

QUAL PEDAGOGIA PARA OS ALUNOS EM DIFICULDADE ESCOLAR?

MARCEL CRAHAY

Faculdade de Psicologia da Educação da Universidades de Genebra e de Liege
marcel.crahay@pse.unige.ch

Tradução: Neide Luzia de Rezende

RESUMO

Repetir o ano não ajuda os alunos com dificuldades. Numerosas pesquisas sobre os efeitos da repetência mostram que em regra geral os alunos fracos que recomeçam o mesmo ano progredem menos que os alunos fracos que são promovidos. Que medidas pedagógicas são necessárias para combater o fracasso escolar? O objetivo deste artigo é proceder a uma revisão do que as pesquisas empíricas revelam sobre os procedimentos mais freqüentemente mencionados como meios de obter um aperfeiçoamento pedagógico. Os primeiros dizem respeito à composição e ao manejo de turmas homogêneas e heterogêneas, de classes grandes ou pequenas, e de grupos de alunos com necessidades específicas; os segundos remetem aos dispositivos de individualização da aprendizagem ou à estratégia oposta, de aprendizagem cooperativa e tutorial; os terceiros são relativos à avaliações formativas, seguidas ou não de ações corretivas. Discutindo o conjunto dos resultados, o autor defende um redirecionamento das pesquisas pedagógicas no sentido de investigar prioritariamente os procedimentos mais genéricos, deixando para casos extremos a mobilização de dispositivos baseados num diagnóstico fino das dificuldades de aprendizagem.

PEDAGOGIA – REPETÊNCIA – AVALIAÇÃO DA APRENDIZAGEM – DIDÁTICA

ABSTRACT

WHICH PEDAGOGY FOR STUDENTS WITH SCHOOL DIFFICULTIES? To repeat the year does not help students with difficulties. Several studies about the effects of grade repetition show that in general students who do poorly and start again the same school year progress less than weak students who are promoted. Which pedagogical measures are necessary to combat school failure? The aim of this article is to review what empirical studies reveal about the most often mentioned procedures as a means to attain pedagogical improvement. The first procedures relate to the composition and management of homogeneous and heterogeneous groups, small or big classes, and groups of students with specific needs; the second procedures relate to the learning individualization devices or to the opposite strategy of cooperative and tutorial learning; the third procedures relate to the formative evaluations, either followed or not by corrective actions. When discussing the results, the author defends to redirect the pedagogical studies so as to investigate primarily the most generic procedures,

leaving for the most extreme cases the mobilization of devices based on a fine diagnosis of the learning difficulties.

PEDAGOGY – GRADE REPETITION – LEARNING EVALUATION – DIDACTICS

A REPETÊNCIA É INEFICAZ

Desde a publicação por Jackson (1975) de uma primeira síntese das pesquisas sobre os efeitos da repetência, as provas atestando a ausência de efeitos positivos dessa prática pedagógica não param de se acumular. Em 1984, a meta-análise de Holmes e Matthews veio confirmar as conclusões de Jackson e, em 1989, Holmes acrescentaria aquilo que permanece hoje como a análise mais conhecida dos efeitos da repetência. Mais recentemente, Jimerson (2001) recen-

QUADRO I
EFEITO DA REPETÊNCIA SOBRE O DESEMPENHO DOS
ALUNOS EM TERMOS DE AMPLITUDE DO EFEITO (AE)

| Critérios Avaliados | Jimerson (2001) | | | Holmes e Matthews (1984) AE | Holmes (1989) AE |
|--|-----------------|----------------------|---------------------------------|--------------------------------------|------------------------|
| | Número de AE | Número de estudos | AE (ponderada por efeito) | | |
| Amplitude do efeito geral | 246 | 20 | -.31 | -.37 | -.15 |
| Desempenho acadêmico | 169 | 18 | -.39 | -.44 | -.19 |
| Desempenho em língua materna | 11 | 5 | -.36 | -.40 | -.16 |
| Desempenho em leitura | 52 | 11 | -.54 | -.48 | -.08 |
| Desempenho em Matemática | 48 | 10 | -.49 | -.33 | -.11 |
| Total/Compostos das médias de notas escolares | 13 | 8 | -.20 | nd | nd |
| | 45 | 6 | -.18 | -.58 | -.58 |
| Reações afetivas gerais | 77 | 16 | -.22 | -.27 | -.09 |
| Desenvolvimento social | 12 | 5 | -.08 | -.27 | -.09 |
| Bem-estar emocional | 13 | 6 | -.28 | -.37 | .03 |
| Comportamento na classe | 30 | 11 | -.11 | -.31 | -.13 |
| Imagem de si | 16 | 6 | -.04 | -.19 | -.13 |
| Ajuste social (medida composta) | 4 | 4 | -.15 | -.16 | nd |
| Freqüência escolar | 2 | 2 | -.65 | -.12 | -.18 |

Nota: Os valores negativos da amplitude do efeito (AE) significam que o desempenho dos repetentes é inferior ao dos alunos promovidos; entretanto, ambos os grupos tinham desempenhos equivalentes antes do ano da repetência. (Nd = não disponível).

QUADRO 2

| | T1 maio 2001 | T2 nov. 2001 | T3 maio 2002 | T4 maio 2003 |
|--------------------------|-----------------|-----------------|-----------------|-----------------|
| Repetentes | 2P | 2P | 2P | 3P |
| Alunos fracos promovidos | 2P | 3P | 3P | 4P |

seou vinte estudos produzidos entre 1989 e 2001 e calculou a amplitude do efeito sobre os dados que não puderam ser levados em consideração por Holmes (1989). O quadro 1 compara os resultados dessa última meta-análise com aqueles obtidos anteriormente por Holmes e Matthews (1984) e Holmes (1989): todas as consequências do efeito são negativas nas três meta-análises.

Os raros estudos realizados sobre o tema nos países francófonos trazem conclusões pouco favoráveis à repetência. A seguir, apresentamos dois estudos recentes, um realizado na Suíça e outro, na França.

Bless, Bonvin e Schüpbach (2005) realizaram um estudo com 4.248 alunos da Suíça alemã e da Suíça romanche, egressos de classes do segundo ano primário (2P), ou seja, de crianças de 6-7 anos, no momento da primeira medida, quer dizer no momento em que as repetências são decididas. Com base nos dados recolhidos no mês de maio de 2001 (T1) no interior dessa grande amostra, os pesquisadores puderam compor 83 pares de alunos repetentes *versus* alunos fracos promovidos (ou não repetentes), pareados sobre a base de características demográficas, do desempenho em Matemática e em língua, bem como nos seus escores de QI. O desempenho desses 166 alunos foi novamente testado em novembro de 2001 (T2), depois em maio de 2002 (T3) e por fim em maio de 2003 (T4). Esse plano de pesquisa (Quadro 2) permite a Bless, Bonvin e Scüpbach (2005) proceder a quatro comparações, mantendo a idade constante, e a uma comparação, mantendo constante a série dos alunos (ao final do terceiro ano – 3P).

Do conjunto dos resultados analisados, percebe-se que ao final de um ano (medidas tomadas no T3), o desempenho em matemática dos alunos repetentes é significativamente inferior ao dos alunos fracos promovidos. Porém, os pesquisadores não observam diferenças no nível do desempenho em língua. Ao final do terceiro ano primário, os alunos repetentes têm desempenho em Matemática e em língua superior ao dos alunos fracos promovidos, mas são um ano mais velhos. Isso leva os autores a dizer:

Se se consideram unicamente as comparações mesma-idade, a promoção é, apesar dos fracos desempenhos escolares, a curto e médio prazo, preferível à repetência. Ao contrário, nas comparações mesma-série, a repetência se mostra, a curto e médio prazo, superior à promoção. (p.127)

Ressaltemos, contudo, que na comparação com um grupo “padrão” de alunos promovidos (quer dizer, um grupo “proveniente” de alunos que freqüentam pela primeira vez o segundo ano primário (2P)), os repetentes têm desempenho mais fraco e progridem menos rapidamente.

