistema cognitivo podem ser activadas por estímulos recebidos pelos órgãos dos sentidos e por outras estruturas do conhecimento (pensamento); dá-se assim a *entrada* de informações.
- A *memória de curta duração* retém um número limitado de informações, durante um curto período de tempo. Os factos isolados são os que se esquecem em primeiro lugar.
- A *memória de longa duração* armazena, de forma codificada e permanente, os conceitos e construções mentais derivados da experiência passada da pessoa (acontecimentos, relações, processos, afectos, ...) e dirige as operações de todo o sistema de tratamento da informação. Esta *base de dados* constitui o coração do sistema cognitivo. A capacidade deste tipo de memória é ilimitada.
- O *sistema muscular* executa os actos motores: falar, correr,...

O extraordinário nível de desempenho do cérebro só se explica devido à existência de:

- processos de tal modo rápidos que já não exigem muita atenção (automatismo);
- estratégias que permitem seleccionar, planificar, exercer a pilotagem e, se necessário, modificar as acções (competências metacognitivas);

- estruturas cognitivas (*esquemas*) que especifiquem as propriedades gerais de um objecto ou de um acontecimento. Trata-se, pois, de abstracções que representam conjuntos significativos de informações armazenadas na memória (por exemplo todo o normal desenvolvimento de uma operação cirúrgica).

A aprendizagem consiste na automatização de processos, na aquisição e desenvolvimento de competências metacognitivas e na construção, revisão e substituição de esquemas.

Em relação aos “noviços” na matéria, os “peritos” caracterizam-se pelo número crescente de conexões entre uma quantidade de factos cada vez maior, pela amplidão e qualidade dos esquemas que organizam e por uma avançada automatização de processos, o que lhes permite encarar novos aspectos dos problemas e colocar-se em perspectivas diferentes.

Quanto às *representações* são “teorias pessoais”, sistemas de crenças relativos a fenómenos. É o modo como um indivíduo ou um grupo os comprehende ou explica. Por exemplo, acreditar que a electricidade corre nos fios como a água nos tubos porque em ambos os casos se fala de *corrente*... Muitas vezes fortemente arraigadas, as representações não são fáceis de desalojar nem de avaliar. Quanto mais incorrectas e inadequadamente articuladas mais dificultam as novas aprendizagens.

3.1. Implicações na melhoria dos testes de rendimento escolar

À luz dos contributos da psicologia cognitiva, a avaliação escolar deveria desempenhar as seguintes funções (H. Walberg, 1991):

- Medir a quantidade de conhecimentos declarativos (factos, conceitos, fórmulas,...) que o aluno possui e identificar o tipo de organização que utiliza para os guardar.
- Medir a velocidade de execução das tarefas cognitivas. A taxa de atenção requerida pode ser manipulada, mediante a introdução de tarefas concorrentes, ou devido ao aumento de dificuldade das tarefas.
- Fornecer rapidamente um feedback aos alunos e chamar a sua atenção para os erros de processo ou de conteúdo.
- Diagnosticar o tipo de erros e as suas fontes.
- Determinar, pela análise dos erros, o grau de adequação, de extensão e de flexibilidade dos esquemas.

- Identificar as estratégias de resolução dos problemas.
- Observar em que medida o aluno é capaz de variar as suas estratégias de resolução dos problemas, quando não encontra as soluções directamente.
- Identificar as estratégias metacognitivas utilizadas pelo aluno; identificar as que o ajudam e as que não o ajudam. Há que distinguir as competências metacognitivas gerais e específicas de determinado conteúdo.
- Para avaliar a qualidade das estratégias de resolução dos problemas, ter como referência o modo como os peritos encaram o problema, elaboram e optam por um plano de resolução, e o modo como põem em acção a estratégia adoptada.
- Avaliar a qualidade das redes semânticas.
- Avaliar o grau de automatização das competências intelectuais.

3.2. Em conclusão

- Há técnicas de avaliação que permitem um certo conhecimento dos processos cognitivos do aluno.
- Impõe-se a utilização de testes de conhecimentos de múltiplas facetas (ver mais adiante a nota sobre a capacidade de generalização). Em especial, o modo de pôr as questões deve ser variado, assim como o nível de dificuldade. Pode, deste modo, avaliar-se em que medida a maneira de colocar as questões influencia os resultados observados.
- A avaliação deve ser dinâmica. Não nos devemos contentar com observar o que os alunos são capazes de fazer, por si mesmos, num dado momento. Há que determinar, também, até onde é que eles podem ir, uma vez postos no caminho da solução (estudo do *desenvolvimento proximal*, segundo a expressão de Vygotsky, 1978).
- Impõe-se contudo uma atitude de prudência, em relação a todas estas propostas. Há poucos estudos estatísticos sobre o assunto. Deve determinar-se a validade de conteúdo e a capacidade de generalização das provas. Além do mais, as diligências propostas implicam custos.

4. VALIDADE E FIDELIDADE DAS MEDIDAS

Como qualquer outra medida, os indicadores devem possuir duas qualidades essenciais sem as quais deixa de ter muito sentido qualquer outro tipo de considerações: validade (quer dizer, medir efectivamente o que pretendem medir) e fidelidade.

A validade interna de um indicador fica ameaçada se, por exemplo operarmos apenas com uma medida parcial de um fenómeno multidimensional. Se, por exemplo, reduzirmos um teste de inteligência a provas de memória é evidente que os resultados obtidos não reflectirão, verdadeiramente, o nível de inteligência da pessoa. De igual modo, um indicador que trate apenas de um aspecto marginal de um fenómeno educativo, não oferece muito interesse e arriscamo-nos a que nos conduza a conclusões incorrectas.

O mesmo se passa com as relações que se podem estabelecer entre certos indicadores. Assim, por exemplo, tomar a taxa de repetências como indicador da qualidade pedagógica sem ter em conta quaisquer outras considerações, corre o risco de ser um procedimento falacioso, enquanto a decisão de reter ou não um aluno depender, em última análise, de um poder arbitrário que se ignora.

A validade fica, pois, falseada:

- se o fenómeno a avaliar for definido de forma inadequada;
- se a variável que veio substituir outra por razões de segurança, não for válida. Por exemplo, o número de anos de estudo da formação inicial dos professores, não nos indica rigorosamente nada de certo sobre a sua qualificação (este número de anos não pode servir como variável substitutiva);
- se o objecto da medida mudar com o decorrer do tempo;
- se houver “corrupção”. Esta falsificação surge, por exemplo, quando o professor sopra as respostas às questões de um teste.

Quanto à fidelidade está evidentemente relacionada com a validade: se as repetidas pesagens de um mesmo objecto indicam pesos diferentes, carecem não só de fidelidade como de validade. Temos muito que aprender a este respeito com a teoria da capacidade de generalização.

4.1. Capacidade de generalização

Por capacidade de generalização de uma medida entende-se o nível a que se pode, a partir de uma observação ou de uma prova particular, tirar conclusões sobre o valor teórico pretendido. Mais concretamente, permitirão as questões postas avaliar o nível de conhecimentos dos alunos de forma fiel? Deparamo-nos aqui com o problema da fidelidade e da validade.

Os resultados observados num exame variam em particular:

- de acordo com o conteúdo e a forma das questões colocadas (faceta questões);
- de acordo com o momento em que a prova é aplicada (faceta ocasiões);
- de acordo com os pontos de vista e qualidade dos examinadores (faceta juízes).

Por exemplo, a propósito da faceta questões há que considerar:

1. A *linguagem utilizada*

- Influência do vocabulário utilizado e da sintaxe (legibilidade psicológica).
- Modo de apresentação material e psicológica. Por exemplo, embora tratando exactamente do mesmo problema, as questões seguintes podem ter resultados diferentes:
 - $3 + 5 = ?$
 - Se comprar um lápis por cinco francos e um bombom por três francos, quanto tenho de pagar?

2. A *natureza da questão*

J. Cardinet gosta de dar o seguinte exemplo, a propósito do cálculo da área do triângulo.

Num grupo que estudou este assunto durante o ano lectivo, obtêm-se diferentes resultados conforme se recorre a um ou a outro dos seguintes modos de pôr a questão:

- Qual é a fórmula para calcular a área do triângulo?
- Qual é a superfície de um triângulo, com oito metros de base e cinco de altura?
- Tem aqui um triângulo e uma régua graduada. Calcule a superfície deste triângulo.
- A questão é idêntica à anterior mas a forma do triângulo é tal que a perpendicular traçada a partir do vértice superior não faz intercepção com a base.

Em resumo, o conhecimento de uma fórmula e da sua utilização em casos esteriotipados não é prova suficiente de uma capacidade real. Uma avaliação

válida só é possível com resultados concretos suficientemente variados nas suas condições para constituirem uma amostra representativa dos vários casos que podem colocar-se.

A propósito desta faceta "ocasiões" podem, também, colocar-se as seguintes questões:

- Teve o aluno, efectivamente, ocasião de aprender a matéria em questão?
- Essa aprendizagem é recente ou já antiga? Neste segundo caso, o conhecimento foi reactivado ou permaneceu adormecido?
- Qual seria o grau de fadiga do aluno quando fez a prova?
- A mesma questão é colocada de diversas formas?

Quanto à faceta "juízes", é muitas vezes referida em docimologia: inconsistência das apreciações, efeito de halo, efeito de contraste, estereotipia, efeito de fadiga do corrector ou do interrogador,...

4.2. Estimativa da ocasião de aprender

Antes de alguém se pronunciar sobre um resultado aparentemente insuficiente, é conveniente verificar se o aluno teve, efectivamente, oportunidade de aprender os conhecimentos ou competências testadas.

Para fazer uma estimativa de qual a percentagem de alunos que receberam um ensino adequado pode-se, quer interrogar separadamente os professores, quer pedir ao conjunto dos docentes implicados que se ponham de acordo acerca de uma estimativa. Nos trabalhos da NAEP, só são mantidas as questões que abordam matérias que, pelo menos 75% dos alunos, tiveram oportunidade de aprender. E de novo são utilizados tratamentos estatísticos complexos (entre os quais a estatística de Mantel-Haenszel) para exprimir os diferentes aspectos dos resultados obtidos inicialmente.

