

Conferência

Avaliação do trabalho experimental dos alunos no ensino das ciências

Joan Solomon

Open University, Faculty of Science

1. Por que razão queremos avaliar o trabalho dos alunos?

Algumas pessoas respondem a esta pergunta em termos de "responsabilidade". Ao professor é atribuída, tendo por base a confiança, uma turma de jovens para serem educados, e embora o professor e as crianças estejam juntos na sala de aula durante muitos dias durante o ano, a relação entre eles não é uma relação parental. O que acontece entre a criança e os pais habitualmente não diz respeito a mais ninguém, mas o mesmo não é verdadeiro em relação ao que acontece entre o professor e os jovens. Os pais, a escola e até mesmo o Ministério da Educação e o país, têm o direito de ser informados sobre os progressos que os jovens fazem ao longo das aulas. Também os próprios jovens têm o direito de o saber. Por isso, pensamos que devemos realizar avaliações para demonstrar às diferentes partes interessadas "o quê" e "o quanto" os alunos aprenderam.

Gostaria também de acrescentar que uma nota relativa à aprendizagem nas ciências, quer no que respeita à parte prática quer à teórica, NÃO é claramente o único resultado esperado do ensino e da aprendizagem das disciplinas da área de Ciências. Existem outros resultados possíveis, tais como gostar-se de aprender ciência, ter interesse pela disciplina, ou o contrário — desinteresse ou mesmo repulsa pela matéria — estes resultados são tão importantes como as notas, mas provavelmente não são passíveis de avaliação. Um interesse pela ciência que se prolongue ao longo da vida é um grande benefício, e é por acreditarmos que o trabalho experimental contribui para o surgimento desse interesse de uma forma significativa que nos compraz constatar que as crianças portuguesas estão a ter mais oportunidades de beneficiar dessas condições.

Esta pergunta pode ser formulada e respondida de outro modo: "Qual é o objectivo da avaliação? Poderemos responder que queremos avaliar o que os alunos sabem num dado momento, porque estes precisam de atingir um determinado padrão. Poderemos afirmar, por exemplo, que concluíram a sua educação básica e existe a necessidade de saber se estão preparados de acordo com um padrão adequado para prosseguirem os seus estudos. Quando decidimos constatar, numa determinada fase dos seus estudos, qual o nível da aprendizagem que atingiram, estamos a fazer uma avaliação sumativa. No Reino Unido este tipo de avaliação escolar era realizada aos 11, 16 e 18 anos de idade. Facilmente se constata a razão pela qual a avaliação sumativa era importante nessas alturas. A avaliação afectava decisões que tinham de ser tomadas relativamente ao tipo de escola, ou ao tipo de curso, ou até mesmo ao futuro emprego do aluno em questão. Só pelo facto de ser tão importante os alunos ficavam muito nervosos. O sistema de atribuição de notas era tão rígido, justo e rigoroso quanto possível. Lembro-me bem de como todos aparecíamos muito pálidos e mal dormidos no primeiro exame que fizemos aos 16 anos.

Hoje em dia, no Reino Unido, existe também uma avaliação sumativa semelhante aos 7 e aos 13 anos. No entanto, nestas idades não é preciso tomar decisões importantes. Infelizmente o exame (designado SATs) produz nos alunos exactamente o mesmo stress que acabei de descrever e, pela minha parte, não vejo qualquer justificação para este procedimento.