Na França, Troncin (2005) chega a conclusões ainda mais desfavoráveis quanto à repetência. Concentrando-se em alunos que começaram o primeiro ano do curso primário (alunos de seis-sete anos) em setembro de 2003 e acompanhando a evolução deles por dois anos, esse pesquisador mostra uma queda significativa no desempenho, ao longo do verão que se seguiu ao primeiro ano, dos alunos que repetiram: desvios médios entre o teste de junho e o de setembro (provas idênticas) dos repetentes oferecem sistematicamente sinais negativos, enquanto os dos alunos fracos promovidos oferecem sistematicamente sinais positivos. Os primeiros tiveram, em setembro, em média um escore global inferior de três pontos ao de junho, enquanto os segundos “ganharam” quatro pontos durante esse mesmo período. Parece, pois, que a decisão da repetência, uma vez comunicada aos alunos, tem um efeito desmobilizador. No entanto, os repetentes começam o ano com um nível escolar superior ao de seus colegas que entram no primário, mas essa vantagem inicial estanca ao longo do tempo a ponto de desaparecer totalmente no fim do ano. Em relação aos alunos fracos da mesma idade que foram promovidos, os repetentes mostram um atraso que indica que teria sido preferível para eles não ter que recomeçar do início o programa do 1º ano. É evidente que, como se vê em todos os outros estudos sobre o assunto, os repetentes progridem durante o ano que eles repetem, mas sua margem de progresso médio é inferior à dos alunos que, apresentando o mesmo nível no fim do 1º ano, foram promovidos para o 2º. Mais precisamente, sobre 103 duplas compostas de um repetente e de um não-repetente pareados em função de diversas características comuns, Troncin (2005) observa uma progressão significativa dos promovidos fracos em 77 dos casos e uma progressão significativa dos repetentes em 26 dos casos. As progressões dos repetentes são, contudo, de uma amplitude

menor do que aquelas dos promovidos fracos. Assim, quando nos detemos nos casos em que as diferenças nas duplas são superiores ou iguais a um desvio padrão da distribuição dos escores do grupo dos promovidos fracos, menos de uma evolução sobre dez é a vantagem dos repetentes¹. Por sua vez, no plano socioafetivo, as entrevistas realizadas com as crianças repetentes confirmam as conclusões de Crahay (2003); o fato de repetir um ano gera um "estigma" que vem acrescentar problemas afetivos e de relação às dificuldades de aprendizagem, o que poderia ser poupadão a esses alunos.

O balanço das pesquisas disponíveis sobre os efeitos da repetência não tem ambigüidade: em regra geral, os alunos fracos que repetem progredem menos que os outros alunos fracos que são promovidos. Ou seja, pode-se considerar que a repetência constitui um meio contraprodutivo de fazer face às dificuldades de aprendizagem dos alunos fracos. É, pois, necessário e urgente procurar outros meios de ação. É esse o objetivo deste trabalho. Mais precisamente, o objetivo deste texto é passar em revista aquilo que as pesquisas empíricas revelam quanto a diferentes procedimentos pedagógicos freqüentemente evocados como meios de melhoria pedagógica. Primeiramente, destaca-se a composição e a gestão dos grupos, em seguida, trata-se de observar os dispositivos de individualização e, por sua vez, as estratégias de aprendizagem cooperativas e tutoriais. Num terceiro momento, comentam-se os testes formativos seguidos ou não de procedimentos corretivos.

COMPOSIÇÃO E TAMANHO DA CLASSE

Muitos são os professores convencidos de que a composição das classes, e, mais ainda, seu tamanho constituem os determinantes da eficácia pedagógica. Ou seja, a eficácia pedagógica poderia ser aumentada de modo significativo se os professores se deparassem com classes homogêneas de tamanho reduzido. A pesquisa tratou separadamente essas questões. Examinaremos portanto as pesquisas voltadas para a composição das turmas, depois aquelas sobre a redução do número de alunos por classe, antes de levar em consideração os efeitos da flexibilização das classes pela organização de grupos com necessidades semelhantes.

I. Observa-se ainda que não se conseguiu identificar nenhum preditor para explicar certas progressões de repetentes mais elevadas que a progressão média dos repetentes.

Classe homogênea ou classe heterogênea

A idéia subjacente à constituição de classes homogêneas pode ser formulada como segue: quando o professor está diante de uma classe composta de alunos de nível escolar homogêneo, ele pode mais facilmente encontrar atividades adequadas, bem como um ritmo de ensino que convenha a todos. No caso das classes heterogêneas, o professor sempre mais rápido e mais lentamente: mais rápido para os alunos mais fracos, e mais lentamente para os mais fortes. Em resumo, todo mundo perderia: os melhores porque seriam freados na sua progressão, e os mais fracos porque não conseguiriam seguir o ritmo dos outros e terminariam por recuar.

O que a pesquisa diz a esse respeito?

É preciso antes de mais nada distinguir dois tipos de estudos. Num primeiro tipo, chamado estudos naturais, os pesquisadores observam classes por nível de desempenho criadas pelos diretores de estabelecimentos e se abstêm de intervir. Num segundo tipo, os estudos experimentais ou quase-experimentais, os pesquisadores se propõem a apreender o efeito específico desse modo de agrupamento. Para tanto, eles buscam neutralizar o efeito de outras variáveis a fim de respeitar o princípio da comparação em condições semelhantes. Eles intervêm então para evitar que os professores menos motivados acabem ficando com as classes dos alunos fracos e, sobretudo, para evitar que as oportunidades de aprendizagem oferecidas a esses alunos sejam de menor qualidade em relação ao que é oferecido aos alunos considerados fortes. Nesse caso, procura-se, pois, proceder a comparações de conteúdo de ensino igual e de qualidade pedagógica equivalente.

Dupriez e Draelants (2004) realizaram uma recensão completa dos trabalhos do primeiro tipo. Concluíram que a constituição de classes homogêneas quanto ao nível escolar tem com freqüência efeitos prejudiciais no plano socioafetivo, principalmente para os alunos situados nas classes dos fracos. Ocorre que os professores das classes dos fracos têm geralmente tendência a adotar uma atitude fatalista. Assim, segundo Oakes (1992), os alunos situados nos grupos ou turmas dos fracos recebem um ensino de menor qualidade: nessas turmas, o tempo dedicado ao ensino é inferior àquele registrado nas turmas consideradas fortes, o número de unidades de conteúdo é menor, os

encorajamentos são mais raros e menos entusiastas e os exercícios de repetição (*drill and practice*) mais numerosos.

Em certos países do mundo, a partir do ensino médio, os alunos de acordo com seu desempenho são encaminhados a cursar certos arranjos curriculares mais ou menos prestigiados socialmente, os quais facultam o acesso aos cursos universitários mais conceituados ou mesmo limitam as possibilidades ulteriores de estudo e profissionalização: os melhores são dirigidos para as trajetórias mais nobres e os outros para as trajetórias menos exigentes (como a *realschule* na Alemanha) ou profissionalizantes. A maioria das pesquisas que comparou os avanços dos estudantes em trajetórias diferenciadas observa importantes diferenças: os inseridos numa trajetória nobre progridem de maneira significativa enquanto aqueles vinculados a trajetórias menos prestigiadas não progridem e, em certos casos, regridem (Crahay, 2000).

Quanto ao segundo tipo de estudos – as comparações de conteúdo e qualidade pedagógica equivalentes – deve-se a Slavin (1987, 1990) duas cuidadosas revisões: a primeira referente ao ensino fundamental e a segunda, ao ensino médio. Os resultados são similares para os dois níveis.

Em certos programas inovadores realizados nos Estados Unidos no ensino fundamental, os alunos são agrupados em função de seus resultados num teste de aptidão geral. As classes assim constituídas são homogêneas no sentido de que elas agrupam, durante o ano inteiro, alunos que possuem QI relativamente equivalente; há classes de alunos fortes, médios e fracos. Slavin (1987a) recenseou 14 estudos e comparou a evolução de crianças freqüentando esse tipo de classes homogêneas e um outro grupo, inicialmente do mesmo nível, freqüentando classes heterogêneas. A amplitude do efeito foi calculada sobre o conjunto dos alunos e, sempre que possível, sobre os três grupos de alunos: fortes, médios e fracos. Para o conjunto dos alunos, a amplitude do efeito vale zero. Nesses 14 estudos nenhuma tendência surge no que concerne os alunos fortes, médios ou fracos. Em resumo, não se pode concluir, dessas comparações de conteúdo e qualidade pedagógica equivalentes, que o agrupamento em classes homogêneas seria favorável aos melhores e prejudicial aos mais fracos, como se poderia pensar a partir dos estudos em meio natural.