4.3. Provas de rendimento ou provas de memorização?

Se os sistemas de indicadores devem ajudar a centralizar a acção educativa sobre os aspectos mais importantes do saber-fazer, então há que medi-los verdadeiramente. Isto só se consegue num contexto de validade. O debate sobre este assunto é particularmente vivo nos Estados Unidos onde muitos apenas reconhecem as provas de rendimento, isto é provas que obrigam o aluno a produzir de facto o comportamento em questão ou, pelo menos, a construir a sua

resposta, não indo apenas buscá-la, já pronta, à memória, nem recorrendo a processos rotineiros e estereotipados de resolução dos problemas.

Não se coadunam com provas de rendimento questões como as seguintes:

- Qual a fórmula do ácido sulfúrico? (Pura memória).
- Sabendo que a área do triângulo se calcula de acordo com a fórmula $base \times altura / 2$, qual a área de um triângulo com 25 cm de base e 30 cm de altura? (Compreensão elementar).
- Calcule a área de um triângulo com 25 cm de base e 30 cm de altura. (Aplicação estereotipada).

Na taxonomia de B.S. Bloom (domínio cognitivo), e de acordo com o sentido convencional que aí se atribui aos termos *memória*, *compreensão*, *aplicação*, estas três questões fazem apelo a processos cognitivos inferiores.

Não se coadunam igualmente com provas de rendimento, exames em que o indivíduo sabe responder por ter aprendido de cor o que se deve fazer em determinado caso, mas em que o não faz efectivamente. Expor, oralmente ou por escrito, a pedagogia progressista que se praticaria se se fosse professor, não prova nem a existência de capacidade para a pôr em prática, nem a possibilidade de fazê-lo, dadas as limitações circunstanciais que existem, na maior parte das situações escolares.

Será que os questionários de escolha múltipla excluem, na verdade, qualquer avaliação de rendimentos relacionada com processos cognitivos superiores? Concreta que não. A questão seguinte, por exemplo, exige claramente o recurso a processos cognitivos superiores e equivale pois a muitas provas de avaliação de rendimento.

O soalho de um quarto deve ser inteiramente coberto de tacos de madeira. Para isso são necessários M tacos de a cm por b cm. Quantos tacos de x por y cm serão precisos?

A	B	C	D	E
$\frac{Ma\bar{b}}{xy}$	$\frac{ab}{Mxy}$	$\frac{(a + b) M}{x + y}$	$\frac{ab \cdot xy}{M}$	$\frac{Mxy}{ab}$

A dificuldade de construção de provas de avaliação de rendimentos está, pois, na concepção de questões ou de tarefas que permitam avaliar, de forma válida, o que o indivíduo é, efectivamente, capaz de fazer.

No que respeita a formas de testar rendimentos, oferecem-se-nos três grandes vias que começam a ser exploradas na área dos indicadores, sobretudo em vários Estados norte americanos (Vermont, Nova Iorque, Connecticut) e no Reino Unido:

1. Quanto às disciplinas teóricas (filosofia, lógica, álgebra, física,...), pôr questões que exijam o recurso a processos cognitivos superiores.
2. Demonstrar a competência considerada como objectivo (por exemplo, dar uma aula de língua estrangeira perante os examinadores ou, outro exemplo, pedir aos alunos que façam uma experiência científica, que pode ser demorada e que elaborem um relatório da maneira como ela decorreu).
3. Examinar produtos reais através dos quais a qualidade da competência pode ser avaliada. Trata-se, neste caso, do *exame de um 'dossier'* (o *portfolio* dos anglo-saxões). Por exemplo, para valorizar as suas capacidades pedagógicas, um professor pode constituir um *dossier* em que reuna a preparação das suas aulas, material didáctico por ele construído, gravações vídeo das aulas, trabalhos dos alunos, registos de conversas tidas com os pais, artigos publicados, etc.

O recurso a *dossiers* para avaliar os alunos, oferece pistas interessantes mas não deixa de ter problemas. De facto, a qualidade do *dossier* varia, não só em função das aquisições dos alunos mas também de acordo com a sua motivação, a sua maior ou menor compreensão do uso que será dado ao *dossier* e do modo como este irá ser avaliado. Observam-se também diferenças segundo os estilos cognitivos.

Muito embora correspondam a um ideal, as provas de avaliação de competências reais levantam grandes problemas práticos, a começar pelo muito tempo que é necessário para uma observação directa satisfatória da execução de tarefas complexas como, por exemplo, a realização de experiências científicas seguidas de interpretação das mesmas. Não basta apenas uma experiência. São necessárias várias para se obter uma amostra realmente representativa que possibilite uma avaliação do saber e do saber-fazer dos alunos, com fidelidade e validade.

4.3.1. Necessidade de processos de moderação

A utilização de testes de avaliação de competências para fins comparativos (a nível regional, nacional, internacional) levanta também problemas ainda mal

resolvidos. De facto, a diferença de tarefas torna difícil a comparação. Além disso, o grau de familiarização dos alunos com este tipo de provas influencia os resultados. A ocasião de aprender varia também.

No Reino Unido – que adquiriu grande experiência no assunto – recorre-se, sistematicamente, a processos de “moderação” das notas (estas técnicas vêm descritas no nosso *Resumo de docimologia*, 1992).

Professores e especialistas na construção de currículos reunem-se para analisar os trabalhos dos alunos, para determinar em que medida se podem comparar as tarefas propostas, as exigências e os processos de avaliação.

O facto de associar os professores aos processos de moderação, leva à melhoria da sua qualificação profissional e torna possível a comparação entre tarefas e critérios de correção.

Sobre os *portfolio* de competências profissionais ver J. Aubret (1992).

5. A ADMINISTRAÇÃO DAS PROVAS

5.1. Amostragem multimatricial

Para verificar em que medida os alunos realizaram as principais aprendizagens previstas nos programas, a primeira ideia que surge é testar toda a população sobre o conjunto das matérias. Ora isto é praticamente impossível.

A amostragem multimatricial consiste em constituir amostras ao mesmo tempo de alunos e de tarefas. Por exemplo, divide-se o conjunto das questões a colocar em vinte grupos de tarefas (*itens*). Se se acha, por exemplo, que o exame de uma amostra de trinta turmas permite fazer uma ideia correcta do rendimento num conjunto de tarefas, haverá que testar seiscentas turmas.

Os resultados de conjunto reunidos dão informações sobre a percentagem de sucesso por grupo de tarefas e sobre a percentagem de alunos que atingem os níveis de desempenho estabelecidos. Avalia-se, pois, o rendimento dos grupos e não o dos indivíduos.

Podem também determinar-se índices de evolução se se conservar uma parte comum de tarefas nas avaliações seguintes.

5.2. Comparação de testes

Para comparar, à medida que o tempo passa, os progressos realizados na aprendizagem de uma mesma matéria, pode-se, à semelhança do que se faz

correntemente em estudos longitudinais, conservar um núcleo de questões comuns nas provas administradas a diferentes níveis escolares (*comparação vertical*).

A *comparação* diz-se *horizontal* quando os testes são aplicados ao mesmo nível de ensino e, eventualmente, aos mesmos alunos, o que implica a construção de testes paralelos, isto é, equivalentes do ponto de vista de conteúdo e de dificuldade.

A *comparação (equating)* baseada na hipótese da existência de traços latentes é um método recente. Por exemplo, A. Beaton (1987) utilizou-a na *comparação vertical* de provas de leitura da NAEP (9, 13 e 17 anos).

Sobre a teoria das respostas aos *itens* (IRT) e sobre a noção de traços latentes ver, mais adiante, a nota sobre a abordagem das escalas.

6. A INTERPRETAÇÃO DOS RESULTADOS

6.1. Noção de padrão

Alguns indicadores não passam de simples listas de dados numéricos. Nos casos mais elementares limitamo-nos a verificar a existência ou não de determinado fenómeno. Exemplo: percentagem de determinado grupo etário que prossegue os estudos para além do ensino secundário de longa duração.

Mas quando se trata do progresso na educação, do desenvolvimento do indivíduo, deve haver a possibilidade de recorrer a referências descritivas, de definir níveis a atingir. É a estas referências qualitativas que, geralmente, se dá o nome de *padrões*.

Um *padrão educativo* consiste na definição, não ambígua, de um *corpus* de conhecimentos que devem ser ensinados, ou na especificação dos saberes, do saber-fazer e do saber-ser a adquirir pelos alunos em diferentes momentos da sua escolaridade (final de ano, final de ciclo,...).

A noção de padrão educativo pode estender-se à qualidade dos diferentes aspectos da ação educativa: formação de professores, manuais, locais, equipamento,...

Quando se trata de avaliações qualitativas os padrões surgem, muitas vezes, na forma de escalas descritivas. Eis um exemplo rudimentar:

Indicador: Qualidade da organização no ensino da leitura

1. Medíocre: Todos os alunos seguem o mesmo ritmo. Não há trabalho de grupo.

2. Satisfatório: Como no nível I, mas um aluno muito lento recebe, por vezes, um pouco menos de trabalho do que os outros.
3. Médio. Formação de dois ou três grupos de nível. Pouca flexibilidade na formação dos grupos.
4. Muito bom: Formação de grupos de acordo com os níveis de competência. Flexibilidade.
5. Excelente: Formação de grupos após um estudo aprofundado das aptidões e das dificuldades sentidas. Grande flexibilidade.

Além disso, podem também definir-se níveis mínimos de aprendizagem a atingir. Neste caso estamos a falar de *padrões de competência* e deparamos, assim, com o difícil problema da definição objectiva de *competência mínima*.

Até há pouco tempo, os programas e o discurso sobre educação, preocupavam-se mais com as intenções, com o que se desejava ver aprendido, do que com os níveis efectivos das aprendizagens a realizar. A pilotagem não consente que se continue a negligenciar este aspecto.

6.2. Determinação da competência mínima

Qualquer prova termina com a verificação do sucesso ou do insucesso. A *competência mínima* marca o limite entre ambos. Defini-la objectivamente não é tarefa fácil.

Há já alguns anos que a questão da competência mínima suscita debates apaixonados. De facto, apesar do notável prolongamento da escolaridade nos países industrializados, verifica-se que muitos alunos abandonam a escola sem dominarem competências básicas como a leitura, a expressão escrita e a aritmética elementar. Ora, um domínio satisfatório destas competências é mais que nunca necessário, quer no mundo empresarial, quer na vida quotidiana dos cidadãos. Esta a razão por que surgem, em numerosos países, normas regulamentadoras subordinando, por exemplo, a atribuição de um diploma de ensino secundário, à passagem numa prova preliminar, onde se verifique se o aluno domina, de forma satisfatória, as competências tidas como essenciais.

