The effects of year repetition (redoublement) on the progress of pupils in the first three years of French schooling.

by

Harvey Goldstein

University of Bristol

h.goldstein@bristol.ac.uk

INTRODUCTION

In the French educational system, difficulties which occur in the first years of primary school are perceived to be rarely overcome and to have a negative effect on the child's future achievement at school. It is therefore important to attempt to identify precursors of failure in the first years of elementary education, both individual ones and contextual ones at the school level. The present study uses a panel of pupils followed for the first two years of compulsory schooling.

This report was prepared for the French Ministry of Education, Paris, with the assistance of Gerard Bonnet. It was submitted to the Ministry in June 2003.

How the panel was recruited.

The French ministry of national education has followed several panels since 1973. The present analyses were carried out on the 1997 panel which is a representative sample of the pupils who arrived in CP ("cours préparatoire") in September 1997. It comprises 9,259 pupils recruited in 1,401 public/state or private primary schools in metropolitan France who have been followed since then. None of the previous panels had a pre-test score at start of CP so that we cannot carry out comparisons with these.

A two-stage sample was used. This method set out to sample a limited number of schools, for reasons of cost, and also to select a limited number of pupils from the same school not to give extra work to teachers and principals of schools. The sample included French primary schools in which there were some pupils in CP in September 1997. Schools were divided into 9 strata according to the number of pupils in CP, the school's sector and to whether the school was in a ZEP ("Zone d'Education Prioritaire"). Each stratum was again divided according to the classification of the school by type of area - urban, Paris and suburbs, country, etc. Pupils were then recruited from the selected schools using a different procedure according to the number of children in CP in 1997.

In the strata constituted by the schools with less than 11 pupils in CP, one in 90 schools was selected. All the pupils in CP in those schools were then recruited into the sample.

In the strata which comprised the schools in which there were between 11 and 30 pupils in CP, one in 30 schools was selected. In each of these, one in three pupils was recruited

with the following method. Heads were asked to write down the names of all the pupils in CP in alphabetical order. The pupils ranked 2, 5, 8 etc. were then recruited.

Another method was used for the schools with more than 30 pupils in CP altogether. Since the pupils of the panel were to take evaluation tests, it was decided to select pupils from one class only in schools to make it easier to organise the tests. Schools were thus given weights which corresponded to their number of classes comprising pupils in CP in 1997 before the selection took place. Classes were then chosen in the selected schools on the basis of the teacher's name. All classes were thus given the same probability of being selected. Pupils were then recruited from the selected classes with the procedure that was used for schools with 11 up to 30 pupils in CP. Finally one in 90 pupils was selected.
It was not possible, in the analysis, to identify each class uniquely, so that we are unable to estimate separate class and school effects.

How the data were collected.

The data were collected at different periods of time. Every year all heads in whose school there are pupils from the panel are sent a questionnaire to update the pupil's situation at primary school level - the number of pupils in the class, class composition, etc.

On joining the panel in 1997 the heads of the schools of the selected pupils were sent an additional questionnaire. This provided information on the child and family as well as schooling at nursery school level and the situation at primary school level.

Information on familial background was also collected by means of a questionnaire sent to families in 1999 and giving detailed information about the composition of the family - number of children, birth order etc. - and the parents - profession, degrees, nationality, parental language, etc. It also asked questions about the child's progress at nursery school and his/her relations to the primary school as well as his/her family's involvement with the school, etc.

The pupil's attainments at the start of CP was measured by evaluation tests carried out in 1997 when the sample was recruited. The tests concern general knowledge, verbal abilities and familiarity with written language, logical skills and familiarity with numbers, time- and space-related concepts, behaviour and attention. Teachers were also asked to fill in an observation grid about the behaviour and abilities of the pupil at the start of CP.

At the start of CE2 the pupils took evaluation tests in French and Maths. Their scores were collected in 1999 and also in 2000 for those who repeated CP or CE1. The pupils who enter into the analyses comprise those for whom information on all the variables used is present. This number may vary somewhat between analyses.