A tendência dos resultados é idêntica no ensino médio. Slavin (1990) conseguiu encontrar seis estudos experimentais e nove quase-experimentais

nos quais os pesquisadores comparam o progresso de alunos agrupados em classes homogêneas ao de alunos agrupados em classes heterogêneas; ambos os grupos são submetidos ao mesmo programa de estudo. Nos seis estudos experimentais, os alunos são distribuídos pelos grupos de modo aleatório; nos estudos quase experimentais, os alunos são pareados de modo minucioso.

Ao lado desses quinze estudos que apresentam sólidas garantias de legitimidade, Slavin (1990) considera 14 estudos, nos quais os pesquisadores tomaram como ponto de partida classes homogêneas (ou heterogêneas) já constituídas e procuraram em seguida classes heterogêneas ou homogêneas comparáveis². Calculada sobre o conjunto das 20 pesquisas que permitem essa medição, a amplitude do efeito vale zero (-0,02 exatamente). Se nos limitarmos, rigorosamente, aos estudos experimentais e quase-experimentais, obtemos uma amplitude do efeito equivalente a -0,06. Ou melhor, nove dos treze estudos dão uma leve vantagem às classes heterogêneas. Quinze estudos fornecem resultados tais que a amplitude do efeito pode ser calculada separadamente para os alunos bons, médios e fracos. Novamente, constata-se a ausência de efeito substancial: os valores médios obtidos são +0,1 para os alunos fortes, -0,08 para os médios e -0,02 para os fracos. A hipótese segundo a qual o agrupamento em classe homogênea seria benéfico aos alunos fortes cai por terra. Ou seja, significa que os alunos fortes não são prejudicados pela companhia de alunos médios ou fracos.

Assim, as comparações com conteúdo e qualidade pedagógica equivalentes indicam um efeito nulo da composição das classes. Esse resultado é oposto àquele que advém dos estudos naturais. Esse contraste é em si interessante. A nosso ver, indica que o efeito próprio ou específico do agrupamento dos alunos por nível de desempenho é nulo. Pela expressão "efeito próprio" pretendemos sublinhar que, se a qualidade e a quantidade de ensino dispensadas permanecem constantes, a maneira com que os alunos são agrupados não afeta o rendi-

2. Nesses últimos estudos, os pesquisadores asseguram somente a equivalência dos grupos no nível da classe, enquanto que, nas pesquisas quase-experimentais, foram pareados os alunos. Além disso, os estudos abordados por Slavin (1990) se referem no mínimo a um semestre, a maioria a um ano. A comparação foi estabelecida ao menos em três classes homogêneas e três classes heterogêneas. Em certos estudos, um mesmo professor se responsabilizava por uma classe homogênea e uma classe heterogênea.

mento, o que se verifica tanto em relação aos alunos fortes quanto aos alunos fracos. Ao contrário, quando os pesquisadores não intervêm para manter a qualidade e a quantidade de ensino iguais nas diferentes classes, é possível que, ao se constituírem trajetórias diferentes, estas sejam valorizadas de forma hierarquizada, portanto, o ensino dispensado varia em função do prestígio social concedido a cada uma dentre elas. A diferenciação do ensino em função da trajetória constitui assim um componente do dispositivo escolar que o torna desigual.

É preciso diminuir o tamanho das classes?

No clássico *School class size*, Glass et al. (1982) apresentaram uma minuciosa meta-análise de 77 pesquisas experimentais, realizadas essencialmente nos Estados Unidos, a fim de responder à questão: a redução do tamanho das classes permite aumentar a eficácia pedagógica? Eles mostram, por um lado, que a relação entre tamanho da classe e rendimento escolar é particularmente visível através dos 14 estudos experimentais em que os alunos foram distribuídos aleatoriamente nas classes e, por outro lado, que a relação entre essas variáveis é logarítmica: o aumento do rendimento é claramente mais importante se forem eliminados cinco alunos numa classe de 15 do que se se eliminar o mesmo número numa classe de 25. Mais precisamente, a redução dos efetivos mostra seus efeitos sobretudo abaixo de dez alunos.

Glass e seus colaboradores também recensearam 60 pesquisas nas quais foram relacionados tamanho da classe e atitudes dos alunos. A conclusão desses autores é igualmente favorável à hipótese da redução de tamanho da classe e se traduz de novo por uma função logarítmica que evidencia que os efeitos de uma redução do número de alunos se manifestam mais rapidamente sobre a motivação dos alunos e o clima da classe do que sobre o aprendizado.

Os resultados de Glass et al. (1982) foram amplamente confirmados durante um vasto programa de pesquisa realizado no Tennessee e intitulado Star (Molnar, 1999). Essa pesquisa de grande amplitude (6.500 alunos e 79 escolas), lançada em meados dos anos 80, acompanhou os mesmos alunos até o fim do ensino superior e desse modo avaliou os benefícios a longo prazo da freqüência a uma classe de tamanho pequeno durante quatro anos (do 3º maternal ao 3º ano primário). Essa pesquisa longitudinal é também experimental: os sujeitos foram escolhidos de maneira aleatória em classes com poucos

alunos ou em classes mais numerosas; os pesquisadores verificaram em seguida se os alunos dos dois tipos de classe apresentavam características comparáveis do ponto de vista de etnia, sexo e estatuto socioprofissional dos pais. Os resultados dos alunos nas classes mais numerosas e menos numerosas foram comparados em diferentes momentos de sua escolaridade mediante diversos indicadores (sucesso nos testes, taxa de fracasso, taxa de conclusão de curso).

Todos os resultados confirmam a vantagem das classes de tamanho reduzido. Resumimos as constatações principais:

- do 1º ao 8º ano (quer dizer, do primeiro ano primário ao equivalente do 2º secundário), observam-se diferenças estatisticamente significativas entre o desempenho em língua e em Matemática dos alunos das classes pequenas e das classes mais numerosas;
- ganhos importantes são registrados em todos os meios – urbano, rural, bairros de periferia, guetos urbanos;
- em cada nível escolar, observa-se que são as crianças mais desfavorecidas – as crianças negras e hispânicas, aquelas dos bairros “difíceis” – que se beneficiam mais da redução do tamanho das classes. A vantagem é duas ou três vezes mais importante para os jovens “vulneráveis” do que para os alunos brancos de classe média. O tamanho tem tipicamente um papel “compensatório”; trata-se literalmente de uma “discriminação positiva”, que corrige certas desigualdades devidas ao meio social de origem das crianças;
- os efeitos são importantes, cumulativos, e se mantêm, quer dizer se consolidam com o tempo: os alunos que permaneceram os quatro anos nas classes de tamanho reduzido têm, no 4º ano, um avanço da ordem de seis a nove meses em termos de aquisição sobre seus colegas de classes mais numerosas. No 8º ano, a vantagem sobe ao equivalente a mais de um ano de escolaridade – apesar de a experiência estar limitada aos três primeiros anos do primário;
- os alunos que freqüentaram as classes menores no início do primário são mais bem-sucedidos nos exames de ingresso no ensino superior (*college*). Para os jovens de origem afro-americana, por exemplo, a taxa de sucesso sobe de 31,7% para 40%;
- do mesmo modo, aqueles que se beneficiaram da experiência de uma classe pequena obtêm mais freqüentemente o diploma do ensi-

no superior sem vivenciar fracassos (72%) do que aqueles que freqüentam outros tipos de classe.

Na França, o estudo de Piketty (2004) chegou a uma conclusão análoga. Esse pesquisador explorou as descontinuidades ligadas ao franqueamento de níveis no início e no término do ano, para estimar o impacto de classes pequenas sobre o sucesso escolar. Graças a essa técnica, Piketty (2004) estima que a redução de um aluno em uma classe de 2º ano conduz a um aumento de ao menos 0,7 ponto do escore médio obtido pelos alunos desfavorecidos nas provas de avaliação de matemática no início do 3º ano. Tal estimativa leva o pesquisador a recomendar um direcionamento mais importante de recursos para as escolas desfavorecidas a fim de permitir uma redução do tamanho das classes.