A questão escaldante é: como fixar, de forma válida, o ponto ou a zona a partir dos quais se pode afirmar que alguém é "minimamente competente"?

6.2.1. Fixação da nota de sucesso

Tem-se ou não sucesso num exame, ou num conjunto de exames? A nota de sucesso corresponde ao rendimento mínimo aceitável, isto é, à competência

mínima. Numa perspectiva de pedagogia para a mestria, deparamos com um problema semelhante quando se pretende saber a partir de que momento o aluno atinge este estado de mestria.

Há várias formas de designar este ponto crítico: *nota de sucesso, nota de ruptura, nível de sucesso, nível de exigência mínima,...* A fixação desta nota levanta, quase sempre, um enorme problema. Como determinar, por exemplo, a partir de que momento se conhece uma língua estrangeira?

A questão só é simples se a avaliarmos – se tal for possível – em termos de tudo ou nada. Assim, podemos decidir que um aluno não terá sucesso enquanto cometer um ou vários erros na multiplicação dos dez primeiros números. Que só se é bom piloto-aviador, a partir do momento em que se efectuem com sucesso *todas* as aterragens.

Desde que se admita uma parte de erro, e que a aquisição a avaliar seja complexa, a nota de ruptura deverá flexibilizar-se numa espécie de ramificação, no âmbito da qual seja possível o sucesso. Fala-se, neste caso, de *zona de ruptura*.

A nota de sucesso é quase *absoluta* e determinada *a priori* (a nota mínima é fixada antes da prova, em relação a um objectivo a atingir), ou *a posteriori* (tomam-se como referência, por exemplo, os resultados de um grupo exterior ou os do próprio grupo).

I. DECISÕES EMPÍRICAS

I.1. NOTA DE SUCESSO NOS EXAMES TRADICIONAIS

Uma das regras mais frequentes nas escolas belgas era (e, por vezes, ainda é) obter pelo menos 50% em cada disciplina e 60% no conjunto. Não se apresentava qualquer justificação objectiva para estes níveis, a não ser razões para colocar ou não todas as disciplinas ao mesmo plano.

Entre os processos mais sistemáticos para fixar a nota de ruptura, Ebel (1979) propõe os seguintes:

a) *Para os testes de escolha múltipla*

Partindo do princípio de que, num teste de escolha múltipla bem elaborado, um aluno ignorante pode obter um resultado que corresponda a respostas dadas inteiramente ao acaso (50% nas respostas verdadeiro-falso; 25% nas respostas de quatro escolhas, etc.), e que um aluno muito bom, deverá atingir quase o máximo,

Ebel pensa que a nota de ruptura pode, neste caso, ser calculada do seguinte modo:

ESCOLHA	Quatro Escolhas	Verdadeiro-Falso
Número de itens	100	100
Nível obtido respondendo ao acaso	25	50
Nível médio ideal	$(100 + 25)/2 = 62,50$	$(100 + 50)/2 = 75$
Nível de ruptura	$(25 + 62,50)/2 = 43,75$	$(50 + 75)/2 = 62,50$

O nível médio ideal situa-se, pois, a meio caminho entre o nível máximo possível e o nível previsto, no caso de se responder inteiramente ao acaso.

O nível de ruptura situa-se a meio caminho entre o nível médio ideal e o nível previsto, no caso de se responder inteiramente ao acaso.

b) Para os exames tradicionais

O nível de sucesso é fixado, arbitrariamente, numa dada percentagem; é em seguida ajustado em função dos resultados observados.

Exemplo: a estimativa é que 3/4 dos alunos dêem respostas correctas, desde que no mínimo 60% e no máximo 80% deles ultrapassem este nível.

Se pelo menos 60% dos alunos obtiverem 75% de sucesso, o nível de sucesso é fixado, exactamente, a meia distância entre o nível correspondente a 75% e o nível que, pelo menos 60% dos alunos consigam ultrapassar.

Se mais de 80% dos alunos ultrapassarem o nível de sucesso, este é situado a meio caminho entre o nível correspondente a 75% e o nível que 80% dos alunos consigam ultrapassar.

Nota Ebel (1979:341): "60% e 80% são números considerados a título de exemplo. Podem ser aumentados ou diminuídos conforme as circunstâncias. O objectivo a alcançar na fixação dos números é o de manter o nível de conhecimentos e a percentagem de sucesso dentro de limites razoáveis, tendo também em devida conta o erro de medida. Este método não é frequentemente utilizado e, contudo, tem o mérito de constituir uma solução racional para um problema difícil."

I. 2. NOTA DE SUCESSO OU PROVAS DE COMPETÊNCIA MÍNIMA

Na pedagogia para a mestria o nível de sucesso exigido situa-se, em princípio, nos 100%. Esta exigência é muitas vezes suavizada quanto mais não seja para se terem em conta os erros fortuitos. Em geral admitem-se os 90% ou, até mesmo os 80%, mas menos não.

Na avaliação normativa acontece que o nível fixado corresponde ao rendimento médio de um grupo de referência. Por exemplo, V. de Landsheere (1988:124) refere que, em vários Estados americanos, a atribuição do diploma no final dos estudos secundários depende da obtenção, num teste de leitura, de uma nota pelo menos igual ao nível médio das turmas do terceiro ano.

V. de Landsheere escreve, precisamente (1988:130): "Fundamentalmente, é impossível fixar um nível único de exigência, porque o grau de perfeição pretendido varia na prática de acordo com o objectivo a atingir e a natureza da aprendizagem a realizar. São raros os casos simples, tais como: sabe-se andar de bicicleta quando, não ocorrendo qualquer acidente, não se tem nenhuma queda e se consegue começar a andar e parar facilmente. Neste caso, o nível de exigência é praticamente de 100%. Como é que se pode aplicar uma regra semelhante a 'saber tocar piano'? Qual a justificação para, nas fichas informativas de avaliação, se exigir uma nota de pelo menos 50%, tanto em língua materna e matemática,... como em ginástica e desenho? Que significa, aliás, saber desenhar parcialmente?"

Em resumo, nenhum dos processos acabados de descrever satisfaz plenamente. Quando muito, conferem um pouco menos de anarquia às decisões.

II. COM VISTA A UMA MAIOR OBJECTIVIDADE

Desde os anos setenta que numerosas investigações se dedicam ao estudo dos métodos científicos de determinação da nota de ruptura e à análise da taxa de erros cometidos aquando da sua aplicação. Em 1986, Berk chama a atenção para mais de trinta e cinco destes métodos. Encontramos também na obra de V. de Landsheere (1988) sobre a competência mínima, uma apresentação e uma análise crítica dos quinze métodos que mais despertaram a atenção. Há um fenómeno que deve ser objecto de profunda reflexão por parte dos educadores: conforme o método escolhido assim a taxa de insucesso varia de 9% a 78%!

Descreve-se aqui apenas o método de Angoff (1971), por reunir um conjunto de qualidades amplamente reconhecidas e por parecer o mais sensato.

1. Pede-se a um grupo de educadores com experiência que indiquem, para cada questão de um exame ou de um teste, que hipóteses (em %) tem um aluno minimamente competente de responder correctamente no grupo a examinar.

10 – 20 – 30 – 40 – 50 – 60 – 70 – 80 – 90 – 100%

2. Cada juiz regista as suas apreciações no quadro seguinte:

Juiz nº 1		
Em vossa opinião, entre cem alunos minimamente competentes, quantos responderiam correctamente às questões?		
Questões	Percentagem	Percentagem / 100
1	80%	0,8
2	40%	0,4
3	50%	0,5
4	0%	0
5	20%	0,2
Nota mínima de sucesso:		1,9

Para obter a nota de ruptura do conjunto de juízes, calcula-se a média das notas mínimas de sucesso por eles atribuídas.

Neste processo a subjectividade está patente, o que convida a uma permanente ponderação das decisões.

Saunders (1984) propõe um interessante aperfeiçoamento deste método.

Uma vez feitos os cálculos segundo o método de Angoff, os juízes revêem as suas apreciações em três ocasiões:

1. Depois de terem conhecimento das notas de ruptura atribuídas pelos outros juízes.
2. Depois de terem, também, tomado conhecimento da distribuição dos resultados obtidos pelos alunos submetidos à teste.
3. Depois de terem estudado as estatísticas descritivas (média, mediana, amplitude de variação, desvio padrão), elaboradas a partir das notas de

ruptura fixadas no estádio anterior (principalmente a percentagem de alunos que iriam ter sucesso).

A nota de sucesso é finalmente fixada, com base nas posições tomadas na fase 3, por ocasião de uma discussão entre todos os juízes.

CONCLUSÃO

Muito embora representem um inegável progresso, em relação à exigência tradicional de um mínimo de 50% ou de 60% em cada disciplina, os processos acabados de descrever continuam a ser muito pouco seguros.

Há uma preocupação principal que nunca se deve esquecer: a função primordial da avaliação não é classificar os alunos mas antes prestar informações sobre o que eles aprenderam efectivamente e sobre o que lhes falta fazer para serem, pelo menos, minimamente competentes.

Esta noção de “competência mínima” é em última análise tão inacessível como a de “nota verdadeira” sobretudo porque varia em função dos contextos. Ser minimamente competente para pilotar um pequeno avião de turismo não qualifica ninguém para tomar conta do comando de um grande avião das linhas internacionais... Além disso, embora não seja inútil, um atestado de incompetência não tem interesse real para o educador se não vier acompanhado de um diagnóstico claro que abra caminho à remediação.

6.3. Análise da percentagem de sucesso

Numa operação deste tipo começa-se por calcular a percentagem global de respostas correctas e o erro padrão para cada questão e cada grupo de questões.

Calcula-se em seguida para cada população testada e para cada questão a percentagem ponderada de respostas correctas (a maior parte das vezes são excluídos dos cálculos os resultados dos alunos que não chegaram a terminar o teste).

Calcula-se o **erro padrão de medida** para cada percentagem de respostas correctas

Este erro é devido ao facto de nunca se medir de forma perfeita, seja qual for a qualidade dos instrumentos. O que é particularmente evidente nas ciências humanas. Por exemplo, quando no final de um exame se atribui uma nota ao conhecimento de uma língua estrangeira, ou ao grau de mestria em piano, é

evidente que não estamos a atribuir **a nota verdadeira**, a que corresponderia, exactamente a tudo o que o aluno sabe e pode fazer. Há técnicas estatísticas que permitem determinar, com maior ou menor probabilidade, um espaço correspondente à *nota verdadeira*.

