

É POSSÍVEL TIRAR CONCLUSÕES SOBRE OS EFEITOS DA REPETÊNCIA?

MARCEL CRAHAY

Faculdade de Psicologia e de Ciências da Educação/Universidades de Liège e de Genebra
marcel.crahay@ulg.ac.be

Tradução: Neide Luzia de Rezende

RESUMO

Desde o início do século XX, os pesquisadores da educação se esforçam para apreender com rigor os efeitos da repetência. Os estudos feitos a esse respeito tiveram uma evolução importante, marcada pela melhora dos dispositivos e por uma sofisticação cada vez maior das técnicas de análise estatística. O objetivo desses trabalhos é circunscrever melhor o verdadeiro efeito da repetência, o que implica uma reflexão metodológica sobre a origem dos vieses e os métodos mais apropriados para controlá-los. Este artigo retoma essa evolução metodológica, comentando os principais resultados das pesquisas.

REPETÊNCIA – FRACASSO ESCOLAR – PROMOÇÃO AUTOMÁTICA – PESQUISAS QUASE-EXPERIMENTAIS

ABSTRACT

IS IT POSSIBLE TO COME TO CONCLUSIONS ON THE EFFECTS OF STUDENT REPETITION? Since the early 20th century, researchers on education have striven to fully understand the effects of student retention. The studies led to an important evolution characterized by the improvement of resources and by a greater sophistication of statistical analysis techniques. The purpose of these studies is to know better the true effect of student retention, which means a methodological reflection on the origin of biases and the most appropriate methods to control them. This article resumes this methodological evolution and comments on the main results of researches.

RETENTION – GRADE REPETITION – GRADE AUTOMATIC PASSING – QUASI EXPERIMENTAL RESEARCH

Este artigo foi publicado originalmente na *Revue Française de Pédagogie*, n.148, jul./set. 2004, com o título: "Peut-on conclure à propos des effets du redoublement?".

REPETÊNCIA, OPINIÃO PÚBLICA E ESTUDOS COMPARATIVOS DO RENDIMENTO ESCOLAR

A opinião pública foi recentemente abalada pelos resultados da pesquisa do Programa Internacional de Avaliação de Alunos – Pisa, 2000 – da qual participaram 31 países. Com frequência, a imprensa destaca o escore nacional, se entusiasmando ou se lamentando com as competências em leitura dos alunos segundo a posição do país na classificação internacional. Mas os estudos comparativos do rendimento têm felizmente outras virtudes além de afaçar ou ferir o orgulho das nações. Eles levam a pensar sobre os fatores que afetam o rendimento dos sistemas educativos.

Em razão disso, surpreendem-se, tanto os professores e os pais quanto os gestores políticos, com as boas performances dos países que, a exemplo da Finlândia, efetivamente abandonaram a prática da repetência.

O próprio fato de países como Dinamarca, Finlândia¹, Noruega, Suécia e Reino Unido, e também Coreia, Malásia, Japão, Nova Guiné etc. terem abolido a possibilidade de fazer repetir o aluno é já objeto de surpresa para muitos. Essa informação refuta por si mesma a afirmação de certos professores que proclamam de maneira peremptória que “a não-repetência é impossível” ou os autores que escrevem que “a escola tem necessidade do fracasso escolar” (Milner, 1984). Evidentemente, alguns sistemas educativos funcionam sem fracasso e, sobretudo, sem repetência. Mas é preciso saber se eles funcionam bem, uma vez que para um bom número de professores, pais e políticos, sob a idéia de impedir a repetência, se esconde o espectro do baixo nível do ensino.

Percebemos de imediato a utilização que se pode fazer dos estudos comparativos do rendimento escolar quanto à questão que preocupa aqueles que receiam os efeitos perversos da supressão da repetência: se é verdade que a abolição da repetência provoca automaticamente a queda do nível², os países

-
1. A repetência não é legalmente proibida na Finlândia, mas não é mais praticada. Nas páginas que se seguem, consideramos esse país como um sistema de promoção automática.
 2. Esse medo está presente entre os professores da Communauté Française de Belgique – CFBW. Portanto, na pesquisa dirigida por Stegen (1994) junto a 263 professores nessa comunidade, 62% dentre eles estimam que a supressão da repetência no início do curso secundário, preconizada pelo Ministério da Educação à época, provocará um nivelamento por baixo.

que fizeram a opção pela promoção automática só podem então se caracterizar por rendimentos medíocres. Para verificar a validade dessa asserção, basta comparar o rendimento dos países que conservaram a prática da repetência e aquele dos países que a abandonaram.

Os dados do Pisa constituem um bom material para proceder a esse tipo de comparação. Cada país testou uma amostra representativa de alunos de 15 anos capaz de refletir a diversidade do quadro educativo do país. Em cada país, um mínimo de 150 escolas, selecionadas segundo um procedimento aleatório, fizeram parte do estudo. No interior das escolas, 35 alunos foram designados ao acaso para compor a lista de todos os alunos de 15 anos. As mesmas regras de seleção foram aplicadas no conjunto dos países, sendo que tal aplicação foi rigorosamente controlada por um especialista em amostragem. O teste comporta questões de múltipla escolha e uma proporção importante (45%) de questões abertas para que o aluno elabore sua própria resposta. Todos os alunos de todos os países participantes passaram por provas de compreensão de textos idênticos, traduzidos de diferentes línguas, a partir da versão de duas fontes, uma em inglês e outra em francês. Os resultados da compreensão de texto são apresentados numa escala de competências hierarquizada em cinco níveis. A cada nível correspondem tarefas específicas de leitura³.

No quadro 1, concentramos a apresentação dos resultados da União Europeia, (expressos sobre uma escala normalizada cuja média internacional vale 500) em cujos países sabemos claramente se a repetência foi ou não abolida.

As tendências são claras; podemos, pois, constatar com segurança o seguinte:

- Nos países nórdicos, assim como na Irlanda e no Reino Unido, sistemas em que a promoção automática foi adotada, a porcentagem de bons leitores (nível 4 ou 5) é superior à média da OCDE (com exceção da Dinamarca, onde ela é ligeiramente inferior), enquanto a porcentagem de maus leitores (nível 1 ou abaixo) é inferior à média OCDE

3. Graças a essas escalas, pode-se, para cada um dos níveis, apreender o tipo de tarefa que os alunos são capazes de realizar com certo sucesso.

(de novo, com exceção da Dinamarca, onde essa porcentagem é igual à média OCDE).