GRUPOS COM NECESSIDADES ESPECÍFICAS DE APRENDIZAGEM

Claparède (1920) já pensava em flexibilizar o funcionamento das classes. A flexibilidade foi implementada nos Estados Unidos por ocasião do Plano Joplin (Floyd, 1954). Segundo esse plano, os alunos são vinculados, em função de suas idades, a uma classe heterogênea. Contudo, eles deixam essa classe em razão de certas aprendizagens – leitura ou matemática – que se desenvolvem em grupos homogêneos. Essa flexibilização das classes respeita escrupulosamente as três regras seguintes:

- tempo passado em grupos homogêneos é nitidamente inferior ao tempo passado em classes heterogêneas; consequentemente, o grupo com o qual as crianças mais se identificam é este último;
- a constituição dos grupos homogêneos repousa sobre a avaliação de uma competência específica e não mais sobre uma aptidão geral;
- os grupos são flexíveis: podem ser reorganizados em função dos progressos dos alunos.

Amplamente difundido nos países anglo-saxônicos, o Plano Joplin foi objeto de muitas avaliações. Dentre estas, 14 estudos foram reanalisados escrupulosamente por Slavin (1987). Dois foram conduzidos segundo um plano estritamente experimental: os professores e 360 crianças no estudo de Morgan

e Stucker (1960) e 52, naquele de Hilson et al. (1964) foram distribuídos aleatoriamente no grupo experimental (classes do programa Joplin) ou do grupo de controle (classes tradicionais). No estudo de Morgan e Stucker (1960), conduzido na Escócia em classes de 5º e de 6º anos, a amplitude do efeito é de +0,30. Quando se comparam especificamente os progressos realizados pelos alunos fracos nos dois grupos, obtém-se uma amplitude do efeito de +0,94. No estudo de Hilson et al. (1964), de duração maior (três anos para um no estudo de Morgan e Stucker) que incide sobre o início do curso primário, os resultados em leitura dos alunos do programa Joplin são claramente superiores aos dos alunos que freqüentam classes tradicionais; a amplitude do efeito é de +0,72. Nos 12 outros estudos arrolados por Slavin (1987), o plano é quase-experimental: os alunos que freqüentam as classes Joplin são pareados (em função de critérios diversos) a sujeitos que freqüentam classes tradicionais. Dez estudos apresentam garantias de validade suficientes (equivalência inicial dos grupos experimentais e grupos de controle), dos quais é feita uma síntese. Deve-se notar que o segundo estudo se refere ao ensino de Matemática, ao passo que todos os outros estudos se referem à leitura. É preciso assinalar também que nenhum estudo conduz a resultados negativos; para três somente, a amplitude do efeito é inferior a +0,20, valor considerado como limite nas meta-análises. A amplitude do efeito calculado sobre os 14 estudos referentes ao Plano Joplin é de +0,45, o que é notável.

Definitivamente, os resultados das pesquisas passados rapidamente em revista dão indicações precisas quanto à maneira de compor e de fazer funcionar as classes no interior das escolas. As classes heterogêneas com poucos alunos devem ser privilegiadas, sobretudo nos primeiros anos do ensino primário, a fim de evitar possíveis efeitos desiguais da constituição de classes heterogêneas (conforme os estudos ditos naturais). Sobre esse consenso básico, é interessante introduzir um funcionamento por grupos com necessidades específicas de aprendizagem, respeitando as três regras erigidas em princípios no Plano Joplin.

A INDIVIDUALIZAÇÃO É IGUALITÁRIA?

Muitos são os pedagogos e professores que erigem a individualização do ensino em um dos dogmas principais da escola justa e eficaz. Ou seja, quanto

mais o ensino for individualizado, mais efeitos positivos engendrará sobre todos os alunos. O núcleo duro dessa crença – acreditamos – foi desvendado por Claparède (1920) em *L'école sur mesure* [A escola sob medida]. Consiste em postular que quanto mais o ensino for adaptado às características do aluno, mais este aprenderá. Esse axioma pedagógico extrai sem dúvida sua verossimilhança do preceptorado, uma prática que, durante muitos séculos, foi a forma dominante do ensino, ainda que reservado a uma minoria: as crianças da aristocracia e de grandes famílias burguesas. Nas classes, contudo, o professor se encontra diante de um grupo de alunos; ele não pode se desdobrar. Para contornar essa impossibilidade, alguns pedagogos conceberam dispositivos³ de individualização compostos por fichas, jogos pedagógicos, materiais didáticos etc. Mais fundamentalmente, o objetivo é se aproximar o máximo possível de um ideal onde cada aluno progride a seu ritmo no domínio do conteúdo de ensino, selecionando ele próprio suas tarefas de aprendizagem. Na verdade, essa organização resulta em uma redução do aspecto relacional do ato pedagógico: confrontado ao material pedagógico, o aluno se encontra numa condição de aprendizagem solitária, o que a distingue radicalmente dos dispositivos de ensino individualizado do preceptorado.

Desde 1982, Birzea denunciava a ilusão que levava a crer que os dispositivos de ensino individualizado contribuíam para reduzir as desigualdades entre alunos. Segundo esse autor, num contexto onde cada aprendiz se acha sozinho em face do material didático e progride em seu próprio ritmo, há o risco de ver os mais rápidos progredirem cada vez mais rapidamente e os mais lentos, cada vez mais devagar. O argumento principal de Birzéa (1982), que encontramos também em Snow (1985), pode se resumir como segue: quanto mais o ritmo e as condições de aprendizagem forem determinadas pelos próprios aluno, maior é a probabilidade de aumentarem as diferenças interindividuais.

Hoje temos à disposição resultados de pesquisas empíricas para enriquecer o debate (Crahay, 2000). Assim, Burns (1984) mostrou, por meio de vários estudos, que existe uma relação negativa entre o tempo consagrado pe-

3. Entre os dispositivos de ensino individualizado mais conhecidos no mundo anglo-saxão podemos citar: Individually Prescribe Instruction – IPI –, Program for Learning in Accordance with Needs – Plan –, Individually Guided Education – IGE –, plano Keller, ou Personalized System of Instruction – PSI.

los alunos à aprendizagem e o fato de seu professor criar as condições para que cada um administre seu ritmo de progressão a sua maneira. Antes, Shimron (1976) observara que, quando os alunos são deixados livres para avançar em seu próprio ritmo, os mais fortes finalizam mais unidades de aprendizagem do que os fracos. Bangert-Drowns, Kulik e Kulik (1983) procederam a uma meta-análise das pesquisas, comparando a eficácia dos dispositivos de individualização do ensino⁴ à do ensino secundário tradicional. Abordando 51 estudos em razão de sua validade, esses pesquisadores obtêm uma amplitude do efeito médio de +0,10. Esse resultado de conjunto é decepcionante; tal amplitude do efeito é banal e deve ser interpretada como uma ausência de efeito realmente significativo. Além disso, o erro *standard* é elevado (+0,53) e as margens de variação da amplitude do efeito são importantes (entre -0,84 a +1,24), o que exclui toda possibilidade de conclusão geral.

É com uma outra meta-análise que Kulik, Kulik e Bangert-Drowns (1990) trouxeram um elemento essencial ao debate. Num balanço bastante completo de trabalhos influenciados pela aprendizagem por domínio – AD –, esses pesquisadores compararam duas modalidades de experiência: uma, reservando um lugar central ao ensino coletivo; outra, privilegiando a individualização. Kulik, Kulik e Bangert-Drowns recensearam de 36 experiências de AD, realizadas em contexto de ensino coletivo, segundo as recomendações de Bloom (1979), e 68 experiências de AD privilegiando a individualização das aprendizagens (segundo as prescrições do Plano Keller⁵) e realizadas no nível do college (ou seja, no ensino superior). Os resultados são favoráveis à manutenção de uma fase de ensino coletivo, como é o caso na AD clássica. Comparadas ao ensino tradicional, as experiências realizadas segundo o Plano Keller engendram um efeito de amplitude de +0,48, o que atesta a eficácia dos procedimentos de ensino individualizado inspirados no Plano Keller, em relação ao ensino tradicionalmente dispensado nos colleges. Entretanto, a amplitude do efeito induzida pelas experiências de AD que conservam uma fase de apresentação coletiva é superior

-
4. É bom lembrar que por "dispositivos de ensino individualizado" deve-se entender modalidades de organização pedagógica em que os alunos são confrontados sozinhos a um material didático que eles descobrem por meio de seu próprio ritmo.
 5. Consulte-se Crahay (2000) para uma descrição mais detalhada. Esclarece-se apenas que, nesse sistema, os estudantes são levados a progredir em seu próprio ritmo mediante um conjunto de materiais didáticos individualizados.