Estas técnicas partem do princípio de que os erros de medida se distribuem ao acaso à volta da nota verdadeira. Se esta distribuição for *normal* (curva de Gauss), a estimativa é que, mais ou menos duas vezes o erro padrão do valor observado, determine uma zona no interior da qual, em cem, há noventa e cinco possibilidades de se situar a *nota verdadeira*.

Entre as técnicas sofisticadas utilizadas para a estimativa deste erro cada vez se recorre mais à técnica complexa de *jackknife*.

A técnica de jackknife, introduzida por Quenouille (1956) e retomada por Tukey (1958), tem como objectivo suprimir ou, pelo menos, reduzir a influência de quem faz a estimativa. Consiste em construir a partir de uma amostra de efectivos n , n subamostras de efectivos ($n - 1$), obtidas, cada uma delas, pela supressão de uma das observações de cada vez, e em calcular, para cada uma delas, um pseudovalor. O estimador de Quenouille-Tukey é a média destes pseudovalores.

Esta técnica é útil, sobretudo quando se trata de estatísticas de relações entre variáveis (por exemplo, o coeficiente de correlação) cujo valor é difícil de calcular. O termo jackknife faz alusão a uma faca suíça, com variados préstimos, comummente utilizada para resolver os pequenos problemas do dia a dia.

Nos exemplos dados de trabalhos da NAEP, o erro padrão de medida é calculado segundo esta técnica.

Em geral, pretende-se também conhecer as percentagens de sucesso e os erros padrão por população, em relação a cada grupo de *itens*, que correspondam a aspectos particulares da disciplina (aspectos definidos por peritos das disciplinas).

6.4. Probabilidade de resposta correcta

O modelo de Rasch exprime a probabilidade de resposta correcta a uma questão como uma função da dificuldade da questão e da aptidão do sujeito.

A complexa teoria de G. Rasch (1960) parte da seguinte interrogação: "Em que condições se poderão comparar os resultados de um mesmo indivíduo observados em diferentes testes, ou então, comparar os resultados de diferentes grupos de estudantes num mesmo teste?". A questão coloca-se também por exemplo

quando um teste é traduzido em diferentes línguas, com vista a um estudo internacional (G.N. Masters, 1993).

G. Rasch contesta que um teste bem construído administrado a populações diversas nas mesmas condições permita, necessariamente, obter resultados comparáveis. A comparação é de duvidar se o teste não representar uma tarefa igual para todos. É o que acontece se, por exemplo, nem todos tiverem seguido o mesmo programa de ensino.

G.N. Masters (1993) sintetiza a problemática nos termos seguintes:

“Para que um teste possa ser considerado como verdadeiro instrumento de medida é preciso que as questões:

- tratem do mesmo aspecto de determinada competência;
- permitam construir uma escala de medida com uma unidade bem definida;
- possam ser “calibradas” com base nesta escala de modo que se possam comparar, directamente, diferentes selecções de questões;
- constituam uma tarefa comparável para os diferentes alunos que tiverem de lhes dar resposta.”

Segundo o modelo proposto por Rasch, a probabilidade de um estudante responder correctamente a uma questão i , depende do grau de mestria que o aluno tiver da matéria e da dificuldade intrínseca da questão. A fórmula de Rasch permite calcular se um estudante n , que responda a uma questão do tipo *verdadeiro ou falso*, tem uma probabilidade de 1 ou de 0 de responder correctamente.

A possibilidade de comparar, directamente, resultados conseguidos em diferentes testes é particularmente interessante para as medidas de rendimento destinadas à pilotagem. Por exemplo, já não é necessário que todos os alunos que frequentam diferentes estabelecimentos respondam às mesmas questões, quanto mais não seja pelo facto de terem seguido programas diferentes.

O preço a pagar por esta comodidade é o exercício de um controlo permanente da conformidade das respostas dos alunos com os modelos.

6.4.1. Algumas particularidades sobre a teoria dos traços latentes

O modelo de construção de escalas de rendimento utilizado por Rasch supõe que determinado rendimento se baseia num traço latente comum à natureza da tarefa a efectuar e a uma modalidade do conhecimento ou da competência do indivíduo que a efectua. As tarefas ou itens utilizados fazem uma estimativa do rendimento dos indivíduos que dependam deste traço, de modo a situá-lo num continuum.

Este processo não exige que se disponha, necessariamente, de um universo hipotético de itens, donde se possam retirar amostras. Implica, mais precisamente, que a resposta do aluno a um determinado item, é função do erro aleatório e da posição ocupada pelo aluno no continuum do traço latente.

Num conjunto de itens não deve haver intercorrelação, nem entre os erros aleatórios com que cada item é afectado, nem entre o erro e o traço latente. A assunção deste facto implica que o traço seja unidimensional. Se os itens escolhidos estiverem bem correlacionados com um objectivo específico de aprendizagem ou com uma aptidão particular, um modelo como o de Rasch funcionará bem.

I.I. Bejar (citado por J. Keeves, 1992:116) sublinha, com razão, que a unidimensionalidade não implica que o rendimento obtido num item se deva apenas a um único processo psicológico: "Realmente, na elaboração das respostas a um conjunto de itens, intervêm diversos processos psicológicos. Mas, desde que os processos funcionem em uníssono, isto é, se a resposta dada a um item estiver relacionada, do mesmo modo, a um processo idêntico, a hipótese da unidimensionalidade continua de pé. Não é respeitada se, por exemplo, certos itens de um teste de conhecimentos exigirem cálculos, e outros exigirem apenas a simples reconstituição de factos memorizáveis. O teste deixa de ser unidimensional se na população testada, o nível de aptidão numérica variar".

O modelo de Rasch utiliza a lógica matemática (função logística), para estabelecer a relação entre a posição dos itens na escala do traço latente e a probabilidade de resposta correcta ao item.

J. Izard (1992:27) chama a atenção para a multiplicidade de casos de utilização do modelo:

- Aferição de testes de conhecimentos.
- Construção de bancos de itens.
- Organização de grupos de testes com propriedades idênticas.
- Distribuição dos alunos por turmas, muito embora não tenham respondido às mesmas questões.
- Detectar questões particularmente sensíveis à forma de ensino praticado.
- Construção de testes que coloquem questões em função das respostas dadas aos itens anteriores (os chamados testes adaptáveis ou por medida).
- Identificação de alunos que apresentem um perfil pouco habitual de competências.
- Identificação das diversas estratégias utilizadas para a resolução de problemas (esta diversidade pode andar ligada a aspectos culturais).
- Assinalar as semelhanças entre itens, o que ajuda a descrever a competência manifestada pelo aluno em determinado nível.

- Análise das respostas às questões abertas, a fim de verificar se as notas a que deram origem são coerentes com os resultados dos questionários de escolha múltipla.
- Pôr os testes em paralelo uns com os outros.

O modelo de Rasch constitui um dos melhores exemplos da teoria das respostas aos itens (IRT) de que se falará a propósito da fixação das escalas.

É este modelo que está na base das escalas utilizadas pela NAEP. São muito explícitas, mesmo para um público pouco informado; por outro lado, permitem comparar facilmente os rendimentos de diferentes populações numa mesma altura, ou observar como uma mesma população evolui à medida que o tempo passa.

6.5. Fixação de escalas

Sendo os indicadores destinados, também, a informar um vasto público acerca das aquisições dos alunos e a fomentar o debate sobre este tema, é importante que os resultados sejam apresentados de forma a serem facilmente compreendidos e que se indique claramente quais os níveis de rendimento que podem ser considerados satisfatórios. Uma das vantagens da *teoria das respostas aos itens* (Item Response Theory - IRT) é permitir uma interpretação criterial (isto é, determinar em relação a um objectivo definido qual o grau de mestria a atingir).

“A teoria das respostas aos itens é uma das formas de analisar os itens de uma amostra matricial. Em geral, parte-se da hipótese que um único parâmetro p caracteriza a aptidão de um indivíduo, para responder aos itens que fazem parte de uma amostra matricial; supõe-se, além disso, que cada item possui um ou vários parâmetros que descrevem a relação entre os p de todos os indivíduos e as respostas aos itens sobre as quais incidiu a observação. Admite-se, deste modo, a hipótese de que a probabilidade que um indivíduo tem de responder correctamente a um item é função de certas propriedades deste mesmo item, da aptidão do indivíduo e da relação entre a aptidão e o item. De uma maneira geral pensa-se que esta relação de acumulação não é linear. Elaborado um modelo da relação de acumulação, um programa de computador faz uma estimativa dos parâmetros individuais p e do ou dos parâmetros associados a cada item” (A.E. Beaton).

No sistema adoptado pela NAEP dos Estados Unidos os resultados podem distribuir-se entre 1 e 1000. Admite-se, contudo, a hipótese de que 99% dos resultados se situarão entre 200 e 500. Os cinco níveis escolhidos para apresentar os resultados em todos os relatórios publicados são 300, 400, 500, 600 e 700 (escala de cinco níveis). Vimos, no capítulo III, exemplos simplificados da aplicação deste tipo de aferição.

Cada nível é descrito pelas tarefas que a maioria dos alunos que o atingem são capazes de efectuar correctamente; como exemplo, são apresentadas algumas questões-tipo.

Os resultados publicados indicam, para cada nível e para cada população, a percentagem de alunos que se limitaram a atingi-lo ou que o ultrapassaram.

Para fixar as escalas assinalam-se, por ocasião dos estudos prévios, as questões com boa capacidade de discriminação entre os diferentes níveis de complexidade. Uma questão é tomada como referência em determinado nível, se os alunos que a atingem têm entre 65 e 80 oportunidades em 100 de responder correctamente. Por exemplo, a probabilidade dos alunos de um nível imediatamente inferior responderem correctamente situa-se entre 20 e 50. As questões só são finalmente seleccionadas após o acordo dos peritos (tiveram de indicar quais as competências consideradas como pré-requisitos e que competências deviam ser transmitidas aos alunos para que estes pudessem responder correctamente).

6.5.1. Exemplos de escalas de competências

Os responsáveis pela *Avaliação Nacional dos Progressos da Educação* dos Estados Unidos (NAEP) adoptaram, em especial, as duas escalas seguintes - sucintamente referidas no capítulo III - para as populações de 9, 13 e 17 anos (Nation's Report Cards, 1988). Nas publicações da NAEP cada um dos escalões é ilustrado com exemplos de questões de exame.