Também existem outros tipos de avaliação. A grande maioria das crianças têm testes no final de cada módulo de trabalho. Um dos meus professores de Latim começava a aula com um teste de vocabulário. Para dizer a verdade eu fiz o mesmo quando fui professora de Física numa escola secundária. Os alunos fingiam queixar-se, mas na verdade não se importavam. Os resultados não eram divulgados, mas como as respostas eram corrigidas imediatamente na aula, os alunos tinham um feedback imediato que lhes permitia melhorar os seus conhecimentos e ter informação sobre a forma como poderiam ter melhores resultados na avaliação seguinte — o que demonstra que se trata de um tipo de avaliação muito diferente. Estes resultados dos testes foram úteis para mim, porque

mostravam quais eram os problemas dos alunos e se alguns tópicos tinham sido introduzidos demasiado depressa. Os resultados também eram úteis para os alunos, indicando-lhes o que não tinham acertado e qual era a resposta certa. Esta avaliação é designada por avaliação formativa, porque ajuda ao desenvolvimento da aprendizagem dos alunos. Há quem designe este tipo de avaliação como avaliação diagnóstica, pois tal como o diagnóstico de um médico quando examina um doente, este tipo de avaliação tenta detectar quais os problemas que existem e, ao fazê-lo, como podem ser resolvidos. Trata-se de uma parte crucial de qualquer bom processo de ensino/aprendizagem.

A diferença entre a avaliação sumativa e a formativa é um pouco como a diferença entre as perspectivas de Piaget e Vigotsky sobre a aprendizagem. Piaget acreditava que a aprendizagem só acontecia depois de um aumento da maturidade intelectual. Não se podia acelerar esse processo, só era possível observar quando tal tinha acontecido. Quando Barbel Inhelder e os colegas pediram a Piaget para participar num projecto que iria tentar pressionar as crianças a passar para a fase epistémica seguinte, este recusou, mas permitiu-lhes que prosseguissem. Quando terminaram, pediram ao "Maitre" para escrever o prefácio do livro, e assim aconteceu. Neste prefácio aparecem as habituais frases simpáticas, e também a crítica: "Fragmentos de aprendizagem isolados não são desenvolvimento". (Tal como diríamos em inglês — com amigos destes quem é que precisa de inimigos!) Os seguidores de Piaget estão sempre a realizar testes, mas estes testes têm por objectivo verificar qual o estágio que os alunos atingiram, mas não aquilo que estão prontos a aprender. Vigotsky acreditava que, com ajuda de um professor ou dos pais, as crianças podem progredir e aprender mais, desde que o ensino partisse daquilo que já tinha sido compreendido e que as crianças estivessem preparadas para aprender. Vigotsky afirmou que o que era importante era ter a certeza que as crianças se encontravam na "zona de desenvolvimento próximo", o que quer dizer que este autor acreditava no ensino. O objectivo da avaliação, segundo Vigotsky, seria descobrir se os alunos estavam prontos para entrar na fase seguinte, isto é, tratar-se-ia de uma avaliação formativa, orientada para a melhoria do ensino.

2. Vantagens e desvantagens da avaliação formativa e sumativa

Tem sido praticamente impossível descrever as diferenças entre os dois tipos de avaliação sem recorrer às suas vantagens e desvantagens. O que ainda não discutimos é se há resultados de investigação que venham corroborar os argumentos de cada uma das correntes.

1. Será que a avaliação sumativa apresenta provas irrefutáveis a partir das quais podemos prever o sucesso dos alunos nas aprendizagens futuras?

Os resultados das investigações são ambíguos. Por um lado sabemos que ter uma boa nota num teste é uma espécie de recompensa e existe a tendência para que o desempenho melhore, na mesma medida em que a vergonha devido a um mau desempenho faz com que os alunos tenham a tendência para piorar os seus resultados. É como a aversão clássica ao tratamento que tem sido estudada nos ratos de laboratório!

Para os seres humanos, o empenho para alcançar bons resultados, ou seja, a motivação, tem a ver tanto com a probabilidade de êxito como com a auto-estima em relação à tarefa. Os dados obtidos a partir de testes concretos mostram grandes diferenças no resultado afectivo (orgulho ou vergonha), relacionadas não só com o êxito ou o insucesso, mas também com as próprias características do teste. (Os resultados do teste traduzem, na opinião do aluno, a sua aptidão, ou correspondem mais a uma questão de sorte?)