Analyses

We have carried out multilevel analysis (Goldstein, 2003) that recognises the fact that schools differ in terms of the achievement of their pupils. This form of analysis, an

extension of multiple regression, explicitly incorporates differences among schools and thus allows us to make valid inferences about relationships.

Three sets of analyses were carried out. The first is an analysis of the proportion who arrive in CE2 without repeating a year, and the other two analyses use the French and Mathematics scores at the start of CE2. For the latter we use only the scores for those pupils taking the tests in 1999, i.e. omitting the repeaters who took the tests a year later. In our conclusions, therefore, we look separately at the probability of repeating and, for those who do not repeat, the prediction of their test scores. An alternative way of modelling these data is to view the repeating pupils as conceptually having Mathematics and French scores, but that these are unobservable. Thus, we would regard their scores as 'informatively missing' and carry out a suitable procedure to correct the estimates obtained in the separate multilevel analyses for Mathematics and French (see Goldstein, 2003, Chapter 14). Inferences from such an analysis would then refer to the whole population starting in CP. There are difficulties with such an approach. First, the removal of these pupils creates a different context for those who remain and it is not just the fact that we have no test scores on the repeaters that changes the relationship. Secondly, if we wish to make predictions of test scores at start of CE2, these should be carried out conditionally on those who do not repeat. For these reasons, therefore, we restrict ourselves to a conditional analysis. We do however obtain estimates for certain relationships at the school level, below.

All the analyses use an initial pre-test score as an adjustment so that we can interpret our results in terms of the relative progress made between entry and start of CE2. This pre-test score is a global test score compiled from 163 separate items reflecting language and general understanding, measured at the start of the first year (CP). We have looked at the separate components of this score, but using these jointly does not substantially change conclusions. The analyses are also multilevel so that between-school differences are incorporated. The pattern of description is as follows. We first describe the data in terms of the basic distribution of each response variable, then we adjust for the pre-test and then report a 'final' model which includes the variables found to be statistically significant during the course of extensive exploratory analyses involving different combinations of explanatory (predictor) variables. We do not present all the intermediate results and we also retain variables, in general, that are significant at the 10% level, together with certain other variables where the lack of significance is of particular interest. Nevertheless, omitting these variables with very small effects does not change to any marked degree the remaining values in the model. They are left in the tables for the reader to emphasise those variables where there may be a special interest in seeing if effects do exist.

Description of variables used.

We include here a brief description of the variables used in the analyses.
The French evaluation concerns comprehension, language tools and written production and the Maths evaluation includes geometry, measuring, numbers and problems.
The information on individual characteristics was collected from the questionnaire filled in on joining the panel in 1997 and also from the family questionnaire in 1999.

Gender, nationality and birth order of the child are included in the models. Birth date is included as follows. If the child was born in 1991, the quarter is used and this is the case

for 97 % of the pupils in the panel. Otherwise the variable gives the year when he/she was born i.e. 1989, 1990 or 1992.

As regards the family, we have information on the socio-economic status of the household head which is classified into 10 categories as shown in the tables. We also use whether the mother is working or not as well as the parental situation: whether the child is living with both his real parents, with one parent only, with one real and one other parent or is in another parental situation.

The highest level of qualifications of both father and mother are included among the socio-demographic characteristics. The language spoken by the parents to the child is used and differentiates the families in which both parents always speak French at home from those in which only one parent always speak French and those where neither of the parents uses French.

The involvement of parents in their child's education is also used. Two types of involvement are defined according to their relationship with school. The variable which describes explicit involvement was based on the answers to the family questionnaire that described the relations between parents and their child's school: whether the child is regularly helped in his/her homework in the evening, whether the parents asked to meet a member of the school staff, attended meetings with teachers, accompanied a school outing ; whether they are parent representatives, mediators, members of a parent-teachers association or parent correspondents - in areas where there are many immigrants for instance, certain parents help to keep the other parents in touch with the school.