— Ciências

Nível 1: Conhecimento de factos científicos da vida quotidiana

Os alunos adquiriram, na vida quotidiana, alguns conhecimentos científicos rudimentares sobre animais e plantas. São capazes de descobrir o modo de utilização de aparelhos com os quais estão familiarizados e que funcionem segundo princípios simples de mecânica.

Nível 2: Compreensão de princípios científicos simples

Os conhecimentos relacionados com as ciências da vida e, especialmente, sobre o sistema biológico humano, aumentaram.

Capacidade para aplicar alguns princípios simples da física.

Os alunos começam a compreender o raciocínio científico, especialmente no que respeita a classificação e interpretação.

Nível 3: Aplicação de informações científicas básicas

Interpretação de dados constantes de quadros simples e de gráficos; retirar conclusões a partir dos resultados de experiências científicas. Os alunos dão provas de conhecimento e compreensão no domínio das ciências da vida e, em especial, revelam familiaridade em relação a certos comportamentos animais e a determinadas relações ecológicas. Conhecimento de certas informações básicas de ciências físicas e de ciências da terra e do espaço.

Nível 4: Análise de modos de proceder e de dados científicos

Conhecimentos científicos aprofundados e nítido aumento da capacidade de interpretar textos e gráficos em função dos conhecimentos adquiridos. Progresso na compreensão dos princípios que regem as ciências físicas.

Os alunos são agora capazes de conceber experiências e de interpretar os seus resultados. Sabem distinguir relações e retirar conclusões, aplicando o conhecimento de factos e de princípios, principalmente no domínio das ciências físicas.

Nível 5: Capacidade para integrar informações científicas especializadas

Capacidade para detectar relações e retirar conclusões por inferência, utilizando conhecimentos científicos pormenorizados no domínio das ciências físicas e, em especial, da química. O aluno tem consciência dos limites das extrapolações.

Capacidade para aplicar certos princípios de genética e interpretar as implicações na sociedade, das investigações nesta área.

— **Matemática**

Nível I: Conhecimento factual de simples aspectos da aritmética

Neste nível, os alunos conhecem algumas noções básicas relativas à adição e à subtração e, a maior parte deles, são capazes de adicionar números de dois algarismos desde que isso não implique a execução de transportes. Reconhecem situações simples que façam apelo a adições ou subtrações. Possuem uma capacidade de ordenar rudimentar.

Nível II: Aparecimento da competência e da compreensão

Os alunos revelam um leque mais alargado de competências e uma compreensão mais profunda do que no nível I. Contudo, a utilização que fazem destas competências continua a ser imperfeita e pouco flexível. É-se levado a crer que, neste nível, os alunos têm dificuldade em fazer raciocínios que exijam algo mais do que simples cálculos. Apenas sabem resolver problemas simples que não exijam mais do que uma única operação. Compreendem o conceito de percentagem.

Nível III: Domínio das operações básicas e início da resolução de problemas

Os alunos possuem agora uma compreensão superficial das quatro operações e começam a adquirir competência para resolver problemas. Adquiriram a noção de número par e ímpar e uma compreensão elementar da noção de quociente.

Os alunos são capazes de utilizar números decimais, de converter fracções e percentagens e de calcular médias.

Resolução de problemas em várias fases. Capacidade para medir comprimentos, identificar formas geométricas, calcular a área do rectângulo. Utilização de informações a partir de gráficos ou quadros.

Nível IV: Processos e raciocínios de complexidade moderada

O raciocínio numérico torna-se mais sofisticado. O aluno recorre a um leque mais alargado de competências matemáticas, incluindo a álgebra e a geometria.

Multiplicação de fracções e de números decimais. Utilização de uma gama de processos que possibilitam a resolução de problemas bastante complexos.

Maior compreensão da medida e dos conceitos geométricos: medidas de ângulos nas figuras simples, cálculo do perímetro, da superfície e do volume.

Nível V: Resolução de problemas em várias fases; álgebra

Conhecimento das propriedades da média aritmética. Resolução de problemas utilizando dados que fazem parte de quadros complexos.

O aluno tem cada vez mais capacidade de utilizar noções matemáticas aprendidas na escola para resolver problemas da vida corrente.

6.5.2. Crítica da teoria das respostas aos itens (Goldstein e Wood, 1989)

A crítica principal diz respeito à dificuldade em estabelecer a unidimensionalidade das escalas.

Para estabelecer uma escala de respostas aos *itens*, parte-se do suposto que eles se referem todos, apenas a um único aspecto ou *traço* (por exemplo, competência na leitura) e que o aluno se caracteriza por um único valor deste traço. Se assim fosse, a escala seria *unidimensional*. Resta é provar que de facto assim é, o que está longe de ser tarefa fácil.

Em relação a cada um dos *itens* que caracterizam cada nível da escala, admite-se a hipótese que qualquer melhoria de resultados, à medida que o aluno avança nos estudos, é reflexo de uma mudança nesse mesmo aluno. Se os resultados obtidos num *item* não sobem à medida que se avança na escolaridade, o *item* é considerado anormal e eliminado. Pelo contrário, se os resultados melhoram de um ano para o outro, o *item* é conservado. Resta provar se este progresso é devido a uma melhoria efectiva do conhecimento e não, por exemplo, a uma maior familiarização com a prova aplicada.

Além disso, não é certo que o “*traço*” que se julga ter detectado em determinada população seja, necessariamente, igual ao de uma outra pertencente a uma cultura diferente ou frequentando um sistema de educação diverso. A tradução para várias línguas basta para criar diferenças de dificuldade nas questões. E se por esta razão, nas avaliações comparativas internacionais, se for levado a suprimir grupos de *itens* por diferirem de país para país, então a comparação pode tornar-se pouco digna de confiança.

Depois, o problema do valor acrescentado também se coloca neste caso. Para calcular em que medida o ensino contribuiu para o progresso das aprendizagens, há que medir o estado inicial das aquisições dos alunos. Deste ponto de vista, a perspectiva longitudinal é, evidentemente, a mais desejável.

6.6. Funcionamento diferencial dos *itens*

Considera-se que uma questão funciona diferencialmente se alunos com aptidões iguais mas pertencendo a diferentes populações tiverem uma diferente probabilidade de responder correctamente. Esta diferença pode por exemplo derivar de problemas linguísticos. São utilizadas técnicas estatísticas complexas para determinar este funcionamento diferencial.

Conclusão

As considerações anteriores estão longe de cobrir o conjunto dos problemas complexos colocados pelas avaliações destinadas a dar corpo aos indicadores.

Pensamos que deixam entrever o grau de tecnicidade de muitos processos. Estes derivam muitas vezes da elevada especialização, o que explica que, em todos os países onde os sistemas de indicadores da educação foram introduzidos, os centros de investigação universitária desempenhem um importante papel.

PRINCIPAIS ABREVIATURAS

- APU Assessment of Performance Unit (Unidade de Avaliação do Rendimento Escolar) (Reino Unido)
- CERI Centre for Educational Research and Innovation (Centro de Investigação e Inovação no Ensino) (OECD)
- CITE Classificação Internacional Estandardizada do Ensino
- DEP Direcção de Avaliação e Prospectiva, Ministério da Educação (França)
- ETS Educational Testing Service (Princeton)
- IAEP International Assessment of Educational Progress (Avaliação Internacional do Progresso da Educação) (Estados Unidos)
- IEA International Association for the Evaluation of School Achievement (Associação Internacional de Avaliação do Rendimento Escolar)
- INES Indicadores Internacionais do Ensino (Projecto da OCDE)
- IRT Item response theory
- k Factor “Conhecimento” (*knowledge*)
- ISCED International Standard Classification of Education - Francês: CITE
- ISCO International Standard Classification of Occupations
- NAEP National Assessment of Educational Progress (Avaliação Nacional dos Progressos da Educação) (Estados Unidos)
- NOBS Non-curriculum bounded objectives (Objectivos Independentes do Currículo)
- OCDE Organização de Cooperação e Desenvolvimento Económico
- OTL Opportunity to learn. Ocasão de aprender: índice que indica em que medida os alunos tiveram possibilidade de aprender os saberes, as competências e o saber-ser sobre que incidem as avaliações.
- PIB Produto Interno Bruto
- SEAC School Examinations and Assessment Council (Conselho de Avaliação Escolar) (Reino Unido)
- SSE School Self-Evaluation (Auto-Avaliação das Escolas).

BIBLIOGRAFIA

- ALLISON, G., *The Essence of Decision*, Boston, Little, Brown and Co, 1971.
- APPLEBEE, A.N., LANGER, J.A. e MULLIS, V.S., *The writing report card: Writing achievement in American schools*, Princeton, ETS (NAEP), 1986.
- APPLEBEE, A.N. et al., *Who reads best? Factors related to reading achievement in Grades 3, 7 and 11*, Princeton, ETS (NAEP), 1988.
- ASSOCIATION FRANÇAISE DES ADMINISTRATEURS DE L'EDUCATION, *Le Pilotage du système éducatif, Acte du XVe colloque, Revue de l'AFAE*, 1993, 3.
- ATKINSON, R.C. e JACKSON, G.B., Eds, *Research and Education Reform*, Washington, D.C., National Academy Press, 1992.
- AUBRET, J., *Reconnaissance et validation des acquis: le portefeuille de compétences*, Paris, Université de Paris VIII, 1992.
- BAUER, R., Ed., *Social Indicators*, Cambridge, Mass., MIT Press, 1966.
- BEATON, A.E., *Implementing the new design: The NAEP 1983-84 Technical Report*, Princeton, ETS, 1987, 786 p.
- BEJAR, I.I., *Achievement testing - Recent advances*, Beverly Hills, Calif., Sage Publications, 1983.
- BERK, R.A., A consumer's guide to setting performance standards on criterion-referenced tests, *Review of Educational Research*, 1986, 56, 1, 137-172.
- BERLEUR, J., *Technology Assessment*, Namur, Faculté Universitaire N.- D. de la Paix, 1991 (syllabus de séminaire).
- BIJKER, W.E. e LAW, J., *Shaping technology / Building society*, Cambridge, Mass., MIT Press, 1992.
- BOTTANI, N. e WALBERG, H., Eds, *The OECD International Education Indicators. A framework for analysis*, Paris, OCDE, 1992. (Edition française: L'OCDE et les indicateurs internationaux de l'enseignement. Un cadre d'analyse).
- BURSTEIN, L., Ed., *The IEA Study of Mathematics III*, Oxford, Pergamon, 1993.
- CARR-HILL, R. e MAGNUSEN, O., *Les indicateurs de résultats des systèmes d'enseignement*, Paris, OCDE, 1973.
- CARROLL, J.B., *The teaching of French as foreign language in eight countries*, Stockholm, Almquist and Wiksel, 1975.
- CLIFT, P.S., NUTTALL, D.L. e McCORMICK, R., Eds, *Studies in school self-evaluation*, Lewes, Palmer Press, 1987.