Fig. 1 - Teoria da Atribuição (Weiner)

Eixo vertical: reacção emocional (por exemplo, orgulho) ; Eixo horizontal: dificuldade da tarefa. As curvas de cima correspondem a tarefas cujo sucesso é reconhecido como dependendo de uma aptidão (a cheio, êxito; a tracejado, insucesso). As curvas mais abaixo correspondem a tarefas reconhecidas como dependendo essencialmente do factor sorte (a cheio, êxito; a tracejado, insucesso)

Tal como a Teoria da Atribuição (Weiner) afirma, o insucesso num teste em que o resultado é apenas uma questão de sorte, ou que se sabe, à partida, que é difícil, não produz um grande impacto na auto-estima dos alunos. Pelo contrário, um aluno ou aluna a quem é dito que ele ou ela obteve maus resultados numa tarefa fácil, que depende de uma aptidão, ficará envergonhado(a). Do mesmo modo, se um(a) aluno(a) obtém bons resultados numa tarefa que depende de alguma aptidão e era difícil, ficará muito satisfeito(a) e muito orgulhoso(a). Daqui decorre que a forma como os professores apresentam o método de avaliação e os seus resultados tem um impacto muito forte, tanto sobre a auto-estima dos alunos como sobre o seu futuro desempenho, sobretudo no caso dos alunos com mais dificuldades. Se os professores os tentarem encorajar dizendo que o teste é fácil, sentir-se-ão envergonhados se falharem. Por outro lado, se as questões forem apresentadas como difíceis, não ficarão preocupados em falhar, e poderão mesmo nem se aplicar muito. Quer a aptidão quer a auto-estima afectam os resultados desta avaliação, tornando-

os pouco fiáveis, sobretudo em situações em que se pretenda avaliar um esforço contínuo .

No Reino Unido existem também alguns dados sobre a capacidade de previsão dos testes sumativos em relação à capacidade real dos alunos para estudar Ciências, a partir dos dados da admissão às universidades (Solomon 1996). A maior parte das universidades recolhe os dados obtidos no exame de admissão — o Advanced Level (Nível Avançado), o nosso exame sumativo mais rigoroso — e faz a sua correlação com a classificação obtida três anos depois no "honor degree". Para aqueles raros estudantes que atingem as duas classificações mais elevadas no exame de admissão (A ou B) verifica-se, em certa medida, uma correlação com a classificação obtida três anos depois. No entanto, para os outros alunos, os que têm classificações C, D, E e F no exame de admissão, os resultados não eram melhores do que se se tivesse usado qualquer número aleatório para prever a classificação final, e era mesmo possível que os alunos tivessem um insucesso total e desistissem da universidade. Mais uma vez, a Teoria da Atribuição dá-nos algumas pistas sobre o motivo por que isto acontece.

Concluimos assim que os testes sumativos afectam tanto a motivação dos alunos que distorcem os resultados obtidos. Esta poderá ser a razão pela qual não é possível obter previsões certas quanto ao desempenho futuro dos alunos, que, no fundo, é a razão primordial pela qual submetemos os nossos alunos a este tipo de avaliação.

2. Será que a avaliação sumativa ajuda os estudantes a estudar melhor?

Desta vez os resultados são menos ambíguos. Os métodos de aprendizagem diferem e é comum distinguir entre aprendizagem superficial e aprendizagem profunda. Aprendizagem superficial é uma aprendizagem de cor, para reter os factos, ao passo que a aprendizagem profunda tenta compreender o significado e os elos de ligação entre os efeitos nos diferentes domínios. John Biggs (1979) e outros demonstraram que é possível colocar a aprendizagem dos alunos numa ou noutra categoria com uma série de questões bem testadas. Não constituirá uma surpresa para ninguém aqui presente