The other variable measures implicit involvement. It describes aspects of parents' behaviour outside school: taking their children to the library or the cinema or the theatre or the museums, taking part together in sports or having walks together, doing handiwork together, playing games on the family computer, playing board games, making the child go to bed before 9 p.m.; whether the child is enrolled in a club for sport or handiwork, attends a school of music or dance, belongs to a young people's association, or is a member of a library.

We take into account whether the child has lunch at home or is supervised after school and the number of years spent at nursery school are also included in the model. To obtain an idea of how the child settled at nursery school we constructed a variable, using answers to the family questionnaire, using information about whether the child enjoyed attending nursery school, had many friends and learnt much there.

We also have information about the level of the pupil at the start of CP. This was measured by evaluation tests carried out in 1997 when the panel was recruited. They concern general knowledge, verbal abilities and familiarity with written language, skills in reasoning and familiarity with numbers, time- and space-related concepts, behaviour and attention in class. A behaviour score was calculated using teachers' answers to an observation grid about the behaviour and abilities of the pupil at the start of CP.

Characteristics of the school such as the sector and the location in a ZEP ("Zone d'Education Prioritaire") are included in the models.As regards the class, the number of

pupils is taken into account as well as the proportion of foreign pupils. We also include the composition of the class whether or not it has multiple year groups.

We describe now each analysis in detail.

Analysis of repetition (Redoublement).

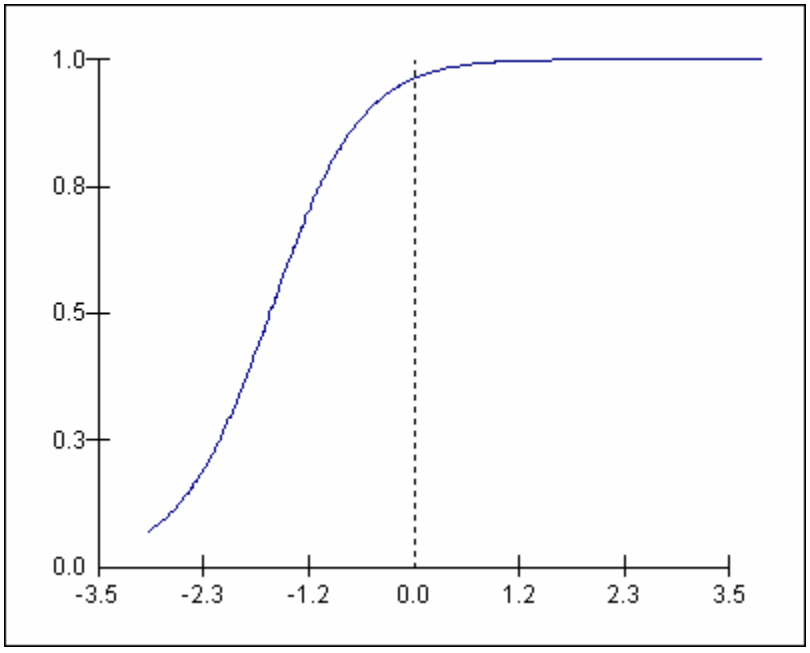
A total of 8902 pupils have information on repetition and 12.0% repeat overall. The first analysis reported uses the following statistical model that incorporates just an intercept term and terms for the pre-test.

$$\begin{aligned} &\log it(\pi_{ij}) = \beta_0 + \beta_1 x_{1ij} + \beta_2 x_{1ij}^2 + u_j \\ &y_{ij} \sim binomial(1, \pi_{ij}), \quad u_j \sim N(0, \sigma_u^2) \end{aligned} \qquad (1)$$

This is a 2-level logistic model where the response, *y*, is 1 if the pupil arrives in CE2 without repeating and 0 if the pupil does. The predictor, *x*, is a Normal score transformation of the global pretest score and we can satisfactorily describe the relationship using the terms in model (1). The term *u* is the school 'random effect' so that the model allows for schools to differ in terms of their average prediction of the repetition probability. Table 1 presents the results of this analysis.