(+0,68). Por sua vez, 13 estudos apresentam dados que permitem examinar em que medida os efeitos da AD variam segundo as aptidões iniciais dos alunos: quatro entre eles adotam a AD coletiva; os outros nove adotam o Plano Keller. O exame atento desses estudos permite uma constatação crucial: a abordagem coletiva da AD caracteriza-se por um efeito particularmente sensível sobre os alunos fracos (+1,09), sem que os alunos fortes sejam prejudicados, pois, com esse tipo de alunos, a amplitude do efeito permanece amplamente positiva (+0,47). Os benefícios da abordagem individualizada da AD são mais uniformes: a amplitude média do efeito é de +0,36 para os alunos fortes e de +0,40 para os alunos fracos. Concluindo, se nos remetermos à avaliação das experiências AD, somos levados a relativizar a hipótese resultante da reflexão de Birzéa (1982): as abordagens individuais da AD não aumentam necessariamente as diferenças entre os fortes e os fracos, tampouco as reduzem. Ao contrário, da avaliação ressalta claramente que as abordagens coletivas da AD reduzem essas diferenças. Além disso, constata-se a superioridade geral dos efeitos engendrados por esse tipo de abordagem pedagógica em relação às experiências nas quais o ensino coletivo é totalmente abandonado.

Essa demonstração é particularmente instrutiva. O Plano Keller compartilha, com a AD, a perspectiva igualitária e também vários princípios pedagógicos: definição precisa dos objetivos, segmentação da aprendizagem em unidades ou módulos e regulação baseada na avaliação formativa. A comparação dessas duas categorias de experiência pedagógica (Plano Keller ou AD), nos aproxima da comparação experimental estrita, uma vez que nos encontramos em presença de dispositivos que compartilham numerosos parâmetros, mas se distinguem claramente quanto à dimensão individualização contraposta à de ensino coletivo. A partir dessa análise, somos tentados a concluir quanto à necessidade de uma fase coletiva no processo de ensino, podendo seguir-se momentos de individualização como é o caso na AD clássica, uma vez que os dados rapidamente expostos aqui devem ser interpretados de um modo relativo. Não seria correto apoiar-se neles para afirmar a ineficácia das experiências pedagógicas que privilegiam a individualização dos dispositivos de aprendizagem ou para reivindicar o retorno radical ao ensino tradicional. Neste caso, a pesquisa recomenda uma articulação do ensino coletivo e de fases de individualização.

PREVENIR OU CORRIGIR ERROS DOS ALUNOS?

De uma maneira ou de outra, o pensamento pedagógico parece repousar sobre um postulado segundo o qual o erro é, sob certas condições, formativo. Aderem a esse axioma aqueles que adotam o construtivismo (Huberman, 1988), mas também aqueles que acionam a noção vigotskiana de zona proximal de desenvolvimento. O mesmo se dá com a didática da Matemática e das ciências, cujos seguidores, sustentam de acordo com Bachelard, que é preciso apresentar a questão do conhecimento científico em termos de obstáculos (Martinand, 1986).

O axioma do caráter formativo do erro está igualmente no centro da noção de avaliação formativa, originada da teoria da aprendizagem por domínio. Essa teoria postula que todo processo de aprendizagem implica o funcionamento de um mecanismo de regulação (Allal, 1991) pelo qual o sujeito procede ao ajuste da competência ou da compreensão que ele constrói. A avaliação formativa e as atividades de fixação ou reforço que dela resultam têm por objetivo oferecer aos alunos as oportunidades de regulação que lhes são necessárias. Esse complemento de formação será tanto mais adaptado às dificuldades do aluno e portanto mais eficaz quanto mais o teste formativo tiver um caráter de diagnóstico, quer dizer, se ele permitir reconhecer os obstáculos que o aprendiz não conseguiu superar. Como observa Allal (1988, p. 96), a concepção de Bloom da aprendizagem por domínio “certamente se apóia sobre o postulado de um duplo mecanismo de regulação situado por um lado no nível da ação de aprendizagem do aluno e, por outro, no nível da ação pedagógica do professor”.

Apenas Skinner, com sua concepção da aprendizagem sem erro ousou contestar esse dogma (Crahay, 1999). É nessa perspectiva iconoclasta que ele concebeu o ensino programado. O método visa apresentar a cada sujeito individualmente uma matéria recortada em muitos segmentos sucessivos, a suscitar uma resposta para cada segmento de conteúdo e fazer cada resposta ser sucedida por um *feedback* ou informação que retroaja sobre a resposta. Se a seqüência de aprendizagem for corretamente programada, a porcentagem de erros não deve ultrapassar 5%. Quer dizer, nessas condições, o aluno pode se apropriar daquilo que for essencial para as competências consideradas necessárias para sua inserção social, cometendo um mínimo de erros.

Parece que é hora de falar de dogma tendo em vista as respostas suscitadas pelas reações aos resultados obtidos nos Estados Unidos ao final do importante Beginning Teacher Evaluation Study [Estudo Avaliativo do Professor Iniciante], ou BTES. No início, Fisher, Berliner, Filby, Marlinave, Cahen e Dishaw, os criadores desse vasto programa de pesquisa⁶, influenciados por uma corrente de idéias que se alimentou da taxonomia de Bloom, se lançaram com um *a priori* favorável àquilo que do outro lado do Atlântico denomina-se atividades de alto nível taxonômico (Fisher et al., 1978). Estas dever-se-iam traduzir por uma taxa média de erro (*grosso modo*, tanto erros quanto acertos). Mais precisamente, esses pesquisadores formularam a hipótese segundo a qual a probabilidade de que um aluno aprenda certa competência, depende do tempo investido especificamente nas atividades escolares que tem a ver com aquela competência, mas sob condição que essas atividades conduzam à superação do nível de competência atual do aluno. Essa condição dever-se-ia traduzir por uma taxa relativamente alta de erros, sinal que o indivíduo estava precisamente em fase de elaboração de um novo saber ou *savoir-faire*⁷.

Ao contrário do que os autores supunham no início do estudo, um grau reduzido de erros durante a atividade escolar consagrada a um determinado conteúdo acadêmico, indica uma aprendizagem de grande amplitude. Ou seja, a organização freqüente pelo professor de tarefas à altura dos alunos se traduz por um pequeno índice de erros e favorece a aprendizagem da competência almejada⁸.

-
6. O objetivo desse programa era determinar as competências que um professor iniciante deveria idealmente dominar para mostrar um bom desempenho desde seu início na carreira docente. Os principais pesquisadores envolvidos nessa pesquisa se propunham a constituir uma amostra de professores experientes, a identificar entre eles os mais e os menos eficazes (ou seja, aqueles que proporcionavam aos alunos aprendizagens importantes *versus* pouco importantes) e a observar como todos eles administravam suas aulas. A contribuição desse programa de pesquisa ultrapassa a apresentação que é feita aqui, em que se levam em consideração apenas as taxas de erro dos alunos .
 7. Para dar conta dessa hipótese, um conceito novo é proposto: o de *academic learning time* [tempo de aprendizagem acadêmica] ALT, que remete a um tempo consagrado a uma aprendizagem escolar específica durante a qual as taxas de erro dos alunos flutuam entre 25% e 50%.
 8. ALT corresponde, em definitivo, ao tempo durante o qual os alunos desenvolveram uma tarefa de aprendizagem cujos objetivos coincidem com questões inseridas na prova de avaliação e cujo grau de dificuldade permite a esses alunos produzir um máximo (90%) de boas respostas.

Essa constatação, que vai na contracorrente do ponto de vista das teorias pedagógicas dominantes, mas no sentido da teoria de Skinner, foi discutida de maneira particularmente interessante por Shulman (1986). Segundo esse pesquisador, a definição da qualidade do ensino pelo número de erros em situação de aprendizagem traz problemas. Não se trata de negligenciar o papel do erro na aprendizagem, mas de se perguntar sobre o estatuto realmente causal de uma taxa elevada de acertos produzidos durante o ensino. Para interpretar os resultados da BTES, ele levanta a seguinte hipótese alternativa: o aluno que comete muitos erros não aprende menos porque comete erros; seus erros são sinal de que ele aprende menos. Esse raciocínio repousa essencialmente sobre duas constatações: 1. as questões feitas pelos professores em classe são próximas das questões pertencentes ao universo das questões incluídas nos testes; 2. os professores ensinam ou explicam pouco. Em consequência, as perguntas do professor, por ocasião das discussões ou dos exercícios individuais, já são elementos de avaliação e a taxa de respostas corretas observadas reflete aquilo que o aluno comprehendeu da explicação mais ou menos sucinta do professor. A observação de que uma taxa elevada de erros está associada a baixo desempenho no teste é só uma tautologia, pois o preditor mede a mesma coisa que a variável a predizer⁹.