que é relativamente fácil demonstrar que a avaliação por exames incentiva a aprendizagem superficial (Ramsden et al. 1987). Até mesmo aqueles que estudaram de uma forma aprofundada e significativa ao longo do ano deixam de o fazer e passam a aprender de cor quando o teste se aproxima. Com a introdução dos testes sumativos aos 7 e 13 anos de idade no Reino Unido, as vendas de livros do tipo "Conteúdos Essenciais" subiu, assim como o tempo passado na sala de aula dedicado a "Pontos que devem ser aprendidos", escritos no quadro, durante o período anterior às avaliações! Isto não é uma boa aprendizagem nem um bom ensino. Este método reduz também drasticamente o tempo disponível para aulas novas e interessantes. O tempo assim como os custos devem ser sempre considerados quando se tomam decisões sobre métodos de avaliação.

3. Será que existem provas de que ao melhorar-se a avaliação formativa se aumentam os padrões de qualidade?

As provas que irei apresentar são de um estudo muito completo realizado por Black e William (1998), que utilizou 250 fontes educacionais.

a) A avaliação formativa, sob condições controladas, produz um efeito de dimensões típicas entre +0.4 e +0.7. (Isto poderia colocar Portugal algures entre os dez primeiros nas tabelas do TIMSS — 3º Estudo Internacional de Matemática e Ciências).

b) Uma avaliação formativa melhorada, que forneça um bom feedback aos alunos, ajuda mais os que têm um desempenho mais baixo do que os outros. São precisamente estes alunos que mais precisam da nossa ajuda.

c) Em quase todos os países, os inspectores são de opinião que a avaliação formativa é utilizada de forma insuficiente porque os professores não têm consciência da sua importância.

A avaliação e a atribuição de notas são demasiadamente valorizadas, em detrimento dos bons conselhos e da aprendizagem.

Com frequência, os professores são capazes de prever os resultados dos alunos num teste externo — pois os seus próprios testes imitam-nos —mas ao mesmo tempo sabem demasiado pouco sobre as respectivas necessidades de aprendizagem .

Black e William (1998)

d) Alguns estudos mostram que a rapidez com que é dado aos alunos o feedback sobre as respostas certas é um factor que tem um efeito importante sobre o seu progresso na aprendizagem.

Fig. 2 - Aprendizagem e esquecimento.

Eixo vertical: percentagem relembrada; eixo horizontal: tempo. Estão assinaladas com setas duas sessões de reciclagem.

Tendo em conta todos os dados disponíveis (ver Black e William 1998), não existem provas de que a avaliação formativa seja menos válida nos seus resultados do que a sumativa, quanto ao êxito nas previsões relativamente ao desempenho futuro dos alunos.

3. Métodos de avaliação do trabalho experimental

Existe, em Inglaterra, uma longa tradição de avaliação sumativa do trabalho experimental no nível Avançado (18 anos), e, embora em

menor escala, aos 16 anos. Os exames experimentais eram muito comuns durante os anos sessenta, embora não existisse quase investigação nessa altura sobre o que devia ser avaliado, nem como deveria ser realizada a avaliação. Por isso, existia uma grande variedade de tipologias incluindo:

w aptidões isoladas — tais como fazer a leitura do volume de um líquido numa proveta ;

w experiências holísticas inteiramente baseadas no conhecimento adquirido, tal como a identificação dos iões negativos e positivos numa solução, que era um trabalho prático que se tinha a certeza que os alunos tinham sido ensinados a fazer;

w novas experiências, como, por exemplo, explorar as oscilações horizontais de uma haste com massas adaptadas quando suspensa por dois fios paralelos.

O último tipo de avaliação exigia dos alunos que fizessem a leitura cuidadosa de instruções a seguir, e que depois elaborassem um gráfico com os resultados obtidos. Este tipo de exame prático era muito difícil de realizar no período de tempo previsto. Havia ainda o problema dos instrumentos que não funcionavam durante o exame ou do aparelho que se avariava a meio do exame! Durante os anos oitenta, mas antes da introdução do “National Curriculum”, gerou-se uma insatisfação crescente relativamente à avaliação do trabalho prático através de exames sumativos.