- Dentre os países que conservaram a prática da repetência, observam-se perfis de resultados diferentes. A porcentagem de maus alunos leitores é superior à média OCDE na Alemanha, Bélgica, Itália, Grécia, em Luxemburgo e Portugal, mas esse não é o caso na Áustria, Espanha e França. Quanto à porcentagem de bons leitores, é superior à média OCDE na Áustria, Bélgica e França, mas esse não é o caso em outros países.

QUADRO I
MÉDIA NACIONAL E PORCENTAGEM DE ALUNOS
EM CADA NÍVEL DE COMPETÊNCIA NA ESCALA PISA
(ESCALA COMBINADA DE COMPREENSÃO DE TEXTO)

	Média	Nível 1 ou abaixo %	Nível 2%	Nível 3%	Nível 4 ou acima%
Alemanha	484	22,6	22,3	26,8	28,2
Áustria	507	14,6	21,7	29,9	33,7
Bélgica	507	19,0	16,8	25,8	38,3
Espanha	493	16,3	25,7	32,8	25,3
França	505	15,2	22,0	30,6	32,2
Itália	487	18,9	25,6	30,6	24,8
Grécia	474	24,4	25,9	28,1	21,7
Luxemburgo	444				
Portugal	470	26,3	25,3	27,5	21,0
Países com Promoção Automática					
Dinamarca	497	17,9	22,5	29,5	30,0
Finlândia	546	6,9	14,3	28,7	50,1
Irlanda	527	10,1	17,9	29,7	41,3
Islândia	507	14,5	22,0	30,8	32,7
Noruega	506	17,7	19,5	28,1	34,9
Reino Unido	523	12,8	19,6	27,5	40,0
Suécia	516	12,6	20,3	30,4	36,8
Média	500	17,9	21,7	28,7	31,9

Fonte: Organização para a Cooperação e o Desenvolvimento Econômico – OCDE, 2001.

O resultado dessa comparação não surpreenderá os habituais pesquisadores desse tipo de estudo. Os dados do Pisa vêm simplesmente confirmar com mais clareza as tendências que já desenhavam os estudos da International Evaluation Association – IEA –, realizados em 1970 e em 1990 (para uma síntese mais detalhada ver Crahay, 2003). Enfim, repetidamente, as comparações internacionais do rendimento em leitura (mas também em Matemática e Ciências) fornecem dados empíricos que se opõem à afirmação segundo a qual a promoção automática é acompanhada por um nivelamento por baixo. Os países que adotaram esse modo de gerir o percurso dos alunos se caracterizam por níveis satisfatórios de desempenho em leitura (mas também em Matemática e em Ciências).

ESTUDOS QUASE-EXPERIMENTAIS: EFEITOS DA REPETÊNCIA E VIESES DA AMOSTRAGEM

“Comparação não é razão”, afirma com pertinência o dito popular. Poderia se dar, portanto, que os estudos comparativos do rendimento escolar enviesassem a reflexão naquilo que concerne à repetência. Mais precisamente, bem se poderia suspeitar que a promoção automática se superponha, nos países que adotaram essa modalidade de organização, a outras qualidades e que sejam estas que determinam efetivamente o bom rendimento escolar observado nesses países. Procede, portanto, o exame de um outro *corpus* de resultados de pesquisa.

Sem dúvida, a questão dos efeitos da repetência fustiga o mundo da educação há um bom tempo. Na sua célebre revisão de 1975, Jackson apresenta uma pesquisa realizada por Keyes em 1911 (apud Jackson, 1975). Ou seja, parece que o interesse pelos efeitos da repetência já existe há quase um século sem que o problema seja considerado definitivamente resolvido. Ora, sem poder estabelecer uma contabilidade precisa, é evidente que o número de estudos sobre o assunto é impressionante: mais de uma centena, acreditamos. Mais surpreendente ainda, os resultados colhidos ao longo desses múltiplos esforços de pesquisa são substancialmente coerentes. Contudo nada foi feito: a incredulidade do mundo da prática, mas também de certos gestores políticos, parece irredutível. Porém, a fim de responder às objeções que lhes foram ou continuam sendo dirigidas, os pesquisadores se esforçaram para afinar sua

metodologia de pesquisa. Assim, pode-se considerar as pesquisas sobre os efeitos da repetência como paradigmáticas no que concerne aos obstáculos metodológicos que os pesquisadores em educação tentam contornar para aproximar um registro confiável dos efeitos de uma prática de ensino. Em razão disso, torna-se oportuno fazer novo traçado da evolução das pesquisas nesse campo.

As primeiras pesquisas sobre os efeitos da repetência (aquelas recenseadas por Jackson, 1975) foram realizadas segundo o esquema antes–pós ou pré–teste–pós–teste. Tratava-se simplesmente de medir o estado dos conhecimentos e performances de uma amostra no interior da qual foram encontrados alunos em dificuldade escolar antes e depois do ano repetido. É evidente que os instrumentos de medida (os testes) devem ser os mesmos antes e depois da repetência. Rapidamente os pesquisadores perceberam que não era possível tirar conclusões com esse dispositivo de coleta de dados. Como discute com profundidade Jackson (1975), é altamente provável observar ganhos importantes entre o pré e o pós–teste, mas é impossível determinar com segurança o fator ao qual se deve atribuir esses ganhos: à repetência, à maturidade dos alunos, a alguma mudança pedagógica e/ou psicológica ligada ao fato de se repetir um ano... É, pois, sem surpresa que o autor americano nota que os doze estudos realizados segundo esse esquema concluem em favor de uma clara evolução cognitiva dos alunos que repetiram, não importando qual seja o instrumento de medida utilizado.

Logicamente, os pesquisadores se voltaram para a metodologia experimental, cujo princípio de base consiste em comparar a evolução de dois grupos de alunos, um tendo repetido e outro não. Evidentemente a comparação só tem sentido se se puder garantir a equivalência inicial dos dois grupos de alunos. E aí reside a maior dificuldade. É claro que as pesquisas sobre a repetência referem-se unicamente a alunos que vivenciaram dificuldades escolares, mas como estabelecer com segurança que dois grupos de alunos são realmente afetados pelas mesmas características (em conseqüência, pelos mesmos problemas)? O rigor experimental requer que, a partir de uma amostra de alunos com dificuldades de aprendizagem, se proceda a uma escolha aleatória de indivíduos nos grupos: ou seja, a indicação de alunos no grupo dos repetentes ou no dos promovidos seria regulada por sorteio aleatório. Se essa técnica é aceitável no âmbito de numerosas pesquisas em medicina ou em

psicologia, o mesmo não ocorre no caso da repetência. Por deontologia, a maioria dos pesquisadores se recusou a utilizar esse procedimento, recomendado aliás por todos os manuais de metodologia. Ou seja, eles se colocaram em condições menos confortáveis no plano do rigor. Aproveitando a variabilidade de exigências dos professores, os pesquisadores buscam escolher, no interior de uma amostra de classes, os alunos com dificuldade escolar mas para os quais são tomadas diferentes decisões quanto à promoção. O que é perfeitamente possível: principalmente as pesquisas de Grisay (1984) mostraram que com um mesmo nível de competência alunos de um mesmo grau escolar poderiam ser promovidos ou, ao contrário, ser levados a repetir o ano. O princípio subjacente aos estudos quase-experimentais consiste em seguir a evolução, durante um ou mais anos, de alunos que demonstraram competências escolares equivalentes em um teste mas para os quais a decisão de aprovação foi diferente.