A controvérsia ainda não foi resolvida; até onde sabemos, a questão não tem sido objeto de pesquisas empíricas após aquelas da BTES. No entanto, o desafio é capital: trata-se essencialmente de determinar em que medida é pedagogicamente preferível prevenir os erros dos alunos ou, ao contrário, ajudá-los a se corrigir.

É no contexto das pesquisas sobre a eficácia da AD que encontramos um segundo conjunto de dados empíricos sobre os efeitos das diferentes modalidades de tratar os erros dos alunos. Para explicar em detalhe essas pesquisas (apresentadas em Crahay, 2000¹⁰), é importante, antes de tudo, lembrar que Bloom (1979) postula que "a maioria dos alunos é capaz de realizar aprendizagens de nível elevado se o ensino for elevado e se os alunos forem ajudados

9. Para uma apresentação e discussão mais aprofundadas do BTES, consultar Delhaxhe (1997).

10. Lembremos simplesmente que a eficácia da AD é amplamente demonstrada tanto para alunos fracos quanto para alunos fortes, tendendo, por um suplemento mínimo de tempo, a se reduzir a distância entre os dois.

QUADRO 3

| | | | | | |
|---------------------|------------------------|---|----------------------------|---|------------------------|
| <u>Unidade</u> 1 | Ensino coletivo | → | Teste formativo + feedback | ↔ | Procedimento corretivo |
| <u>Unidade</u> 2 | Ensino coletivo | → | Teste formativo + feedback | ↔ | Procedimento corretivo |
| <u>Unidade</u> 3 | Ensino coletivo | → | Teste formativo + feedback | ↔ | Procedimento corretivo |

quando e onde encontram dificuldades, se lhes for dado tempo para atingir o domínio e se existirem critérios claros do que seja esse domínio" (p.15-16). Aceitando essencialmente os programas escolares tais como se apresentam, a AD se presta a uma economia bem simples: a matéria a ser assimilada é dividida em unidades de aprendizagem; cada uma é ensinada coletivamente, mas, no fim de cada unidade, os alunos são submetidos a um teste formativo e se beneficiam, se não forem bem-sucedidos, de procedimentos corretivos antes de passarem por um segundo teste formativo. Em princípio, o professor não avança para a unidade seguinte sem que a unidade em estudo tenha sido dominada por todos. Essa pedagogia pode ser esquematizada da seguinte forma:

Nesse esquema, mesmo conservando fixa a variável "ensino coletivo", é possível fazer variar os parâmetros seguintes:

- a. número de tentativas aceitas para atingir o nível do domínio;
- b. presença de procedimentos corretivos;
- c. forma dos testes formativos utilizados (itens de escolha múltipla ou itens com respostas dissertativas);
- d. forma dos testes administrados por ocasião das tentativas sucessivas;
- e. número de testes propostos aos alunos;
- f. número de *feedbacks* fornecidos aos alunos referentes a seu desempenho nos testes formativos;
- g. nível de domínio exigido para ser bem-sucedido no teste formativo.

Kulik e Kulik (1987) conseguiram reunir 49 estudos de natureza quase-experimental sobre o efeito desses diferentes parâmetros e aplicaram a esses estudos o cálculo da amplitude do efeito. Os resultados dessa meta-análise são apresentados no quadro 4.

QUADRO 4
AMPLITUDE DO EFEITO LIGADA A DIFERENTES
ASPECTOS DOS TESTES FORMATIVOS

| Aspectos dos testes formativos | | Número de estudos | Amplitude do efeito |
|--|-------------------------|-------------------|---------------------|
| Número de tentativas aceitas para atingir o nível de domínio | Uma ou duas | 24 | + 0,53 |
| | Três | 10 | + 0,52 |
| | Quatro | 14 | + 0,58 |
| Presença de procedimentos corretivos | Sim | 29 | + 0,54 |
| | Não | 20 | + 0,54 |
| Forma dos testes | Múltipla escolha | 44 | + 0,54 |
| | Respostas dissertativas | 4 | + 0,69 |
| Forma dos testes durante avaliações sucessivas | Mesma forma | 6 | + 0,51 |
| | Forma diferente | 36 | + 0,50 |
| Número de testes nos grupos (GE & GC) | Mesmo número | 31 | + 0,48 |
| | Número diferente | 18 | + 0,65 |
| Freqüência dos feedbacks referente aos testes | Idêntica | 21 | + 0,36 |
| | Diferente | 28 | + 0,67 |
| Porcentagem de acertos exigidos como nível de domínio | Entre 70 e 80 | 17 | + 0,38 |
| | Entre 81 e 90 | 15 | + 0,51 |
| | Entre 91 e 100 | 17 | + 0,73 |

Fonte: Kulik, Kulik, 1987.

As amplitudes do efeito são elevadas, o que atesta a eficácia geral das experiências de AD, interessando portanto aqui as diferenças entre os valores. Dois resultados não parecem surpreender: a forma dos testes bem como sua similitude por ocasião de avaliações sucessivas não influencia a amplitude do efeito (análise de variância não significativa). Ao contrário, é surpreendente constatar que o número de tentativas aceitas para atingir o nível de domínio não constitui um parâmetro muito influente. Esse resultado é, de qualquer modo, interessante: sugere que não se deve insistir com o aluno que não atinge o nível de domínio após duas tentativas.

Uma constatação não deixa de surpreender: a ausência de procedimentos corretivos não é prejudicial para a eficácia da AD. Essa observação é crucial; ela explica claramente porque é possível implantar uma prática de AD eficaz e

pouco dispendiosa quanto ao tempo. Ela abala, contudo, uma idéia-chave da teoria de Bloom, que recomenda, há muito, um afinamento do procedimento de regulação das aprendizagens dos alunos. No mesmo espírito, Allal (1988), Cardinet (1988) e Rieben (1988) formularam, com algumas pequenas diferenças, a hipótese de que as práticas de AD teriam muito a ganhar com um diagnóstico mais preciso das dificuldades dos alunos e de procedimentos de regulação estritamente voltados para essas dificuldades.

A análise de variância indica que dois parâmetros fazem essencialmente a diferença: o grau de exigência do nível de domínio e a freqüência dos *feedbacks* fornecidos aos alunos sobre seu desempenho nos testes formativos. É necessário procurar definir os efeitos de suas interações. O quadro 3 contribui para isso. Nele os estudos foram categorizados segundo dois critérios:

- **grau de exigência do nível de domínio:**
- estudos nos quais esse nível é igual ou inferior a 90%;
- estudos nos quais é igual ou superior a 91%;
- **freqüência dos *feedbacks* fornecidos ao aluno após o teste formativo:**
- estudos nos quais esse parâmetro foi controlado de modo a que permaneça igual nos dois grupos;
- estudos nos quais esse parâmetro não foi controlado e provavelmente variou segundo os grupos.

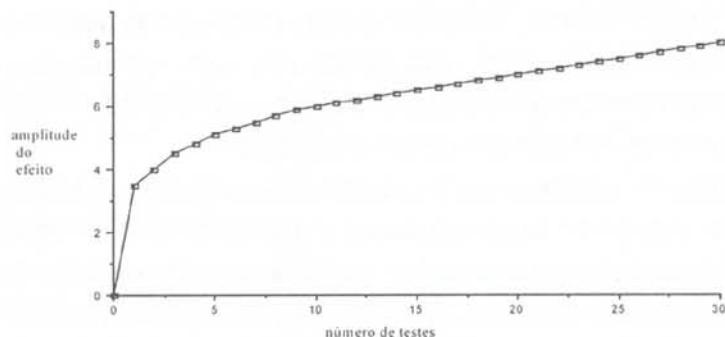
Quando os alunos das classes de AD e aqueles das classes de controle recebem o mesmo número de *feedbacks* sobre seu desempenho nos testes

QUADRO 5
AMPLITUDES DO EFEITO LIGADAS A DIFERENTES
ASPECTOS DOS TESTES FORMATIVOS

| Facetas dos testes formativos | Baixo grau de exigência | | Grau de exigência elevado | | Total | |
|---|-------------------------|---------------------|---------------------------|---------------------|-------------------|---------------------|
| | Número de estudos | Amplitude do efeito | Número de estudos | Amplitude do efeito | Número de estudos | Amplitude do efeito |
| Freqüência idêntica de <i>feedbacks</i> | 15 | + 0,26 | 6 | + 0,63 | 21 | + 0,36 |
| Freqüência desigual de <i>feedbacks</i> | 17 | + 0,60 | 11 | + 0,79 | 28 | + 0,67 |
| Total | 32 | + 0,44 | 17 | + 0,73 | 49 | + 0,54 |

Fonte: Kulik, Kulik, 1987.