Uma alternativa consistia na atribuição pelos professores de uma nota ao trabalho prático dos alunos durante todo o curso, de acordo com uma grelha acordada com a Comissão de Exames, sendo o trabalho dos alunos aberto a inspecção. Tendo em conta que tinha havido informação dada pelos professores aos alunos durante o ano, esta avaliação passou a ser formativa. Nas ocasiões em que tal era realizado, para alunos dos 11 aos 14 anos (GASP, Iredale 1990), e mais tarde para a área vocacional com os alunos dos 16 aos 18 anos (GNVQ, Solomon 1995), a avaliação indicava que a motivação melhorava, porque os próprios alunos estavam envolvidos nessa avaliação. Com frequência eram os alunos que assinalavam aquilo que tinham conseguido alcançar na sua lista de competências, em vez de receberem apenas uma nota sem explicações, tal como acontecia no tipo anterior de avaliação através de exame.

Existia ainda outro tipo de trabalho prático, mais ambicioso, que estava limitado, penso eu, ao nível avançado de avaliação do Nuffield Physics, e a alguns cursos universitários. Neste caso, os alunos faziam uma pequena investigação original, que tinham elaborado sozinhos, em horário extra escolar. Conheço um pouco esta experiência, pois o meu filho e a minha filha mais nova tiveram a oportunidade de participar neste tipo de trabalho. Nós tínhamos cristais piezeléctricos a crescerem em soluções por toda a casa e tínhamos conversas infindáveis sobre as vibrações dos violinos. O trabalho era suposto ser realizado em 20 horas, e eu duvidei que tal fosse possível. Os meus filhos gostaram imenso, aprenderam imenso, mas nem um nem outro conseguiram atingir o resultado final que se tinham proposto, porque seleccionaram um tópico demasiado difícil, mas mesmo assim tiveram notas altas. Claramente os alunos foram avaliados pelos progressos que tinham alcançado, e pela sua compreensão de um projecto ambicioso, e não pelo resultado final. Não será necessário acrescentar que também era uma avaliação muito dispendiosa em termos de tempo para o próprio examinador.

O tempo que quer os alunos quer os examinadores dispendem com a avaliação é um facto importante a ter em conta na avaliação do trabalho experimental. Para os alunos tem de se acentuar que o importante é fazerem a experiência, obterem os dados e tirarem conclusões e não o escrever um longo relatório literário que pouco tem a ver com o trabalho prático realizado. Para todos, alunos, professores e examinadores, essa concisão ajuda a a reduzir o tempo gasto na avaliação, além de aumentar a sua fiabilidade.

Durante os anos oitenta houve ainda um grande projecto de investigação, financiado pelo governo do Reino Unido, com o objectivo de explorar como é que o trabalho prático pode ser avaliado da melhor forma — o APU. Não é possível descrever aqui todos os resultados interessantes obtidos, por isso limitar-me-ei a afirmar que o erro mais óbvio na prática experimental, em todas as idades, estava na identificação e no tratamento das variáveis. As crianças discutiam a cronometragem de carros a descerem um plano inclinado, mas largados de locais diferentes! No entanto, tinham melhores desempenhos a realizar experimentalmente este tipo de comparações do que a planear aquilo que iriam fazer, o que foi surpreendente e