Consideremos, a título de ilustração, o estudo realizado nos Mesa Public Schools por Peterson, Degraice e Ayabe (1987). Em 1981, esses três pesquisadores procuraram encontrar, nas escolas do referido distrito, alunos que se havia decidido reprovar. Dentre os 171 alunos nessa situação, eles circunscreveram uma amostra de 106 repetentes, dos quais 65 deveriam repetir o primeiro ano; 26, o segundo, e 15, o terceiro. Num segundo momento, os três pesquisadores constituíram um grupo de controle por emparelhamento. O emparelhamento consiste em escolher indivíduos para constituição de pares; os indivíduos que constituem um par devem apresentar um máximo de pontos comuns. No caso do estudo de Peterson, Degraice e Ayabe, os alunos emparelhados têm a mesma idade, mesmo sexo e pertencem ao mesmo grupo étnico. O emparelhamento era ainda estabelecido com base nos dados obtidos em 1981 pelos alunos no Californian Achievement Test – CAT –, levando-se em conta os resultados de leitura, língua e Matemática. Em resumo, tudo é feito para comparar o máximo possível os dois grupos. Depois os autores acompanharam esse grupo de alunos por três anos. Mais exatamente, eles lhes aplicaram o CAT em fins de 1982, 1983 e 1984. O plano de pesquisa pode ser esquematizado como no quadro 2⁴.

4. Ressalte-se que a maioria das pesquisas sobre os efeitos da repetência comporta apenas um pós-teste.

QUADRO 2
 ESQUEMATIZAÇÃO DO PLANO DE PESQUISA
 ADOTADO NOS ESTUDOS SOBRE OS EFEITOS DA REPETÊNCIA

	Final 81	Ano Escolar 81-82	Final 82	Final 83	Final 84
Grupo Experimental	Pré-teste	Repetência	Pós-teste 1	Pós-teste 2	Pós-teste
Grupo de controle	Pré-teste	Promoção	Pós-teste 1	Pós-teste 2	Pós-teste

Com tal dispositivo, encontramos-nos próximos de uma situação experimental, sem entretanto nos inserirmos efetivamente nela. É de praxe designar as pesquisas aleatórias pela expressão “estudos quase-experimentais” na medida em que se dispõe de dois grupos supostamente equivalentes. Os indivíduos de um dos grupos, que aqui chamaremos de “experimental”, beneficiam-se de um tratamento: a repetência. O estado inicial dos indivíduos dos dois grupos (pré-teste) bem como seu estado após o tratamento (pós-teste) são medidos com a ajuda do mesmo instrumento ou de instrumentos equivalentes. Notemos, entretanto, que se nos referirmos à obra de referência de Cook e Campbell (1979), o qualificativo “quase-experimental” é dele usurpado. A quase experimentação implica, com efeito, que os pesquisadores tenham o domínio do tratamento, que deve ser aplicado a grupos naturalmente constituídos. A experimentação pura consiste num dispositivo pelo qual os pesquisadores controlam ao mesmo tempo o tratamento, a seleção dos indivíduos e sua inserção (que é aleatória) nos diferentes grupos. Nesse caso, o pesquisador não tem nem o controle do tratamento (repetência), nem o da escolha dos indivíduos que dele se beneficiam. Ele é apenas responsável pela composição do grupo de controle. Para o tipo de estudos que nós discutimos aqui, Cook e Campbell (p.298) utilizam a expressão *passive-observational studies* [estudos de observação passiva]. Poderíamos também utilizar as expressões “observação implicada” ou melhor, “pesquisa retrospectiva com esquema pré-estruturado” (Landsheere, Delchambre, 1979, p.258).

O dispositivo de pesquisa utilizado nos estudos sobre os efeitos da repetência – que nós continuaremos a chamar quase-experimental, como é de tradição no campo – não é portanto isento de vieses eventuais que poderiam afetar a validade das conclusões. Como assinala Jackson (1975), pode-se sempre temer, com tal plano de pesquisa, que diferenças latentes entre alunos repetentes e alunos promovidos não tenham sido corretamente evidenciadas

pelo(s) teste(s) de início e depois que a comparação tenha sido desviada em favor dos alunos promovidos ou em favor dos repetentes.

Esse risco é minimizado se se utilizar um procedimento experimental puro. Numa primeira etapa, uma amostra de indivíduos-alvo é selecionada seja por recrutamento aleatório, se quisermos fazer experiência com qualquer um, seja com a ajuda de uma medida inicial, se nos interessarmos por um grupo específico de indivíduos. Com a repetência, encontramos-nos no segundo caso, ou seja, convém reunir uma amostra de alunos em dificuldade escolar. Numa segunda etapa, os indivíduos são escolhidos aleatoriamente no grupo experimental ou no grupo de controle. Ao final desse procedimento, pode-se considerar como certa a equivalência dos dois grupos.

O problema é decididamente mais delicado num estudo de tipo quase-experimental (ou, mais precisamente, numa pesquisa retrospectiva com esquema pré-estruturado). O grupo experimental, que no caso irá se beneficiar da repetência, é constituído naturalmente: são os professores que escolhem os alunos para os quais consideram oportuna a medida. O pesquisador intervém somente após compor um grupo de controle cujas características ele vai tentar ajustar melhor às do grupo experimental. Os testes pelos quais passam os alunos antes da intervenção (para estabelecer seu nível inicial) e aqueles pelos quais passam durante ou após a intervenção (para medir as mudanças ou os ganhos induzidos) são um dos elementos cruciais do dispositivo. O erro de medida que inevitavelmente comportam pode, no início, mostrar equivalência entre os grupos cujas capacidades cognitivas iniciais são na realidade diferentes. Em consequência, pode-se acreditar, por ocasião do pós-teste, em diferenças de ganho, enquanto, na realidade, existe um viés de seleção dos indivíduos, ocultado pelo erro de medida do pré-teste.