GRÁFICO I
AMPLITUDE DO EFEITO ESPERADO EM FUNÇÃO DO
CRESCIMENTO DO NÚMERO DE TESTES FORMATIVOS



Fonte: Bangert-Drowns, Kulik, Kulik, 1991.

formativos (e provavelmente o mesmo número de testes) e o nível de domínio é inferior a 91% de acertos, a amplitude do efeito é de um quarto de sigma (0,26). Porém, quando os professores das classes de controle não se preocupam em fornecer tantos *feedbacks* a seus alunos quanto aqueles das classes de AD, a distância aumenta: passa de +0,26 para +0,60. Ela aumenta numa proporção equivalente quando se aumenta o grau de exigência do nível de domínio: passa então de +0,26 a +0,63. Quando se combinam esses dois parâmetros, vê-se a amplitude do efeito subir a +0,79.

Numa outra meta-análise referente à aplicação freqüente de testes após o ensino¹¹, Bangert-Drowns, Kulik e Kulik (1991) conseguiram reunir 35 estudos, também de natureza quase-experimental, permitindo apreender o efeito específico desse parâmetro. Utilizando um modelo estatístico desenvolvido por Glass et al. (1982), Bangert-Drowns, Kulik e Kulik (1991) se empenharam em estabelecer uma relação de tipo logarítmico entre o número de testes propostos

11. Desde 1910 os pesquisadores têm afirmado que a simples realização de um teste logo após o ensino (independentemente de todo *feedback* e de todo procedimento corretivo) favorece a memorização das informações que foram transmitidas (Jones, 1923; Turney, 1931; Spitzer, 1939). Essa hipótese foi abandonada entre 1940 e 1960, para retomar força com o advento do cognitivismo. Outros estudos sobre esse tema reaparecem, a partir dos anos 60 (Mc Keachie, 1963; Mawhinney et al., 1971; Proger, Mann, 1973; Richards, 1979).

aos estudantes e a amplitude do efeito. O resultado, conforme o gráfico, é instrutivo. O modelo permite estimar que a amplitude do efeito é de +0,41 quando se passa de nenhum teste a dois testes por período de 15 semanas; ela sobe a +0,49 quando se passa de nenhum teste a quatro testes por período. Isso permite concluir que a aplicação regular de testes aos alunos favorece a aprendizagem, mas os ganhos não aumentam na mesma proporção do número de testes aplicados.

CONCLUSÕES E PERSPECTIVAS

A repetência não ajuda os alunos em dificuldade escolar a superar aquilo que atrapalha seu desenvolvimento. Ao contrário, as pesquisas indicam com insistência que a promoção é geralmente preferível à repetição de um ano. Isso não significa que a promoção automática seja uma panacéia pedagógica. Mais precisamente, a pesquisa sinaliza aos professores que a repetência não traz nenhuma solução para o difícil problema da gestão das dificuldades de aprendizagem daqueles alunos considerados fracos. A saída para o problema não é essa.

A pesquisa denuncia igualmente a ilusão de se acreditar na eficácia das classes homogêneas quanto ao rendimento. Quando se obtém um resultado, em geral é em detrimento dos alunos fracos e é negativo pelo simples fato de ser resultante da resignação dos professores. A variável "composição das classes" é em si mesma inoperante.

Por sua vez, a redução do tamanho da classe e o recurso aos grupos com necessidades específicas de aprendizagem contribuem positivamente para o aprendizado dos alunos fracos.

Há muito a ser feito em matéria de pesquisa para determinar como reagir de modo adequado aos erros dos alunos. Com base nos dados recolhidos pelo BTES, não se pode excluir a hipótese segundo a qual é melhor prevenir que remediar. A revisão literatura acadêmica referente à pesquisa sobre experiências de AD sugere que a maioria dos alunos são capazes de regular seu aprendizado sem que seja necessário recorrer a procedimentos corretivos. Parece que é suficiente oferecer-lhes a oportunidade de se submeter a testes regulares, fornecendo-lhes *feedbacks* precisos sobre seu desempenho nesses testes.

É importante, contudo, determinar o alcance preciso dos resultados que foram analisados nestas páginas. As meta-análises reposam, é bom lembrar, sobre uma série de estudos realizados segundo o modelo grupo experimental *versus* grupo de controle, chamado esquema GE – GC. Ademais, o cálculo da amplitude do efeito utiliza essencialmente as médias obtidas num e outro grupo. A eficácia de que falamos aqui, corresponde, portanto, a uma medida tomada sobre o conjunto do grupo. Ou seja, é preciso interpretar os resultados das meta-análises como indicações sobre o que mais convém ao conjunto dos alunos. Pois, infelizmente, os estudos que abordam o efeito de parâmetros determinados (o nível de domínio estipulado, o número de testes formativos, o uso de procedimentos corretivos etc.) não fornecem estimativa diferencial segundo o nível dos alunos (fortes, médios e fracos). Ou então, e esta é a questão principal: o que convém ao conjunto dos alunos é suficiente para aqueles que se encontram em dificuldade escolar? Isso não quer dizer que não é preciso modular certos aspectos da AD tendo em vista esses alunos. Nada impede de pensar que é necessário diferenciar o nível de domínio em função do nível dos alunos. Da mesma forma, é possível perguntar se não seria o caso de propor mais testes formativos aos alunos que sentem mais dificuldade e, quando eles persistirem em tropeçar em certas dificuldades, propor-lhes procedimentos corretivos. Pois, como já discutimos antes (Crahay, 2000), o fato de a eficácia média das experiências de AD não ser afetada pela ausência de atividade de reforço, significa provavelmente que a maioria dos alunos pode-se regular pelos *feedbacks* recebidos sobre seu desempenho nos testes formativos. Não vale a pena impor explicações ou exercícios suplementares nos casos em que a auto-regulação dos alunos pode ser suficiente. Entretanto, saber se é preciso recorrer a tal procedimento quando se constata que certos alunos estão longe de alcançar o nível de domínio na segunda aplicação do teste formativo já é uma outra questão.

É preciso, a nosso ver, proceder a uma recentralização das pesquisas. De modo geral, as pesquisas indicam que a maioria dos alunos nada ganha com o afinamento das atividades de reforço ou com as pesadas estratégias de individualização. Enquanto a tradição pedagógica francesa postula, desde Claparède (1920), que uma intervenção pedagógica será tão mais eficaz quanto mais for regulada pelo processo individual de aprendizagem, a literatura anglo-saxônica parece contestar esse modo de pensar. É provável que a verdade se encontre

na nuança e no meio-termo. Talvez não seja produtivo desenvolver, em todos os casos, pesados procedimentos de regulação da aprendizagem. É possível supor que os obstáculos mais freqüentes à aprendizagem da maioria dos alunos podem ser superados com a ajuda de procedimentos leves. É provável que um diagnóstico preciso dos problemas com que se deparam os alunos que sofrem verdadeiros bloqueios seja a única saída formativa do ponto de vista pedagógico. Mas esse argumento ainda precisa de demonstração empírica.