- COLLIN, R.D., *Les technologies de l'intelligence*, Genève, Neurope Lab, 1993.
- COLLINGRIDGE, D., *The social control of technology*, Londres, Francis Pinter, 1980.
- COLOMB, J., *L'évaluation des élèves et l'examen dans l'enseignement secondaire*, Rapport, Strasbourg, Conseil de l'Europe, DECS/Rech (1993) 22.
- COMBER, L.C. e KEEVES, J.P., *Science education in nineteen countries*, Stockholm, Almquist and Wiksel, 1973.
- COMMISSION DES COMMUNAUTES EUROPEENNES, *The social and Economic Implications of New Technology*, First Biennial Report, Society and Technology towards the Future, FAST-DGXII, 1991, 2, 101-117 (FAST Document 10).
- CONSEIL DES MINISTRES DE L'EDUCATION DU CANADA, *Mise en application des indicateurs de l'enseignement: L'expérience du Canada*, Communication à l'assemblée générale du project INES de l'OCDE, Semmering, 1989.
- DE GROOF, J. et al., Indicatoren voor het onderwijsbeleid. Een stand van zaken van het OESO-Project "INES", *Tijdschrift voor het onderwijsbeleid*, 1991-1992, 3, 142-156.
- DE GROOF, J. et al., *De School op rapport. Het Vlaams onderwijs in internationale context*, Kapellen, Peleckmans, 1993.
- de LANDSHEERE, G., Comparaison du rendement de l'enseignement des mathématiques dans douze pays, *Education* (Bruxelles), 1967, 106, 61-70.
- de LANDSHEERE, G., *Evaluation continue et examens. Précis de docimologie*, Paris, Nathan; Bruxelles, Labor, 1992, 7e éd.
- de LANDSHEERE, G., *Dictionnaire de l'évaluation et de la recherche en éducation*, Paris, PUF, 1992, 2e éd.
- de LANDSHEERE, G., e POSTLETHWAITE, T.N., *Rendement de l'enseignement des mathématiques dans douze pays*, Paris, Institut Pédagogique National, 1969.
- de LANDSHEERE, V., *Faire réussir – Faire échouer. La compétence minimale et son évaluation*, Paris, PUF, 1988.
- de LANDSHEERE, V., *L'éducation et la formation. Science et pratique*, Paris, PUF, 1992.
- DELORS, J., *Les indicateurs sociaux*, Paris, Futuribles, 1971.
- DEPARTMENT OF EDUCATION OF NEW ZEALAND, *Assessment for better learning*, Wellington, 1989.
- DIETEL, R., What works in performance assessment? *Evaluation comment*, Printemps, 1993.
- DORON, R. e PAROT, F., Eds, *Dictionnaire de psychologie*, Paris, PUF, 1991.

- DOSSEY, J., et al., *Mathematics: Are we measuring up?*, The mathematics report card, Executive summary, Princeton, ETS (NAEP), 1988.
- DUPONT, P., DUPUIS, M. e OSSANDON, M., *A livre ouvert... dans la gestion de l'école*, Bruxelles, Ministère de l'Éducation, 1988.
- EBEL, R.L., *Essentials of educational measurement*, Englewood Cliffs, Prentice Hall, 1979, 3e éd.
- ELLY, W.B., *How in the world do students read?* La Haye, The IEA Secretariat, 1992.
- FITZ-GIBBON, C., *Performance Indicators and Examination Results, Interchange*, 1992, 11 (Edinburgh).
- FLANAGAN, J.C., et al., *The American High-School Student: The Identification, Development and Utilization of Human Talents*, Pittsburg, Pittsburg University Press, 1964.
- FLANAGAN, J.C., et al., *Project Talent: Five Years after High School and Appendix II*, Pittsburg, American Institute for Research, 1971.
- GARDEN, R.A. e ROBITAILLE, D.F., Eds, *The IEA study of mathematics II*, Oxford, Pergamon, 1989.
- GARIN, C., Les régions dans la compétition scolaire, *Le Monde*, 14 janvier 1993.
- GIPPS, C., *National Assessment and School Evaluation in the U.K.*, Communication à la Conférence annuelle de l'AERA, San Francisco, 1989.
- GLASER, R., Instructional technology and the measurement of learning outcomes. Some questions, *American Psychologist*, 1963, 18, 519-521.
- GOLDSTEIN, H. e WOOD, R., Five decades of item response modelling, *Brit. J. Math. Stat. Psychol.*, 1989, 42, 139-167.
- GONOD, F., Problématique de la maîtrise sociale de la technologie, *Analyse des systèmes*, 1990, 16, 3, 3-43.
- GORMAN, T.P., PURVES, A., DEGENHART, R.E., *The IEA study of written composition, Vol.I: The international writing tasks and scoring scales*, Oxford, Pergamon, 1988.
- GRISAY, A., *Rendement de l'enseignement de la langue maternelle en Belgique francophone*, Bruxelles, MENC, Organisation des Etudes, Coll. "Recherche en éducation", 1974.
- GRISAY, A., L'enseignement belge forme-t-il trop peu de bons élèves?, *Revue de la Direction générale de l'Organisation des Etudes*, 1974, 9, 3-12.
- GRISAY, A., La sélection peut-elle aider à combattre l'ennui scolaire? Analyse de quelques résultats de la recherche de l'IEA sur le rendement scolaire en Belgique francophone, *Scientia Paedagogica Experimentalis*, 1976, 13, 2, 241-261.

- GRISAY, A., *Du mythe de la "bonne école" à la réalité (fuyante) de "l'école efficace"*, Liège, Service de pédagogie expérimentale de l'Université de Liège, 1988.
- GRISAY, A. e HENRY, G., La recherche IEA en sciences et en langue maternelle, *Revue de la Direction générale de l'Organisation des Etudes* (Bruxelles), 1973, 8, 35-39.
- GUTHRIE, J.W., *The evolving political economy of education and the implication for educational evaluation*, (conferência), Leicester University, 1989.
- HABY, R., in't VELD, R.J. e TSCHOUMY, J.A., *Examens des politiques nationales d'éducation – Belgique*, Paris, OCDE, 1993.
- HARGREAVES, D.H., Assessment and Performance Indicators: the English Experience. In T. WYATT e A. RUBY, Eds, 1988, 125-134.
- HENRY, G., Le rendement de l'enseignement des mathématiques dans douze pays, *Information. Bulletin de l'Organisation des Etudes* (Bruxelles), 1970, 1, 3-16.
- HENRY, G., Un aspect de l'inégalité socio-culturelle entre sexes. Etude différentielle des résultats de l'IEA en sciences, *Revue de la Direction générale de l'Organisation des Etudes*, 1974, 3, 3-9.
- HENRY, G., Une étude du rendement de l'enseignement des sciences dans dix-neuf pays, *Actes du congrès des professeurs de sciences* (Bruxelles), 1974, 31, 1-31.
- HENRY, G., *Rendement de l'enseignement des sciences en Belgique francophone*, Bruxelles, MENC, Coll. "Recherche en éducation" 8, 1975.
- HENRY, G., Une étude des effets de la sélection au niveau de l'enseignement secondaire supérieur, *Revue de la Direction générale de l'Organisation des Etudes*, 1975, 1, 3-13.
- HERRIOT, R.E., Some Observations on the Condition of Education, *Social Indicators Research*, 1979, 6, 181-194.
- HUTMACHER, W., *Quand la réalité résiste à la lutte contre l'échec scolaire*, Genève, Service de recherche sociologique, Cahier 36, 1993.
- IEA, *Science achievement in 17 countries*, Oxford, Pergamon, 1988.
- IZARD, J., *Assessing learning achievement*, Paris, UNESCO, Coll. Educational Studies and Documents 60, 1992.
- JOHNSON, E.G., The design of the National Assessment of Educational Progress, *Journal of Educational Measurement*, 1992, 29, 95-100.
- KAAGAN, S. e SMITH, M.S., Indicators of educational quality, *Educational Leadership*, Octobre 1985, 21-24.

- KEEVES, J., Scaling achievement test scores. In J. KEEVES, Ed., *Methodology and measurement in international educational surveys*, La Haye, IEA, 1992.
- KEEVES, J. *Learning science in a changing world. Cross-national studies of science achievement: 1970-1984*, La Haye, IEA, 1992.
- KEEVES, J. e ROSIER, M., *The IEA study of science I*, Oxford, Pergamon, 1993.
- KOGAN, M., Monitoring, Control and Governance of School Systems, in P. LADERRIERE, Ed., *Governance and Accountability in Education*, Paris, OECD, 1993.
- KONTTINEN, R., *Research into the role of final examinations in secondary education*, Strasbourg, Conseil de l'Europe, DECS / Rech. 1993.
- KUHN, T.S., *The structure of scientific revolution*, Chicago, University Press, 1970, 2e éd.
- LAPOINTE, A.E., MEAD, N.A. e PHILLIPS, G.W., *A World of Differences. An International Assessment of Mathematics and Science*, Princeton, ETS, 1989.
- LECLERCQ, D., Computerized tailored testing, *European Journal of Education*, 1980, 5, 3.
- LE GUEN, M., *Measuring to succeed. National evaluation in France*, Paris, OCDE-CERI/INES, 1991, 9.
- LEWIS, E.G. e MASSAD, C.E., *The teaching of English as a foreign language in ten countries*, Stockholm, Almquist and Wiksel, 1975.
- LIENSOL, B. e MEURET, D., Les performances des lycées pour la préparation au baccalauréat, *Education et Formations*, 1987, 11, avril-juin, 25-36.
- LINN, R.L., BAKER, E.L. e DUBAR, S.B., Complex, performance-based assessment: expectations and validation criteria, *Evaluation Comment*, Hiver 1991-1992, 3-9.
- MADAUS, G.F., The influence of testing on the curriculum. In L. TANNER, Ed., *Critical issues in curriculum*, Chicago, University of Chicago Press, 1988.
- MADAUS, G.F. e KELLAGHAN, T., Curriculum evaluation and assessment. In P. W. JACKSON, Ed., *Handbook of research on curriculum*, New York, Macmillan, 1992, 119-154.
- MAGY, J., *L'enseignement et la formation en Communauté française. Produire et gérer la qualité*, Bruxelles, CEPES, Coll. Monographies, 1992.
- MANTEL, N. e HAENSZEL, W., Statistical aspects of the retrospective study of disease, *Journal of the National Cancer Institute*, 1959, 22, 719-748.
- MARKLUND, S., Assessment of School Achievement in Compulsory and Post-Compulsory Schools in Sweden. In P. WESTON, Ed. *Assessment of pupil achievement: motivation and school success*, Strasbourg, Conseil de l'Europe, 1991, 31-44.