A questão que se coloca para os pesquisadores preocupados em medir os efeitos da repetência é de distinguir o efeito verdadeiro dessa medida do seu efeito aparente. Cook e Campbell (1979) mostraram que o erro de interpretação dos ganhos toma a mesma direção que a diferença suspeitada no início. Mais precisamente, se a intervenção realizada em benefício dos indivíduos do grupo experimental tem um efeito positivo, tal efeito será subestimado se, em razão do erro de medida do pré-teste, não se percebeu o fato que esses sujeitos são mais “fracos” que aqueles do grupo de controle. No caso oposto, tal efeito será superestimado, ou seja, se se suspeitar que os alunos repeten-

tes são mais “fortes” que os alunos promovidos. Os riscos de sub ou superestimação vão em sentido inverso no caso de um efeito negativo. Arrisca-se uma super estimativa no caso de existirem razões para temer um erro de medida que oculte uma diferença inicial em favor dos alunos promovidos e uma subestimativa no caso oposto. A contribuição de Cook e Campbell (1979) está resumida no quadro que se segue.

QUADRO 3
RISCOS DE SUB OU SUPERESTIMATIVA DOS EFEITOS
DA REPETÊNCIA EM FUNÇÃO DO ERRO DE MEDIDA ESPERADO

Suspeita-se que a repetência		É positiva	É negativa
Teme-se que o erro de medida dissimule uma diferença	em favor dos alunos repetentes (grupo experimental)	risco de superestimativa de efeito	risco de subestimativa do efeito
	em favor dos alunos promovidos (grupo de controle)	risco de subestimativa	risco de superestimativa do efeito

No seu exame dos estudos quase-experimentais voltados para os efeitos da repetência, Jackson (1975) se esforçou por levar em conta o risco de viés de amostragem. Num primeiro momento, ele mostra resultados de todas as pesquisas sem se preocupar com um erro de medida eventual. Ele classifica os resultados desses estudos distinguindo cinco possibilidades: a. diferença significativa em favor dos alunos promovidos; b. diferença não significativa em favor dos alunos promovidos; c. ausência de diferença; d. diferença não significativa em favor dos alunos repetentes; e. diferença significativa em favor dos alunos repetentes.

No quadro 4, as cifras indicam o número de comparações relevantes de uma ou outra dessas categorias de resultados. Dois critérios são utilizados: resultados de testes de conhecimentos e resposta às escalas de ajustamento social.

Nesse quadro de resultados, as tendências são claras tanto no que concerne aos conhecimentos escolares quanto à adaptação socioafetiva. Na maioria das análises referentes ao primeiro critério, observam-se melhores resultados entre os alunos fracos que se beneficiam de uma promoção do que entre os alunos fracos que repetiram ($24 + 45 = 69$). A tendência inversa é observa-

da em 31 análises. Desde que nos limitemos às análises que resultaram em diferenças significativas no plano estatístico, o balanço é mais convincente ainda: 24 pesquisas concluem contra a repetência e duas, a seu favor. Em 67 das 104 análises referentes ao segundo critério, a tendência é favorável aos alunos promovidos. Ela é favorável aos repetentes em 37 casos sobre 104. A relação dos resultados significativos no plano estatístico é semelhante àquela observada para o primeiro critério: 27 para os promovidos contra três para os repetentes.

QUADRO 4
RESULTADOS DOS ESTUDOS DITOS QUASE-EXPERIMENTAIS COMPARANDO
ALUNOS COM DIFICULDADES, REPETENTES E PROMOVIDOS

Autores	Ano de public.	Conhecimentos escolares					Adaptação socioafetiva				
		a	b	c	d	e	a	b	c	d	e
Mac Elwee	1932	-	-	-	-	-	3	9	0	2	0
Cheyney e Boyer	1933	0	1	0	0	0	-	-	-	-	-
Arthur	1936	0	4	1	2	0	-	-	-	-	-
Anfinson	1941	1	0	2	0	0	-	-	-	-	-
Sandin	1944	-	-	-	-	-	17	3	0	0	0
Goodlab	1954	-	-	-	-	-	1	4	0	0	0
Wallihan	1955	1	0	0	0	0	-	-	-	-	-
Coeffild e Bloomers	1956	3	33	1	22	1	-	-	-	-	-
Anderson	1957	0	0	0	2	0	0	5	0	8	0
Worth	1960	2	3	0	2	1	0	2	0	7	2
Kowitz e Armstrong	1961	1	1	0	0	0	-	-	-	-	-
Kamil e Weilkart	1963	4	0	0	0	0	0	3	0	8	1
Chasky	1964	6	0	0	0	0	2	6	0	6	0
Briggs	1966	4	0	0	0	0	-	-	-	-	-
Dobbs e Neville	1967	0	1	0	1	0	-	-	-	-	-
Reinherz e Griffin	1971	0	1	0	1	0	-	-	-	-	-
White e Howard	1973	-	-	-	-	-	4	5	0	34	3
Total		24	45	4	29	2	27	40	0	34	3

Qual é a parte dos erros de medida nesses resultados? Para tentar responder a essa questão crucial, Jackson (1975) examina atentamente as proprie-

dades metodológicas de cada um dos estudos mencionados no quadro precedente e os subdivide em duas categorias, segundo um viés esperado em favor dos repetentes ou, ao contrário, um viés em favor dos promovidos. Como mostra o quadro 5, os resultados observados tendo em vista o primeiro critério, vão no sentido do viés esperado.

QUADRO 5
DISTRIBUIÇÃO DOS RESULTADOS SEGUNDO A NATUREZA DO VIÉS ESPERADO

Natureza do viés temido	Conhecimentos escolares				
	a	b	c	d	e
Viés a favor dos alunos promovidos	22	28	3	6	1
Viés a favor dos alunos repetentes	2	17	1	23	1

Fonte: Adaptado de Jackson (1975).

Das análises, 60 são passíveis de um viés em favor dos alunos promovidos, mas 44 são ameaçadas pelo viés inverso. No total, pode-se pensar em um leque de análises relativamente equilibrado quanto ao viés que poderia distorcer a conclusão num sentido ou no outro. No caso em questão, o balanço está ligeiramente desequilibrado em favor da promoção e, conseqüentemente, o caráter prejudicial da repetência é provavelmente um tanto subestimado. De todo modo, parece que se pode excluir a idéia de a tendência inversa ter sido pouco observada.