Using the results in Table 1, this relationship is graphed in Figure 1.

Figure 1. Proportion not repeating by Normalised composite pretest score.



It is clear that for pupils whose intake score is above the mean (0.0) the probability of not repeating is effectively one. For the very low achievers this probability decreases rapidly towards zero.

Table 1. Response is arriving in CE2 without repeating. PQL2 model using logit link.		
Fixed	Estimate	s.e.
Intercept	3.361	
SCGLO	2.198	0.124
SCGLO^2	-0.012	0.066
Random		
Level 2 variance	0.790	0.116

We note that the quadratic term here is not statistically significant, but is retained because it does become so in subsequent analyses. The PQL2 estimation procedure is used throughout since the numbers in each school are not too small and the level 2 variance not too large. MCMC estimation gives very similar results.

We now present the full model in Table 2 using all the variables. The variable categories will be described below. We fit only variance components models, with a single random effect for each school, since we can find little evidence for any random coefficients.

Table 2. Response is arriving in CE2 without repeating. PQL2 model using logit link.		
Fixed	Estimate	s.e.
Intercept	2.639	
SCGLO	1.860	0.142
SCGLO^2	-0.194	0.080
Female	0.282	0.110*
Artisan-commercial	-0.752	0.484
Higher professional	0.065	0.525
Teacher	0.552	0.891
mid-professional	-0.485	0.479
Employee	-0.833	0.469
Skilled	-0.932	0.459
Unskilled	-1.338	0.474

Unemployed	-0.908	0.490
1 parent	-0.325	0.204
1 real 1 other parent	-0.628	0.288
Other parental situation	0.000	0.000
Private school	-0.120	0.204
Explicit involvement	0.080	0.046
1 parent French	-0.036	0.176
Neither parent French	1.053	0.187
Has lunch at home	0.143	0.119
Mother Primary cert.	0.259	0.207
Mother BEPC	0.455	0.174
Mother CAP, BEC	0.418	0.152
Mother BAC	1.049	0.230
Mother univ	0.942	0.258
Mother working	0.160	0.125
Birth order	-0.133	0.050**
In ZEP year 1	0.198	0.175
Number of pupils in class CP	-0.017	0.018
Child settled well at nursery	0.236	0.028***
No. of nursery years 1-2	-0.566	0.501
No. of nursery years 2-3	-0.107	0.461
No. of nursery years 3+	0.138	0.471
Birth 2nd quarter 1991	-0.022	0.174
Birth 3rd quarter 1991	-0.093	0.166
Birth 4th quarter 1991	0.054	0.171
Birth 1989/90	0.213	0.390
Birth 1992	-0.354	0.528
Multiple year class CP	-0.186	0.158
Random		
Level 2 variance	0.942	0.161

For variables with 2 categories, or for continuous variables, we denote significance as follows:

* P<0.05: ** P<0.01: *** P<0.001.

The following categories are chosen as base categories for the above variables:

Table 3. Base categories

Variable	***Base category***
SES of household head	Agricultural workers
Parental situation	Both real parents
Parental language	Both French
Mother education	No diploma
Number of nursery years	One year or less
Birth year	First quarter 1991

The variables that have been dropped from original table because they were not significant are:

Location of school (urban etc), implicit involvement of parents in school, fathers education, proportion of foreign pupils in class, supported after school.

Significance tests and multiple comparisons

For variables with three or more categories, pairwise multiple comparisons have been done with the following differences detected:

Table 4. Significance tests

Variable: overall χ^2	***Pairwise differences significant at 5% level***
SES of head: 29.8 (8 df) P<0.001	Unskilled *vs* Higher professional (p=.01)
Parental situation: 5.9 (3 df) P>0.10	
Parental language: 33.3 (2 df) P<0.001	1 parent French *vs* both other: Both other *vs* both French
Mother's education: 27.5 (5 df) P=0.000	(CAP, BEP; BAC; Univ.) *vs* No diploma
Years at Nursery school: 9.1 (3 df) P=0.03	1-2 years *vs* 3+ years
Years at nursery school*: 7.8 (2 df) P=0.02	3+ vs 0-2 years
Birth date: 1.9 (5 df) P>0.10	

* from analysis merging < 1 year and 1-2 years

Comments on repeating analysis

All the comparisons are adjusted for the other variables in the model. In particular adjustment for the initial global score. Note that for variables with just 2 categories or continuous variables, we can compare directly with the estimated standard errors.