- MASSOZ, D. e HENRY, G., *Connaissances et attitudes socio-politiques d'élèves de l'enseignement secondaire*, Recherche de l'IEA, Bruxelles, MENC, Direction générale l'Organisation des Etudes, 1980.
- MASTERS, G.N., Rasch measurement and IEA studies, *IEA Bulletin*, 1993, 2,2.
- MATALON, B., *L'analyse hiérarchique*, Paris, Gauthier-Villars, 1975.
- McKNIGHT, C.C. et al., *The underachieving curriculum – Assessing U.S. School Mathematics from an International Perspective*, Champaign, Ill., Stupp Publishing, 1987.
- MENESE, J., *L'analyse modulaire des systèmes*, *Hommes et techniques*, 1972.
- MINISTÈRE DE L'EDUCATION NATIONALE, DIRECTION DE L'EVALUATION ET DE LA PROSPECTIVE, *L'état de l'école*, Paris, DEP, 1992, 2.
- MISLEVY, R.J., Foundations of a new test theory. In N. FREDENKSEN e I. BEJAR (Eds.), *Test theory for a new generation of tests*, Lawrence Erlbaum, 1990.
- MULLIS, I.V.S. e JENKINS, L.B., *The Science report card. Trends and achievement based on the 1986 National Assessment*, NAEP, Princeton, 1988.
- MULLIS, I.V.S. et al., *The science report card 1986*, Princeton, ETS-NAEP, 1988.
- NATIONAL CURRICULUM COUNCIL, *National Curriculum Council*, York, 1991.
- NATIONAL EDUCATION GOALS PANEL, *The National Education Goals Report*, Washington, 1992.
- NEAVE, G., *Education and the European Community*, Paris, Institut européen d'éducation et de politique sociale, 1983.
- NGUYEN, N.T., LOBET-MARIS, C., BERLEUR, J. e KUSTERS, B., Methodological issues in information technology assessment, *The International Journal of the Management of Technology*, 1993.
- NUTTALL, D., Choosing indicators, In OCDE-CERI, 1991.
- NUTTALL, D., School evaluation in the UK, in *Université Libre de Bruxelles, L'enseignement en Europe. L'enseignement en Belgique*, Bruxelles, ULB, 1993, 43-47.
- OAKES, J., *Educational indicators: a guide for policy makers*, New Brunswick, N.J., Center for policy research in education, Rutgers University, 1986.
- OCDE, *Regards sur l'éducation. Les indicateurs de l'OCDE*, Paris, OCDE, 1992.
- OCDE, *Une éducation et une formation de qualité pour tous*, Paris, OCDE, 1992.
- OCDE, *The OECD International Education Indicators: A framework for analysis*, Paris, OCDE, 1992.
- OCDE-CERI, *Manuel sur les indicateurs de l'enseignement*, Paris, OCDE, 1991 (Doc. CERI/INES/91) 17.

- OCDE-CERI, *Curriculum reform: Policies, Practices and Plans for the Future*, Paris, OCDE, 1993 (Doc. CERI/ED/93)12.
- OLMSTED, P. e WEIKART, D., *Profiles of child care and education in fourteen countries*, Ypsilanti, High/Scope Press, 1989.
- PALMER, R.F., *National Testing in Canada*, Strasbourg, Conseil de l'Europe, Rech. 1993, 39.
- PELGRUM, H. e PLOMP, T., *The use of computers in education world wide*, Oxford, Pergamon, 1991.
- PETRELLA, R., Des politiques malades de leur culture. L'homme et l'outil: les valeurs qui basculent, *Le Monde diplomatique*, 20 septembre 1987.
- PIJL, S.J., Practices in Monitoring Student Progress, *Internationale Zeitschrift für Erziehungswissenschaft*, 1992, 38, 117-131.
- PORTER, A., Indicators: objective data or political tools?, *Phi, Delta, Kappan*, 1988, 3, 503-508.
- POSTLETHWAITE, T.N., *The Monitoring of Education and the Responsability of Governments*, Bruxelles, International Academy of Education, 1993.
- POSTLETHWAITE, T.N. e WILEY, D., Eds., *The IEA study in science II: Science achievement in 23 countries*, Oxford, Pergamon, 1993.
- PURVES, A., *Literature education in 10 countries*, Stockholm, Almqvist and Wiksel, 1973.
- RASCH, G., *Probabilistic models for some intelligence and attainment tests*, Copenhagen, Danish Institute for Educational Research, 1960.
- RICHARDS, C.E., Indicators and three types of educational monitoring systems. *Phi, Delta, Kappan*, 1988, 3, 495-498.
- ROGERS, G., *Keeping the school under review. A method of self assessment for schools devised by the ILEA inspectorate*, Londres, ILEA, 1977.
- ROSS, K.N. e POSTLETHWAITE, T.N., *Indicators of the quality of education: a summary of a national study of primary schools in Zimbabwe*, Paris, International Institute of Educational Planning (UNESCO), 1993.
- RUBY, A., The International Indicator Project, In T.J. WYATT e A. RUBY, Eds., *Indicators in Education*, Sidney, Australian Conference of Directors-General of Education, 1988.
- RUBY, A., Indicators, reporting and rationality: understanding the phenomena, in OCDE/CERI, *International educational indicators project. A reader on conceptual issues*, Paris, OCDE, 1991 (documento de trabalho).

- ODDEN, A., Making sense of education indicators: the missing ingredient. In T. WYATT e A. RUBY, Eds., 1990, 33-50.
- RUBY, A. e WYATT, T.J., *Using indicators for school improvement*, Sydney, Australian Conference of Directors-General of Education, 1988.
- SCHEERENS, J., *Process Indicators of School Functioning*, Paris, OCDE-CERI, 1989.
- SCHEERENS, J., Caractéristiques des établissements d'enseignement. Programme à moyen terme, 1992-1994. Note de travail reproduite dans le *Rapport annuel* du Ministère de l'éducation, de la recherche et de la formation de la Communauté française de Belgique (Direction générale de l'organisation des études), janvier 1993.
- SEDLACEK, W.E., Noncognitive Indicators of Student Success, *Journal of College Admission*, 1989, 125 2-10.
- SELDEN, R., Missing data, *Phi, Delta, Kappan*, 1988, 3, 487-491.
- SIMON, H.A., *La science des systèmes, science de l'artificiel*, Paris, EPI, 1974.
- SMITH, M.S., Educational Indicators, *Phi, Delta, Kappan*, 1988, 3, 492-494.
- SKINNER, B.F., *Walden Two*, New York, Macmillan, 1968.
- STEVICK, B.F., *Skinner's Walden Two*, New York, Seabury Press, 1968.
- STEWARD, J. et al., *Assessment for better learning*, Wellington, Ministère de l'Education, 1989, p. 19.
- STEWARD, J. et al., *Tomorrow's standards. The report of the ministerial working party on assessment for better learning*, Wellington, Ministère de l'éducation, 1990.
- STIGGINS, R.J., Design and development of performance assessment, *Education Measurement: Issues and Practice*, 1987, 6, 3, 33-42.
- TALMAGE, H., Evaluation of programs, In H.E. MITZEL, Ed., *Encyclopedia of Educational Research*, New York, Free Press, 1982.
- THORNDIKE, R.L., *Reading comprehension education in 15 countries*, Stockholm, Almquist and Wiksel, 1973
- TORNEY, J., OPPENHEIM, A.N. e FARNEN, R.F., *Civic education in ten countries*, Stockholm, Almquist and Wiksel, 1976.
- U.S. DEPARTMENT OF EDUCATION, National center for education statistics (Research and Development Reports), *International Mathematics and Science Assessments: What Have We Learned?*, Washington, 1992.
- VAN HERPEN, M., Conceptual models in use for education indicators, In N. BOTTANI e H. WALBERG, 1992, pp. 25-52.

- VIGOTSKY, L.S., *Mind in society: the development of higher psychological processes*, Cambridge, Mass., Harvard University Press, 1978.
- VINCK, D., L'évaluation sociétale des choix technologiques, *Courrier du CETHES*, 1988, 5, 5-16.
- WALBERG, H., *The implications of cognitive psychology for measuring school achievement*, Chicago, University of Illinois, 1991, pp. 19-22.
- WEDMAN, I., *Pupil assessment with a focus on assessment at secondary education*, Strasbourg, Conseil de l'Europe, DECS/Rech., 1993.
- WYATT, T.J., *What can be learned from the social indicators movement?*, Sydney, Australian Conference of Director-General of Education, 1988.
- WYATT, T., Indicators. Trends in the educational literature, in OECD-CERI, *International Educational Indicators Project. A reader on conceptual issues*, Paris, OCDE, 1991.
- WYATT, T. e RUBY, A., Eds., *Indicators in Education*, Sydney, Australian Conference of Directors-General of Education, 1988.
- WYATT, T. e RUBY, A., Eds., *Education indicators for quality, accountability and better practice*, Sydney, Australian Conference of Directors-General of Education, 1990.
- X, THE WHITE PAPER, Choice and Diversity: A New Framework for Schools, *School Update*, Autumn 1992, 2-3. (Londres, Department of Education).

CIÊNCIAS DA EDUCAÇÃO

OBRAS PUBLICADAS

COLEÇÃO PRÁTICAS PEDAGÓGICAS

FORMAR CRIANÇAS LEITORAIS

Josette Jolbert

O FIO DA MEADA

Manuela Ferreira

LUTAR PARA DAR UM SENTIDO À VIDA

Isabel Alves Costa e Filipa Baganha

GUIA PARA A ORGANIZAÇÃO DA ESCOLA

Franco Ghilardi e Carlo Spallarossa

O FANTOCHE QUE AJUDA A CRESCER

Isabel Alves e Filipa Baganha

ENSINAR A POESIA

Teresa Guedes

A INTEGRAÇÃO ESCOLAR DAS CRIANÇAS DEFICIENTES

Jean Simon

100 JOGOS MUSICAIS

Ger Storms

O PRAZER NA ESCRITA

Pierre Bach

CLUBE DE MATEMÁTICA

António César de Sá e Margarida Leite Faria

AVALIAR É REFLECTIR SOBRE O ENSINO

Carlos Rosales

100 JOGOS DE MOVIMENTO

Huberta Wiertsema

AVALIAR É MEDIR?