Essa conclusão parece ainda mais legítima ao ser corroborada pelos resultados dos três únicos estudos experimentais realizados sobre os efeitos da repetência. De fato, ao contrário da maioria dos pesquisadores que se recusaram a recorrer à indicação aleatória dos sujeitos, Klene e Branson, de 1929, Farley, de 1936, e Cook de 1941 (apud Jackson, 1975) procederam a estudos que se aproximam bastante das regras da pura experimentação. No essencial, esses três estudos não apresentam diferenças significativas. As análises de Klene e Branson se apóiam em 141 alunos distribuídos entre a 2ª e a 6ª série do ensino obrigatório. Eles concluem em favor dos alunos promovidos, mas não mencionam diferenças significativas. A pesquisa de Farley se apóia em 400 alunos distribuídos entre a 2ª e a 5ª série do ensino obrigatório. Nenhuma diferença significativa aparece no que concerne aos alunos da 2ª e da 3ª séries. Em contrapartida, no que concerne àqueles da 4ª e da 5ª séries, observa-se nos tes-

tes relativos à compreensão de texto uma diferença significativa em favor dos alunos promovidos. Para os mesmos alunos, não se observa diferença significativa no nível dos testes de Matemática. O estudo de Cook se apóia em 700 alunos distribuídos nas sete primeiras séries da escolaridade obrigatória. O autor não observa diferença significativa entre os dois grupos.

O CONTROLE ESTATÍSTICO DOS VIESES DE AMOSTRAGEM

Na França também, os efeitos da repetência foram objeto de pesquisas. A primeira, segundo o nosso conhecimento, é a obra de Seibel (1984). Aproveitando uma avaliação pedagógica executada a pedido da direção das escolas a partir de uma amostra representativa de classes e de alunos do CP (curso preparatório ou 1º ano primário), esse pesquisador analisou a evolução dos escores de 1.100 alunos em provas estandardizadas de conhecimentos de Francês e de Matemática entre junho e dezembro do ano de 1983. No interior dessa amostra, é possível identificar um grupo de repetentes e, complementarmente, constituir um grupo de alunos com dificuldades equivalentes e que não repetem. Nesse caso preciso, a situação é a que Landsheere e Delchambre (1979) designaria como um “estudo retrospectivo com esquema pré-estruturado”. Do exame dos dados, duas constatações importantes, tendo em vista nosso propósito, merecem ser retidas:

1. em junho, os resultados médios dos “não-repetentes fracos” e dos “repetentes” diferem pouco, tanto no teste de Francês (diferença 2,9) quanto no teste de Matemática (3,3);
2. em dezembro, as diferenças entre esses dois grupos aumentaram bastante. Observa-se uma diferença de 16,3 pontos no teste de Francês e de 11,7 pontos no teste de Matemática em favor dos alunos fracos que não repetiram.

A primeira dessas constatações, quer dizer, aquela relativa aos resultados da prova de junho (que, nesse caso funciona como pré-teste) pode supor um viés de amostragem em favor dos não-repetentes fracos na comparação com os repetentes. Isso pode dar lugar a dúvidas em relação aos resultados desse estudo.

O tipo de tratamento de dados utilizado por Grisy (1993), a pedido da Direção de Avaliação e Prospectiva – DEP – do Ministério de Educação Francês, no âmbito de um estudo sobre os efeitos e o funcionamento dos colégios, permite circunscrever os riscos de sub ou de superestimativa dos efeitos da repetência. Mais precisamente, essa pesquisadora utilizou dois tipos de tratamento: o primeiro, a análise de regressão, forneceu uma estimativa do efeito da repetência comprometida por erros de medida; o segundo, atualizado por Heuchenne (1993), permite estimar o verdadeiro efeito do tratamento. Trata-se de uma análise Lisrel⁵.

A amostra é constituída por alunos ingressados na 6ª série (1º ano do secundário) no ano de 1991 em 100 colégios da França metropolitana. Esses alunos realizaram em setembro de 1990, quando entraram na 6ª série, as provas estandardizadas de Francês e Matemática. Foram testados de novo dois anos mais tarde (maio de 1992), quando então a maioria deles encontrava-se na 5ª série (2º ano do secundário). Uma parte dos alunos contudo não conseguiu ir para a 5ª. Esses repetentes, que se encontram ainda no final da 6ª no momento do pós-teste, constituem o grupo-alvo da análise operada por Grisy (1993). Uma matriz completa de dados está disponível para 5.137 alunos (4.873 não repetentes e 264 repetentes).

A bateria do pré-teste comporta oito provas (quatro provas para o Francês, quatro para Matemática), totalizando 213 itens. A fidelidade global do pré-teste é satisfatória (0,79). O pós-teste comporta quatro provas. Duas delas constituem uma ancoragem, quer dizer, são compostas de itens já apresentados por ocasião do pré-teste. As duas outras provas são novas. No total, o pós-teste baseia-se em 91 itens. Sua fidelidade é boa (0,85).

A correlação entre a variável "repetência" e o critério (os resultados aos pós-testes) é negativa (-0,23). O mesmo ocorre em relação ao coeficiente de

5. Convém recomendar esse tipo de análise sempre que não se pode proceder a uma indicação aleatória dos sujeitos nos diferentes grupos estudados e sempre que não se possa garantir a perfeita (ou quase perfeita) fidelidade do pré-teste. O tratamento definido por Heuchenne (1993) leva em consideração: a fidelidade global do pós-teste; a fidelidade do pré-teste do qual nos servimos para manter sob controle as competências iniciais dos alunos; a orientação e a importância da correlação entre tratamento e competência inicial; um componente (tau) que é uma expressão complexa, na qual são levadas em conta as fidelidades já citadas, mas também uma correção global eventual dos erros de medida dos pré-testes e dos pós-testes bem como o coeficiente de regressão do pós-teste sobre o pré-teste.

regressão dessa mesma variável sobre o critério, obtido sob controle das variáveis de pré-teste (-0,207). Isso significa que, com resultado igual nas provas de Francês e de Matemática quando entraram na 6^a, os alunos repetentes obtêm, dois anos mais tarde, um resultado no pós-teste inferior de mais de 20 décimos de desvio padrão em comparação com seus colegas não repetentes. A aplicação do modelo Lisrel chega igualmente a um coeficiente negativo, mas de menor amplitude (-0,11). Ou seja, os repetentes só são lesados por um décimo no desvio padrão. Se esse valor corresponde à metade daquele obtido por equação de regressão, é contudo suficiente para negar a hipótese segundo a qual a repetência age como uma remediação que permite aos alunos recuperar suas lacunas e se lançar no percurso escolar com mais vantagens.

O estudo de Grisay (1993, ver também 2001) permite circunscrever o estado atual do debate científico. O efeito negativo da repetência encontra-se no fato de os alunos fracos terem progredido menos que os alunos que, tendo dificuldades comparáveis, acabaram promovidos. Resta portanto estimar exatamente a amplitude desse efeito negativo. A aplicação da técnica da meta-análise aos dados recolhidos sobre os efeitos da repetência pode trazer elementos que respondem a essa questão.