Girls are less likely to repeat than boys. We can calculate the 'odds ratio' for pairs of categories. The odds for a particular category is simply the probability of not repeating divided by the probability of repeating. For two categories we use the odds ratio which is the odds for one category divided by the odds for the other. Thus, for example, for girls:boys this is 1.3, which means that girls have 1.3 times the odds of arriving in CE2 without repeating. The use of 'odds' is common in risk analysis, and also commonly used in betting situations. There are differences according to the SES of head of household and a clearly identified difference between those from unskilled and higher professional households, with the former more likely to stay on (odds ratio= 4.1). There are no detectable differences between different parental situations. There is no detectable difference between private and public schools. There is no evidence that explicit or implicit involvement in school activities is related to repeating. Parental language has an effect with a definite difference between neither parent speaking French and either one or both speaking French. Having lunch at home has no significant effect on repeating. Mother's education is strongly associated with repeating, but not father's education. The probability of not repeating tends to be greater when the qualification of the mother is higher. Higher birth orders are associated with a higher probability of repeating. There is no statistical evidence that being in a ZEP results in a higher probability of repeating. The number of pupils in the class is not associated with a higher probability of repeating. If parents reported that the child settled well in nursery school they are more likely not to repeat (odds ratio = 1.3).

As regards the number years at nursery school, there is no significant difference for 'one year or less' and 1-3+, but a difference between 1-2 and 3+. This result is a little difficult to explain. This seems to be due to the choice of the categories for the variable. When we put together the pupils who spent less than 1 year at nursery school with those who spent 1-2 years, the coefficient for 2-3 becomes positive (0.361). There is no significant effect for birth date. Having a class with multiple year groups has no significant effect.

There is little evidence of any random coefficients. That is, the relationship between the predictors and the probability of response does not vary from school to school: schools differ only in terms of their average probability of pupils not repeating. The between-school standard deviation is 0.8 which is relatively large compared to the other effect sizes. However, because the number of pupils in each school is only about 8 it is not possible to determine with any accuracy effect values for each school.

Analysis of French and Maths at start of CE2.

The same variables are used for Maths and French so that comparisons can easily be made.

Table 5. Maths and French responses at start of CE2. Variance components model.				
Fixed	Maths		French	
	Estimate	s.e.	Estimate	s.e.
Intercept	-0.697		-0.994	
SCGLO	0.670	0.014	0.620	0.014
SCGLO^2	-0.038	0.009	-0.018	0.009
Female	-0.081	0.019***	0.256***	0.019
Artisan-commercial	-0.009	0.066	0.031	0.066
Higher prof.	0.121	0.065	0.165	0.065
Teacher	0.017	0.079	0.136	0.078
mid-professional	0.042	0.062	0.128	0.062
Employee	0.004	0.064	0.089	0.063
Skilled	-0.003	0.062	0.053	0.061
Unskilled	-0.046	0.071	-0.010	0.071
Unemployed	0.056	0.081	0.189	0.081
In institution	-0.194	0.400	-0.141	0.398
Foreign	-0.019	0.065	-0.041	0.065
Private school	-0.123	0.038**	-0.044	0.037
Explicit involvement	-0.023	0.008**	-0.021	0.008**
Implicit involvement	0.022	0.014	0.041	0.014***
Mother Primary Cert.	0.087	0.056	0.075	0.056
Mother BEPC	0.076	0.041	0.078	0.041
Mother CAP, BEC	0.065	0.037	0.075	0.037
Mother BAC	0.207	0.041	0.235	0.041