REFERÊNCIAS BIBLIOGRÁFICAS

- ALLAL, L. Vers un élargissement de la pédagogie de maîtrise: processus de régulation interactive, rétroactive et proactive. In: HUBERMAN, M. (ed.) *Assurer la réussite des apprentissages scolaire ? Les propositions de la pédagogie de maîtrise*. Neuchâtel: Delachaux & Niestlé, 1988. p. 86-126.
- _____. *Vers une pratique de l'évaluation formative*. Bruxelles: De Boeck, 1991.
- ASTOLFI, J. P.; DEVELAY, M. *La Didactique des sciences*, Paris: PUF, 1991. (Coll.: Que sais-je? 2448)
- BACHELARD, G. *La Formation de l'esprit scientifique*. Paris: Vrin, 1938.
- BANGERT-DROWNS, R. L.; KULIK, J. A.; KULIK, C.-L. C. Effects of frequent classroom testing. *Journal of Education Research*, v.85, n.2, p.89-99, nov./dec. 1991.
- _____. Individualized suystems of instruction in secondary schools. Bruxelles: De Boeck, 1983.
- BIRZEA, C. *La Pédagogie du succès*. Paris: Presses Universitaires de France, 1982.
- BLESS, G.; BONVIN, P.; SCHÜPBACH, M. *Le Redoublement scolaire: ses déterminants, son efficacité, ses conséquences*. Berne: Haupt, 2005.
- BLOOM, B. S. *Caractéristiques individuelles et apprentissages scolaires*. Bruxelles: Labor; Paris: Nathan, 1979.
- BURNS, R. B. How time is used in elementary schools: the activity structure of classrooms. In: ANDERSON, L. W. (ed.) *Time and school learning: theory, research and practice*, London, Canberra: Croom Helm, 1984. p.91-128.
- CARDINET, J. Un point de vue constructiviste sur la pédagogie de maîtrise. In: HUBERMAN, M. (ed.) *Assurer la réussite des apprentissages scolaires? Les propositions de la pédagogie de maîtrise*. Neuchâtel: Delachaux & Niestlé, 1988. p.155-196.

- CLAPARÈDE, E. *L'École sur mesure*. Lausanne: Payot, 1920.
- CRAHAY, M. *L'École peut-elle être juste et efficace? De l'égalité des chances à l'égalité des acquis*. Bruxelles: De Boeck, 2000. (Coll.: Pédagogies en développement)
- _____. *Peut-on lutter contre l'échec scolaire ? 2^e éd. rev. et aug.* Bruxelles: De Boeck, 2003.
- _____. *Psychologie de l'éducation (1^{er} cycle)*. Paris: Presses Universitaires de France, 1999.
- DELHAXHE, A. Le Temps comme unité d'analyse dans la recherche sur l'enseignement, *Revue Française de Pédagogie*, n.118, p.107-125, 1997.
- DUPRIEZ, V.; DRAELANTS, H. Classes homogènes versus classes hétérogènes: les apports de la recherche à l'analyse de la problématique. *Revue Française de Pédagogie*, n.148, p.145-165, 2004.
- FISHER, C. W. et al. *Teaching and learning in elementary schools: a summary of the BTES*. San Francisco: Far West Regional Laboratory for Educational Research and Development, 1978.
- FLOYD, C. Meeting children's reading needs in the middle grades: a preliminary report. *Elementary School Journal*, n.55, p.99-103, 1954.
- GLASS, G. V. et al. *School class size: research and policy*. Beverly Hills (Ca): Sage Publication, 1982.
- HILSON, M. et al. A Controlled experiment evaluating the effects of a nongraded organization on pupil achievement. *Journal of Educational Research*, n.57, p.548-550, 1964.
- HOLMES, C. T. Grade level retention effects: a meta-analysis of research studies. In: SHEPARD, L. A.; SMITH, M. L. (eds.) *Flunking grades: research and policies on retention*. Bristol: Falmer Press, 1989. p. 16-33.
- HOLMES, C. T.; MATTHEWS, K. M. The Effects of nonpromotion on elementary and junior high school pupils: a meta-analysis. *Review of Educational Research*, v.54, n.2, p.225-236, 1984.
- HUBERMAN, M. (ed.) *Assurer la réussite des apprentissages scolaires? Textes de base en pédagogie*. Neuchâtel: Delachaux & Niestlé, 1988.
- JACKSON, G. B. The Research evidence on the effects of grade retention. *Review of Educational Research*, v.45, n.4, p.613-635, 1975.
- JIMERSON, S. R. Meta-analysis of grade retention research: implications for practice in the 21st Century. *School Psychology Review*, v.30, n.3, p.420-437, 2001.

- JONES, H. E. Experimental studies of college teaching: the effect of examination on permanence of learning. *Archives of Psychology*, n.10, p.1-70, 1923.
- KULIK, J.A.; KULIK, C.-L. Effects of accelerated instruction on students. *Review of Educational Research*, n.54, p.409-425, 1984.
- _____. Mastery testing and student learning: a meta-analysis. *Journal of Educational Technology Systems*, n.15, p.325-345, 1987.
- KULIK, C.-L.C.; KULIK, J.A.; BANGERT-DROWNS, R. L. Effectiveness of mastery learning programs: a meta-analysis. *Review of Educational Research*, n.60, p.265-299, 1990.
- KULP, D. H. Weekly tests for graduate students? *School and Society*, n.38, p.157-159, 1933.
- MARTINAND, J. L. *Connaitre et transformer la matière*. Berne: Peter Lang, 1986.
- MAWHINNEY, V. T. et al. A Comparison of students studying-behavior produced by daily, weekly, and three-week testing schedules. *Journal of Applied Behavior Analysis*, n.4, p.257-264, 1971.
- MC KEACHIE, W. J. Research on teaching at the college and university level. In: GAGE, N. L. (ed.) *Handbook of Research on Teaching*. Chicago: Rand McNally, 1963. p.1.118-1.172.
- MOLNAR, A. *Smaller classes and educational vouchers: a research update*. Harrisburg: Keystone Research Center, 1999.
- MORGAN, E. F.; STUCKER, G. R. The Joplin Plan of reading vs. a traditional method. *Journal of Educational Psychology*, n.51, p.69-73, 1960.
- OAKES, J. Can tracking research inform practice? Technical, normative and political considerations. *Educational Researcher*, v.21, n.4, p.12-21, may 1992.
- PIKETTY, T. *L'impact de la taille des classes et de la ségrégation sociale sur la réussite scolaire dans les écoles françaises: une estimation à partir du panel primaire 1997*. Paris-Jourdan: ENS-EHESS, 2004.
- PROGER, B. B.; MANN, L. *An historical review of theoretical and experimental literature on the teaching values of informal (non-standardized), teacher-made achievement tests: 1913-1968*. Blue Bell (PA): Montgomery County Intermediate Unit 23, 1973. (ERIC Document Reproduction Service n. ED 084 292)
- RICHARDS, J. Adjunct postquestions in text: a critical review of methods and processes. *Review of Educational Research*, n.49, p.181-196, 1979.
- RIEBEN, L. Un point de vue constructiviste sur la pédagogie de maîtrise. In: HUBERMAN, M. (ed.) *Maîtriser les processus d'apprentissage: fondements et perspectives de la pedagogie*. Neuchâtel: Delachaux & Niestlé, 1988. p.127-154.

- SHULMAN, L. S. Paradigms and research programs in the study of teaching: a contemporary perspective. In: WITTRICK, M. C. (ed.) *The Third handbook of the research on teaching*. New York: MacMillan, 1986. p.2-35.
- SLAVIN, R. E. Ability grouping and student achievement in elementary schools: a best-evidence synthesis. *Review of Educational Research*, v.57, n.3, p.293-336, 1987.
- _____. Achievement effects of ability grouping in secondary schools: a best evidence synthesis. *Review of Educational Research*, v.60, n.3, p.471-499, 1990.
- _____. Best-evidence synthesis: an alternative to meta-analytic and traditional reviews. *Educational Researcher*, v.15, n.9, p.5-11, 1986.
- _____. Mastery learning reconsidered. *Review of Educational Research*, n.57, p.175-213, 1987a.
- _____. Mastery learning re-reconsidered. *Review of Educational Research*, n.60, p.300-302, 1990a.
- _____. *Preventing school failure: research, policy and practice*. Boston: Allyn and Bacon, 1994.
- SHIMRON, J. S. Learning activities in individually prescribed instruction. *Instructional Science*, v.5, n.4, p.391-401, 1976.
- SNOW, R. Aptitude-treatment interaction models. In: HUSEN, T.; POSTLETHWAITE, T. N. (eds.) *The International encyclopedia of education*. Oxford: Pergamon, 1985. p. 301-305.
- SPITZER, H. F. Studies in retention. *Journal of Educational Psychology*, n.30, p.641-656, 1939.
- TRONCIN, T. *Le Redoublement: radiographie d'une décision à la recherche de sa légitimité*. Dijon, 2005. Thèse (doct.) Université de Dijon.
- TURNEY, A.H. The Effect of frequent short objective tests upon the achievement of college students in educational psychology. *School and Society*, n.33, p.760-762, 1931.

Recebido em: dezembro 2005
Aprovado para publicação em: abril 2006