META-ANÁLISE E ESTIMATIVA DA AMPLITUDE DO EFEITO

A técnica da meta-análise, desenvolvida por Glass (1977), tornou-se um instrumento indispensável desde que os pesquisadores propuseram-se a fazer o balanço das pesquisas realizadas sobre uma questão específica. Essa técnica permite combinar os resultados de diferentes pesquisas, conferindo-lhes validade maior. Ocorre, com freqüência, se duvidar da conclusão de uma determinada pesquisa em razão de as análises terem sido realizadas sobre números pequenos. Além disso, pode-se igualmente suspeitar que essa determinada pesquisa seja enviesada por um erro de medida no ponto de partida e que ao final favorecerá um dos dois grupos. Com a meta-análise pode-se, por um lado, contar com um efeito de agregação dos dados, o que resolve o primeiro problema, e, por outro lado, com um efeito de compensação dos vieses: as pesquisas prejudicadas por um viés em favor dos alunos repetentes estando compensadas por aquelas ameaçadas por um viés em favor dos alunos promovidos. Esses benefícios são cada vez mais assegurados à medida que se é capaz de

combinar numerosas pesquisas. É evidente, por sua vez, que as vantagens apontadas não podem corrigir os defeitos graves eventualmente existentes em uma ou outra pesquisa. A qualidade de uma meta-análise depende portanto da qualidade das pesquisas disponíveis, mas também da severidade dos critérios metodológicos adotados pelo autor na seleção das pesquisas.

A meta-análise de Holmes e Matthews (1984) e, mais ainda, a de Holmes (1990) constituem modelos do gênero. Assim, em 1984, Holmes e Matthews examinaram 650 pesquisas sobre os efeitos da repetência. Foram escolhidas 44, por apresentarem garantias suficientes quanto ao rigor metodológico da investigação e, sobretudo, quanto à equivalência inicial do grupo experimental e do grupo de controle. Quer dizer, os pesquisadores se mostraram severos quanto aos critérios metodológicos utilizados para conservar uma pesquisa na amostra que devia dar lugar à meta-análise. Em 1990, Holmes manteve esse rigor metodológico. Em resumo, ele conservou 63 pesquisas, das quais 44 faziam já parte de sua publicação anterior com Matthews.

Na meta-análise de 1984, todas as amplitudes calculadas do efeito⁶ são negativas. Naquelas de 1990, todos os efeitos são negativos, com uma exceção: de fato, observa-se um efeito positivo sobre a imagem de si, mas de amplitude irrisória: +0,06. O quadro 6 sintetiza os resultados obtidos no caso dessas duas meta-análises.

Os efeitos negativos da repetência sobre os ganhos de conhecimento dos alunos parecem dificilmente discutíveis. A amplitude dos efeitos calculados tomando por critério uma ou outra medida de performance varia de -0,25 (em Matemática) a -0,37 (em Ciências) quando estes são medidos por meio de provas externas; é de -0,78 quando se utilizam provas concebidas pelos próprios professores. Os efeitos "afetivos" da repetência são, em geral, negativos (com exceção da imagem de si), mas de menor amplitude: os valores variam

6. Lembremos brevemente os princípios dessa técnica. As pesquisas baseadas numa mesma variável (aqui, a repetência) são recenseadas. Conservam-se os estudos nos quais os pesquisadores procederam à comparação de um grupo experimental com um grupo de controle. Os resultados são expressos de forma estandardizada: diferença entre grupo experimental e grupo de controle, dividida pelo desvio padrão do grupo de controle. Média do grupo experimental - média do grupo de controle = amplitude do efeito. A importância ou amplitude do efeito da variável estudada é expressa em frações do desvio padrão; ela pode igualmente ser expressa em centis.

entre -0,12 para o bem-estar emocional e -0,23 para as atitudes comportamentais. Na conclusão, voltaremos a essa difícil questão dos efeitos afetivos da repetência, mas desde já é preciso lembrar que se trata de um componente psicológico mais dificilmente apreensível pelos testes do que as aquisições cognitivas.

QUADRO 6
AMPLITUDE DO EFEITO DA REPETÊNCIA
(EM FRAÇÕES DE DESVIO PADRÃO)
EM UMA SÉRIE DE VARIÁVEIS-CRITÉRIOS

Critérios medidos	Holmes e Matthews (1984)		Holmes (1990)	
	Número de estudos	Amplitude de efeito	Número de estudos	Amplitude de efeito
Performances acadêmicas	31	- 0,44	47	- 0,31
Performances em língua materna	-		18	- 0,33
Performances em leitura	24	- 0,48	34	- 0,30
Performances em matemática	20	- 0,33	31	- 0,25
Performances em atividades científicas	-		3	- 0,37
Resultados de provas do professor	-		3	- 0,78
Métodos de trabalho	1	- 0,41	-	
Reações afetivas gerais	-		27	- 0,21
Desenvolvimento social	13	- 0,27	27	- 0,21
Bem-estar emocional	5	- 0,37	10	- 0,12
Atitudes comportamentais em classe	7	- 0,31	10	- 0,23
Imagem de si	9	0,19	11	+ 0,06
Atitudes em relação à escola	8	0,16	10	+ 0,18
Frequência escolar	-		5	-0,22

Fonte: Adaptado de Holmes, Matthews (1984), e de Holmes (1990).

Na meta-análise de 1990, Holmes se ateve a ver em que medida o efeito da repetência sobre as aquisições cognitivas dependia do momento em que ele ocorria. O quadro 7 indica os resultados que o autor conseguiu obter.

É preciso notar ainda que todos os valores são negativos. Contrariamente a uma opinião largamente difundida no corpo docente, a repetência precoce (antes de entrar na escola primária ou no 1º ano) não tem efeito preventivo;

as amplitudes do efeito, medidas respectivamente em oito estudos para o nível maternal e em 12 para o 1º ano, são de -0,28.

QUADRO 7
AMPLITUDE DO EFEITO DA REPETÊNCIA (EM FRAÇÕES DE DESVIO PADRÃO)
MEDIDA PELOS TESTES DE CONHECIMENTO EM FUNÇÃO DO NÍVEL
DE ESTUDO NO QUAL OCORRE A REPETÊNCIA

	Em maternal	P 1	P 2	P 3	P 4	após a 4ª
Número de estudos	8	12	4	7	6	5
Amplitude do efeito	- 0,28	- 0,28	- 0,10	- 0,15	- 0,36	- 0,38

Fonte: Adaptado de Holmes (1990).