Jean Cardinet

COMO CRIAR UMA BOA RELAÇÃO PEDAGÓGICA

M.ª Teresa Gomez Masdevall, Victoria Mir i Costa e M.ª Gracia Serrats i Parets

AVALIAÇÃO FORMATIVA

Roland Abrechit

A COMPREENSÃO NA LEITURA

Jocelyne Giasson

EXPRESSÃO DRAMÁTICA E TEATRO

Jean-Claude Landier e Gisèle Barret

CONSTRUIR UMA FORMAÇÃO

C. Strauven

COMO ENSINAR ATITUDES

Dominique Morissette e Maurice Gingras

COLEÇÃO HORIZONTES DA DIDÁCTICA

COMO PREPARAR UMA AULA DE HISTÓRIA

Clemente Herrero Fabregat e María Herrero Fabregat

DIDÁCTICA DA EDUCAÇÃO INFANTIL

Miguel A. Zabalza

ENSINAR E FAZER APRENDER

Louis Not

DIDÁCTICA DA LÍNGUA ESTRANGEIRA

Ana Isabel Oliveira Andrade e Maria Helena A. B. Araújo e Sá

DIFÍCULDADES DA LEITURA E DA ESCRITA EM ALUNOS DO ENSINO BÁSICO

José A. Silva Rebelo

PROFESSORES E COMPUTADORES

Carlos Afonso

A INTERPRETAÇÃO DE TEXTOS NA AULA DE PORTUGUÊS

Maria de Lourdes Sousa

A ECOLOGIA E A ESCOLA

OCDE

NOVAS ESTRATÉGIAS – NOVOS RECURSOS NO ENSINO DA HISTÓRIA

Ana de Sousa, Aureliana Pato e Conceição Canavilhas

ITINERÁRIOS LITERÁRIOS – VIAJANDO PELA LITERATURA PORTUGUESA

Elvira Manuel Pardinhos Azevedo e Zaida Maria Campos Monteiro Braga

DIDÁCTICA E COMUNICAÇÃO

Georges Guislain

CURRÍCULO E ENSINO

A. Matos Vilar

A EDUCADORA DE INFÂNCIA

Marília Mendonça

DIDÁCTICA DAS CIÊNCIAS

António Almeida e Maria da Conceição Vilela

COLEÇÃO BIBLIOTECA BÁSICA DE EDUCAÇÃO E ENSINO

O ENSINO NA SOCIEDADE MODERNA,
O.C.D.E.

INOVAÇÃO PEDAGÓGICA E FORMAÇÃO DE PROFESSORES,
José Alberto Correia

PEDAGOGIA DOS PROCESSOS DE APRENDIZAGEM,
Antoine de La Garanderie

O ENSINO BÁSICO EM PORTUGAL,
Vários autores portugueses

A CONSTRUÇÃO SOCIAL DA EDUCAÇÃO ESCOLAR,
Vários autores portugueses

CULTURA PESSOAL E ACÇÃO PEDAGÓGICA,
Georges Jean

PSICOLOGIA DA ADOLESCÊNCIA E DA RELAÇÃO EDUCATIVA,
Evaristo Fernandes

CURRÍCULO E DESENVOLVIMENTO CURRICULAR
- PROBLEMAS E PERSPECTIVAS,
Fernando Augusto Machado, M. Fernanda Gonçalves

DESENVOLVIMENTO PSICOLÓGICO DA CRIANÇA,
R. Deldime, S. Vermeren

O IMAGINÁRIO NA RELAÇÃO PEDAGÓGICA,
Marcel Postic

DESENVOLVIMENTO PESSOAL E SOCIAL,
Maria Praia

A REFORMA DOS PROGRAMAS ESCOLARES,
Malcolm Skilbeck / O.C.D.E.

AS ESCOLAS E A QUALIDADE
O.C.D.E.

COLEÇÃO EM FOCO

QUESTÕES DE POLÍTICA EDUCATIVA
Bártolo Paiva Campos

EFICÁCIA NO ESTUDO
Rafael Ibañez Benet, Benito Lopez Andrade,
José Martinez Alvarez e Francisco Bellon

PORTUGAL E A EUROPA – Identidade e Diversidade
Vários autores portugueses

A FORMAÇÃO DO JOVEM – Um Modelo Interactivo
João Rebello de Carvalho, Luís Marques Barbosa, Fernanda Costa Geraldes

EDUCAÇÃO DE ADULTOS – Educação para o Desenvolvimento
Augusto Santos Silva

A EDUCAÇÃO TECNOLÓGICA NOS ANOS 90
Joaquim Azevedo

CRIANÇAS SOBREDOTADAS – Que Sucesso Escolar?
Ilídio de Jesus Coelho Falcão

SUPERVISÃO – Uma Prática Reflexiva de Formação de Professores
Flávia Vleira

INOVAÇÃO E MUDANÇA NA REFORMA EDUCATIVA
Alcino Matos Vilar

A EDUCAÇÃO TECNOLÓGICA E OS NOVOS PROGRAMAS
João Manuel Pereira Dias Baptista

ANALFABETISMO FUNCIONAL E RENTABILIDADE ECONÓMICA
O.C.D.E.

O NOVO MODELO DE GESTÃO DAS ESCOLAS BÁSICAS E SECUNDÁRIAS
Maria Rosa Pereira Tripa

PROJECTO EDUCATIVO DE ESCOLA
Abel Paiva da Rocha

COLEÇÃO CADERNOS DO CORREIO PEDAGÓGICO

O JORNAL ESCOLAR – Porquê e Como Fazê-lo
António Santos e Manuel Pinto

CONTRIBUTOS PARA UMA OUTRA PRÁTICA EDUCATIVA
Ferreira Alves e João Formosinho

ORGANIZAÇÃO, GESTÃO E PROJECTO EDUCATIVO DAS ESCOLAS
José Matias Alves

EXPECTATIVAS ESCOLARES E PROFISSIONAIS DOS JOVENS DO 9º ANO
Joaquim Azevedo

ENSINO ARTÍSTICO
Vários Autores

A ESCOLA E AS AUTONOMIAS
Manuel Jacinto Sarmento

OLHAR E MELHORAR A ESCOLA
Irene Cortesão e Reinalda Malafaia

DESENVOLVIMENTO PESSOAL E SOCIAL
Maria Praia e Maria Fernanda Soares

AJUDE O SEU FILHO A APRENDER
Christian Thomas e Bernard Viselthier

AVALIAR A AVALIAÇÃO
Vários Autores

SER PROFESSOR/A NUM CONTEXTO DE REFORMA
Carlinda Leite e Manuela Terraseca

AJUDE O SEU FILHO A TER SUCESSO
Françoise Brissard

ANIMAÇÃO COMUNITÁRIA
Vários Autores

O PROFESSOR PLANIFICADOR
A. de Matos Vilar

A AVALIAÇÃO FORMATIVA – Que Desafios?
Luiza Cortesão

A AVALIAÇÃO DOS ALUNOS NO ENSINO BÁSICO
A. de Matos Vilar

CRÓNICAS DO QUOTIDIANO ESCOLAR
José Matias Alves

COMUNICAÇÃO NA SALA DE AULA
Jorge da Silva Antão

DO ACESSO AO (IN)SUCESSO – A Questão das Desigualdades
Jorge Silva Rolla

POR UM PROJECTO EDUCATIVO DE REDE
Fernando Diogo

O PROJECTO EDUCATIVO EM EDUCAÇÃO DE INFÂNCIA
Anabela Rolla e Jorge Silva Rolla

O PROCESSAMENTO DE TEXTO E O ENSINO DE LÍNGUAS ESTRANGEIRAS
Anabela Sarmento

AVALIAÇÃO EM TEMPO DE MUDANÇA – Projectos e Práticas nos Ensinos Básico e Secundário
Fernando Augusto Machado

GUIA DE RECURSOS — A Escola e o Meio
Luís Filipe Santos e José Carlos Cruz

AVALIAÇÃO SUMATIVA EXTRAORDINÁRIA
Américo Dias

O PROCESSAMENTO DE TEXTO – Uma Ferramenta para o Ensino da Língua Materna
Zita Maria Romero Gonçalves

VÍDEO EDUCATIVO – Da Produção à Utilização
Armando Silva

INFLUÊNCIA DO PROFESSOR SOBRE OS ALUNOS
Saul Neves de Jesus

A OBRA

Pilotar um avião é conduzir os passageiros aos seus destinos, nas melhores condições possíveis, utilizando da forma mais adequada o material e as informações fornecidas pelos instrumentos de bordo; há sinais que são activados em caso de perigo.

Fazer a pilotagem de um sistema educativo, de uma escola em particular ou até de uma turma, procede do mesmo cuidado de eficácia e segurança, e não é incompatível com uma opção humanista. Assim o entende um número cada vez maior de países, mergulhados na crise cultural contemporânea, da qual a nova revolução industrial é apenas um dos aspectos.

A introdução da pilotagem surge como uma interpelação ao mundo da educação, porque levanta questões de fundo e coloca problemas técnicos complexos. Após uma reflexão de carácter geral sobre a natureza desta tecnologia, e sobre as suas principais orientações, o autor explica o seu funcionamento indicando, em especial, o que são indicadores e o modo de os utilizar. São descritas, de forma rigorosa, as principais realizações nacionais e internacionais em matéria de macropilotagem e de micropilotagem, e são analisadas as suas vantagens e inconvenientes.

A obra termina com um anexo sobre aspectos técnicos particularmente importantes.

O AUTOR

Gilbert de Landsheere é professor emérito da Universidade de Liège onde dirigiu o Laboratório de Pedagogia Experimental. Especialista na elaboração de currículos de ensino e avaliação nos seus diversos aspectos. Autor ou co-autor de dezassete livros, traduzidos em nove línguas e de duzentos e cinquenta artigos. Exerce funções em várias organizações internacionais governamentais, em especial no Gabinete Internacional de Educação (Unesco) em Genebra, e preside à Academia Mundial de Educação.

ISBN 972-41-1887-8



9 789724 118871
Nr de Código 661323