interessante. Este tipo de experiências, que envolvem o controlo de variáveis, atraíram mais atenção do que as outras, porque, suponho, pareciam mais ligadas à abordagem de Piaget sobre a aprendizagem. O controlo de vários factores, como o volume de água utilizado para dissolver cristais, a temperatura da água, se a solução iria ser agitada ou não, a dimensão dos cristais, tudo isto iria fazer apelo ao raciocínio formal muito cedo. Se as tarefas fossem reduzidas somente a uma ou duas variáveis independentes, era mais fácil de ensinar. Rapidamente este tipo de trabalho prático passou a ser o protótipo de toda a boa avaliação e foi designado simplesmente como experiência controlada ("fair test")! O exemplo da dissolução foi tão amplamente adoptado nas escolas britânicas, que os alunos britânicos realizam essa experiência pelo menos três vezes durante o seu percurso escolar. A avaliação sumativa teve o seu efeito habitual sobre o ensino, a aprendizagem e a motivação. Os nossos alunos, que antes corriam para o laboratório perguntando: "Que experiências é que vamos fazer hoje?", agora, com frequência, limitam-se a olhar aborrecidos. Os professores fizeram o melhor que podiam para preparar os alunos para as experiências controladas e para fazer relatórios de modo a obter notas altas de acordo com as complexas grelhas de avaliação. Estas seriam estabelecidas de modo a atingir aquele padrão mítico de "rigor" que asseguraria uma descrição perfeita do nível dos alunos, que já vimos que é irrealista. Todo este trabalho árduo parece ter afastado qualquer divertimento do trabalho prático, o que é muito triste. Não ficou de modo nenhum claro se o padrão de trabalho prático melhorou ou piorou. Embora o "National Curriculum" defenda que os professores de Ciências devem incentivar as crianças a investigar questões colocadas por elas próprias, tal não parece acontecer com muita frequência, quer nas tarefas de avaliação, quer nas aulas normais.

4. Avaliar diferentes tipos de trabalho prático

Existem boas razões para afirmar que diferentes tipos de trabalho prático são realizados, em particular porque a Biologia exige tipos de

aptidões práticas diferentes das que são necessárias nas experiências simples de Física e Química. Nesta altura gostaria de dar um exemplo. Uma vez estava a observar uma turma de alunos britânicos, com cerca de 14 anos de idade, a fazerem um trabalho prático com gerânios tendo por base a transpiração de água através das folhas. O professor e o seu auxiliar tinham passado uma parte considerável do seu tempo a selar as extremidades dos caules com pedaços de tubo capilar com água para que pudesse ser possível medir a taxa de transpiração. Infelizmente, como é habitual acontecer nestes casos, constatou-se na manhã seguinte que apenas algumas tinham ficado bem vedadas. Isto quer dizer que grande parte da turma teve de inventar as suas próprias perguntas sobre a transpiração da água e conceber experiências nas quais mediam a perda de massa do sistema planta+copo um ou dois dias depois, o que por si só não era negativo. Alguns dos alunos queriam fazer esta experiência com vento a soprar sobre a planta, usando um secador, e depois sem o secador. Um grupo de alunas, de aproveitamento médio, decidiu experimentar se a planta consumia mais água da torneira ou mais água "natural". Água "natural" segundo as alunas, era a água do lago de biologia que era verde e tinha algas, mas, tal como argumentavam as alunas, "natural"! Dois dias depois o grupo obteve resultados claros, mas não os que esperava. A planta com água da torneira tinha produzido a maior diminuição de massa de água, ao passo que no outro copo a variação era imperceptível. As alunas ficaram desapontadas. Aqui devia começar uma nova experiência: encorajei as alunas a observar a extremidade do caule que tinha estado na água do lago. Estava macia e com um certo cheiro. Uma das alunas conjecturou que talvez existissem micróbios na água que tinham atacado o tecido do caule da planta. Ninguém pensou em usar um microscópio para comparar as extremidades dos caules das duas plantas ou para examinar uma gota de água do lago de Biologia. Ninguém considerou a possibilidade de utilizar um microscópio para comparar as extremidades dos caules ou uma gota de água do lago de Biologia. Certamente não teria melhorado as suas notas de acordo com a grelha de classificação da experiência controlada!