Mas o interesse metodológico principal da meta-análise de 1990 localiza-se no modo pelo qual Holmes aborda uma das maiores dificuldades dos estudos sobre os efeitos da repetência: enquanto os repetentes devem recomençar aprendizagens já abordadas no ano anterior, os promovidos são submetidos a novas aprendizagens. Um ano mais tarde, no momento em que os pesquisadores colhem as medidas para estabelecer a comparação, o segundo grupo têm um ano escolar à frente dos outros, embora ambos tenham a mesma idade. Dispondo de provas estandardizadas e escalonadas sobre vários anos escolares, é legítimo que se observe os progressos de uns e de outros. É essa abordagem que privilegia a maioria dos estudos utilizados nas meta-análises de Holmes e Matthews (1984) e de Holmes (1990). Parece contudo oportuno se perguntar sobre o que ocorreria se se comparassem alunos com dificuldades equivalentes, mas que repetiram um ano a fim de consolidar as aprendizagens pré-requeridas para os anos posteriores. Na verdade, o princípio da repetência consiste em sacrificar um ano para permitir à criança reiniciar sobre bases melhores e atingir – com um ano de atraso, não obstante – níveis de competências que ela não poderia alcançar se não repetisse. Esse raciocínio leva a um outro tipo de comparação no qual se relacionam os níveis de conhecimento alcançados por alunos fracos que repetiram, com aqueles alcançados por alunos fracos que nunca repetiram, e isso num mesmo nível escolar (por exemplo, ao final de um ciclo). Em alguns dos estudos arrolados por Holmes em 1990, os alunos repetentes e seus colegas do grupo de controle foram acompanhados durante anos. Isso permite que se tenha uma idéia do efeito da repetência a médio prazo.

Mais precisamente, Holmes procede a dois tipos de comparação.

- Num caso, comparam-se alunos que repetiram com alunos fracos não repetentes com a mesma idade. Em conseqüência, os alunos não repetentes se beneficiaram do ensino de um ano suplementar. Nesse caso, considera-se a **comparação com a mesma idade**.
- Num segundo caso, comparam-se os alunos repetentes e não repetentes ao final do ano escolar, sabendo-se que os primeiros têm um ano a mais em relação aos segundos. Será considerada a **comparação com o mesmo grau escolar**.

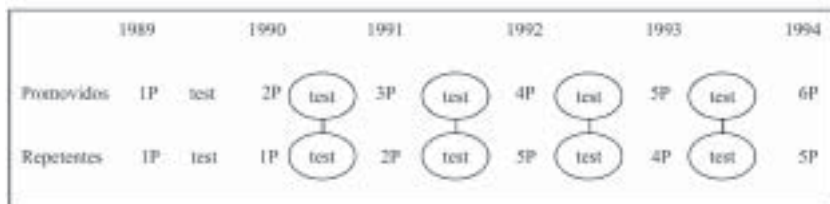
Pode-se tornar mais claro o princípio dessas duas comparações imaginando uma situação concreta. Suponha-se um grupo de alunos que entram no primeiro ano (1P) em 1989. No final do ano, uma parte deles repete e outra parte é promovida. Em 1990, alguns permanecem em 1P e outros vão para 2P. Imaginemos que, para essas duas coortes de alunos, a continuação do percurso se realiza sem obstáculo, o que conduzirá à seguinte situação:

	1989	1990	1991	1992	1993	1994
Promovidos	1 P	2 P	3 P	4 P	5 P	6 P
Repetentes	1 P	1 P	2 P	3 P	4 P	5 P

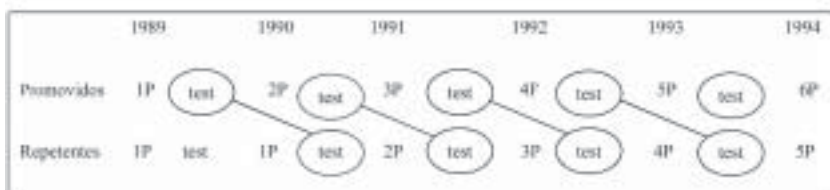
Se os alunos dos dois grupos se submetem a testes de conhecimento ao final de cada ano, será possível realizar as duas comparações explicadas a seguir:

- Em 1991, comparam-se as performances dos promovidos quando eles se encontram em 3P com aquelas dos repetentes quando estão em 2P. Procede-se a uma comparação com a mesma idade.
- Comparam-se as performances dos promovidos obtidas em 1990 no final de 2P com aquelas obtidas pelos repetentes em 1991 no final do mesmo grau. Procede-se a uma comparação com o mesmo grau escolar.

Note-se que, nos dois casos, os repetentes estão em vantagem. Tome-se, para facilitar a explicação, o caso dos alunos que repetiram no primeiro ano:



- Quando se comparam as performances dos alunos de mesma idade, os escores dos repetentes correspondem às provas de 1P por ocasião de seu segundo primeiro ano; os escores dos promovidos correspondem às provas do 2P . Ou seja, esses últimos são submetidos a um teste mais difícil⁷.



- Quando se compara com o mesmo grau escolar, esse viés não existe mais. Contudo, nessa condição experimental, os repetentes têm um ano a mais, portanto, um benefício em termos de maturidade.

Os resultados dessas duas comparações são apresentados nos quadros 8 e 9.

QUADRO 8
EFEITOS DA REPETÊNCIA A MÉDIO E A LONGO PRAZO
RESULTADOS DAS COMPARAÇÕES COM A MESMA IDADE

	Após 1 ano	Após 2 anos	Após 3 anos	Após mais de 3 anos
Número de estudos	28	5	3	3
Amplitude de efeito	- 0,41	0,64	0,74	0,88

Medida em fração de desvio padrão.

Fonte: Adaptado de Holmes (1990).

7. Não obstante existam itens de ancoragem (itens que estão presentes nos testes de primeiro e de segundo anos) e a aferição dos escores torne os dados das duas amostras comparáveis, nem por isso o teste de segundo ano é mais fácil.

QUADRO 9
EFEITOS DA REPETÊNCIA A MÉDIO E A LONGO PRAZO
RESULTADOS DAS COMPARAÇÕES COM O MESMO GRAU ESCOLAR

	Após 1 ano	Após 2 anos	Após 3 anos	Após mais de 3 anos
Número de estudos	10	7	5	6
Amplitude de efeito	0	+ 0,02	- 0,12	+ 0,04

Medida em fração de desvio padrão.

Fonte: Adaptado de Holmes (1990).

Essa dupla constatação é esclarecedora. Quando se compara o nível cognitivo alcançado pelos repetentes e os não repetentes no mesmo grau escolar, as performances de ambos os grupos são equivalentes, mesmo quando as comparações são estabelecidas um ano após a repetência ou mais de três anos após. Isso significa, portanto, que, mesmo tendo sacrificado um ano a fim de reiniciar sobre bases melhores a continuidade escolar, os alunos repetentes não se mostram melhores que seus colegas que não fizeram esse sacrifício. De nosso ponto de vista, essa ausência de efeito por ocasião das comparações em relação ao grau escolar é o que demonstra mais a inutilidade da repetência na sua pretensão de ajudar os alunos em dificuldade. Se, além disso, se levam em consideração os resultados das comparações com a mesma idade (o que coloca em evidência que o atraso dos alunos repetentes vai aumentando com os anos em relação àqueles que foram promovidos), é pertinente concluir que um ano repetido é na verdade um ano perdido.

DEFINITIVAMENTE, O QUE SE PODE CONCLUIR?

A menos que se queira negar a amplitude das provas e a teimosia dos pesquisadores em encontrar a maneira mais válida de apreender os efeitos da repetência, alguém de bom senso deve admitir que os dados de pesquisa não pleiteiam a manutenção dessa prática. Parece aceito hoje que o fato de repetir um ano e de recomeçar toda a programação de um curso não ajuda os alunos em dificuldade a superar os obstáculos que os impedem de ser honrosamente bem-sucedidos na escola. Considerando-se que as pesquisas ditas quase-experimentais raramente focalizam o curso médio, pode-se compreender que alguns resistem a estender aos adolescentes aquilo que foi observado em relação às crianças do ensino fundamental.

Do ponto de vista do pesquisador, parece urgente ultrapassar a polêmica sobre os efeitos da repetência no ensino fundamental para privilegiar outros questionamentos. Se a repetência não constitui um meio de ajuda para os alunos em dificuldade, parece necessário procurar outros meios para resolver esse importante problema. Ou seja, em vez de solicitar novas provas quanto aos efeitos da repetência, talvez seja mais profícuo pedir aos pesquisadores que se debrucem sobre outros objetos de investigação, uma vez que, em vista da qualidade dos esforços mobilizados para conjurar os vieses de amostragem e de medida nos estudos sobre a repetência, parece difícil melhorar ainda mais a validade das demonstrações e bastante improvável alterar a tendência das conclusões. Somente a questão dos efeitos socioafetivos da repetência poderia ainda valer alguns esforços de pesquisa. Nesse domínio, trabalhos de tipo qualitativo tendem a aportar algo mais (Crahay, 2003), o que, de qualquer modo, resulta em mais argumentos contrários à repetência.

Resta saber se, em matéria de educação, os profissionais da área e os gestores políticos estão dispostos a se deixar convencer por um conjunto de pesquisas que, tendo melhorado significativamente o controle dos vieses de medida, chega a resultados convergentes.

REFERÊNCIAS BIBLIOGRÁFICAS

COOK, T. D.; CAMPBELL, D.T. *Quasi experimentation: design and analysis issues for field settings*. Chicago: Rand Mc Nally, 1979.

CRAHAY, M. *Peut-on lutter contre l'échec scolaire?* (2.ed. rev. et augm.). Bruxelles: De Boeck, 2003.

_____. Recherche en éducation et réflexion de l'action éducative. In: HADJI, C.; BAILLÉ, J. (ed.) *Recherche en éducation: vers une nouvelle "alliance"* Bruxelles: De Boeck, 1996. p.125-160. (Pédagogies en développement)

_____. La Recherche en éducation: une entreprise d'intelligibilité de faits et de représentations ancrés dans l'histoire sociale. In: LEUTENEGGER, F.; SAADA-ROBERT, M. (ed.) *Expliquer et comprendre en sciences de l'éducation*. Bruxelles: De Boeck, 2002. p.253-275. (Raisons éducatives, n. 5)

GLASS, G. V. Class size. In: DUNKIN, M. J. (ed.) *The International encyclopedia of teaching and teacher education*. Oxford: Pergamon, 1988. p.540-545.

GLASS, G. V. Integrating findings: the meta-analysis of research. *Review of Research in Education*, Itasca, IL: Peacock, v.5, p.351-379, 1977.

GRISAY, A. Évaluer les dispositifs de prise en charge d'élèves faibles (ou forts): l'utilisation de groupes naturels entraîne des artefacts. *Cahiers du Service de Pédagogie Expérimentale*. Liège: Université de Liège, n.7/8, p.129-151, 2001.

_____. *Le Fonctionnement des collèges et ses effets sur les élèves de sixième et de cinquième*. Paris: Ministère de l'Éducation Nationale, Direction de l'Évaluation et de la Prospective, 2001. (Les Dossiers Éducation et Formation, n. 32)

_____. Les Mirages de l'évaluation scolaire (I): rendement en français, notes et échecs à l'école primaire? *Revue de la Direction Générale de l'Organisation des Études*, v.19, n.5, p.29-42, 1984.

HEUCHENNE C. Effet des erreurs de cité d'un traitement. *Mathématique, Informatique et Sciences humaines*, v.31, n.122, p.5-19, 1993.

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, 1990, 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.

JACKSON, G. B. The Research evidence on the effects of grade retention. *Review of Educational Research*, v.45, n.4, p.613-636, 1975.

LANDSHEERE, G. de. *Dictionnaire de l'évaluation et de la recherche en éducation*. Paris: Presses Universitaires de France, 1992.

LANDSHEERE, G. de; DELCHAMBRE, A. *Les Comportements non verbaux de l'enseignant: comment les maîtres enseignent*, 2. Nathan: Paris, 1979.

MILNER, J.-C. *De l'école*. Paris: Seuil, 1984.

SEIBEL, C. Genèses et conséquences de l'échec scolaire. *Revue Française de Pédagogie*, n.67, p.7-28, 1984.

SLAVIN, R. E. Achievement effects of ability grouping in secondary schools: a best-evidence synthesis. *Review of Educational Research*, v.60, n.3, p.471-499, Fall 1990.

STEGEN, P. *Quelques éléments du contexte dans lequel s'est implantée la réforme du premier cycle de l'enseignement secondaire*. Liège: Service de pédagogie expérimentale de l'Université;

Ministère de l'éducation de la Communauté française de Belgique, 1994. (Rapport d' une recherche)

PETERSON, S. E.; DEGRACIE, J. S.; AYABE, C. R. A Longitudinal study of the effects of retention/promotion on academic achievement. *American Educational Research Journal*, v.24, n.1, p.107-118, 1987.

POPPER, K. R. *La Connaissance objective*. Paris: Éditions Complexes, 1982.

VAN DER MAREN, J. M. *Méthodes de recherche pour l'éducation*. Bruxelles: De Boeck, 1995.

Recebido em: dezembro 2005

Aprovado para publicação em: dezembro 2005