A razão pela qual contei esta história é para demonstrar o seguinte: se o sistema de avaliação permite apenas um tipo de investigação,

facilmente conduzirá a restrições ao trabalho prático a realizar nas escolas, o que, obviamente, é muito grave.

A grande vantagem aparente das "experiências controladas" ("fair tests") é que existem fases bem definidas e lógicas que podem ser ensinadas aos alunos. É também o método que é utilizado para fazer testes de artefactos técnicos. Em particular isto tem sido muito utilizado nas escolas britânicas para elaborar testes para descobrir que tipo de papel é mais resistente, ou que tipo de toalha absorve mais água. Tal como podem constatar, isto não é muito eficaz para promover o interesse nos fenómenos científicos, nem para clarificar conceitos, como por exemplo a luz ou o calor que estão a ser ensinados nas nossas aulas.

Note-se também que as experiências de biologia não estão facilmente em conformidade com o modelo de experiência em que uma gama contínua de leituras da variável dependente é obtida, sob condições controladas, à medida que a variável independente é alterada. Alguns de nós, no King' s College de Londres, temos discutido recentemente os diferentes tipos de trabalho prático que desejaríamos que fossem realizados pelos alunos, se começássemos tudo de novo. O Dr. Rod Watson, cujo interessante artigo sobre este assunto está ainda em publicação (Watson et al 1998), identifica seis tipos diferentes de trabalho prático. Estamos ambos de acordo, e defendemos veementemente que não existe apenas um tipo de trabalho prático, como o que nós designamos na Grã Bretanha por experiência controlada ("fair test"), e que existem muitos outros tipos diferentes de trabalhos práticos válidos e úteis para os alunos de todas as áreas científicas.

5. Conclusões

Já constatámos que a avaliação pode ter uma influência enorme sobre aquilo que acontece na escola, tanto naquilo que é positivo como menos positivo. A avaliação sumativa de acordo com um esquema rigoroso, quer para o trabalho experimental quer para o teórico, não consegue produzir previsões fiáveis, com excepção dos alunos mais

aptos a realizar testes; por outro lado, pode distorcer gravemente o leque de experiências dos alunos em termos de trabalho prático.

Em vez da avaliação sumativa, baseada num único trabalho prático feito no exame, porque não realizar avaliação formativa durante todo o ano, de modo a melhorar a qualidade do trabalho dos alunos?

Por que não estipular um determinado número de critérios diferentes que devem ser atingidos por diferentes tipos de trabalho prático? A avaliação poderá então conduzir a trabalho variado e interessante, em vez do contrário.

Por que não tornar os critérios de avaliação abertos e transparentes para que os alunos possam melhorar o nível do seu próprio trabalho, e para que os professores possam comparar os padrões que os seus alunos obtiveram com os de outras escolas?

REFERÊNCIAS

BIGGS, J., (1979), Individual differences in study processes and learning outcomes, *Higher Education*, 8 381-394.

BLACK, P. and WILIAM, D., (1998), Assessment and classroom learning. *Assessment in Education*, 5(1) (in press).

IREDALE, C., Pupils' attitudes towards GASP, *School Science Review*, 1990, 72(258), 133-137.

RAMSDEN, P., BESWICK, D. and BOWDEN, J., Learning processes and learning skills. In (ed. Richardson, J. Eysenck, M. and Piper, D. *Research in education and cognitive psychology*, 1987 Milton Keynes, Open University Press.

SOLOMON, J. (1995), *Case-studies in Advanced level GNVQ*, Oxford University Department of Educational Studies, Oxford.

SOLOMON, J. (1996), *Good practice in GNVQ Science at Advanced Level: Admission and Learning in Higher Education* Oxford University Department of Educational Studies, Oxford.

WATSON, R., GOLDSWORTHY, A. and WOOD-ROBINSON, V. (in press)
What is not fair with investigations, *School Science Review*.

WEINER, B. (1992), *Human Motivation*, London, Sage.