Mother university	0.240	0.042	0.251	0.042
Father Primary cert.	-0.072	0.055	-0.019	0.055
Father BEPC	0.050	0.043	0.101	0.043
Father CAP, BEC	0.055	0.035	0.115	0.035
Father BAC	0.146	0.043	0.210	0.043
Father university	0.097	0.045	0.229	0.044
Birth 2nd quarter 1991	0.021	0.028	0.016	0.028
Birth 3rd quarter 1991	-0.032	0.029	-0.030	0.029
Birth 4th quarter 1991	-0.050	0.030	-0.044	0.030
Birth 1989/90	-0.728	0.121	-0.749	0.120
Birth 1992	0.165	0.083	0.118	0.083
Birth order	-0.004	0.010	-0.029	0.010***
In ZEP CE1	-0.161	0.046***	-0.103	0.044**
Not supported after school	0.002	0.033	0.039	0.033
1 parent French	-0.017	0.045	-0.007	0.045
Neither parent French	-0.035	0.048	-0.082	0.047
Settled well at nursery	0.024	0.006*	0.034	0.006***
No. of nursery years 1-2	0.247	0.121	0.081	0.120
No. of nursery years 2-3	0.159	0.111	0.077	0.110
No. of nursery years >3	0.183	0.112	0.061	0.111
Multiple year groups	0.066	0.026*	0.077	0.025**
1 parent	-0.120	0.051	-0.140	0.051
1 real 1 other parent	-0.155	0.069	-0.215	0.0689
Other situation	0.0231	0.315	-0.029	0.315
Random				
Level 2 variance	0.136	0.010	0.110	0.009
Level 1 variance	0.424	0.010	0.429	0.010
Percentage of variance at level 2	24%		20%	

For variables with 1 category, or for continuous variables, we denote significance as follows:

* P<0.05: ** P<0.01: *** P<0.001.

Table 6. Base categories	
Variable	***Base category***
SES of household head	Agricultural workers
Parental situation	Both real parents
Parental language	Both French
Mother and father education	No diploma
Number of nursery years	None
Birth year	First quarter 1991

The variables that have been dropped from original table because they were not significant are:

Location of school (urban etc), lunch at home.

Significance tests and multiple comparisons

For variables with three or more categories, pairwise multiple comparisons have been done with the following differences detected:

a). Table 7. Maths

Variable: overall χ^2	***Pairwise differences significant at 5% level***
SES of head: 16.0 (9 df) P=0.07	
Parental situation: 9.6 (3 df) P=0.02	
Parental language: 0.9 (2 df) P>0.10	
Mother's education: 51.4 (5 df) P<0.001	(None, Primary, BEPC, CAP-BEC) *vs* (BAC, Univ)
Father's education: 19.4 (5 df) P=0.002	Primary *vs* BAC
Years at Nursery school: 5.8 (3 df) P=0.12	
Birth date: 48.9 (5 df) P<0.001	1989/90 *vs* every other category

b). Table 8. French

Variable: overall χ^2	***Pairwise differences significant at 5% level***
SES of head: 22.7 (9 df) P=0.007	
Parental situation: 16.0 (3 df) P=0.001	(1 parent, 1 real&1 other) *vs* both real
Parental language: 3.0 (2 df) P>0.10	
Mother's education: 60.5 (5 df) P<0.001	(None, Primary, BEPC, CAP-BEC) *vs* (BAC, Univ)
Father's education: 35.3 (5 df) P<0.001	(None, Primary) *vs* (BAC, University)
Years at Nursery school: 0.9 (3 df) P>0.10	
Birth date: 47.1 (5 df) P<0.001	1989/90 *vs* every other category

Comments on analysis of French and Maths scores.

For comparative purposes we can make a very approximate conversion of SD units to years of learning using the average, unadjusted, difference between those born at different times within the same class, i.e. of different ages. We then multiply these values by 2.6 (see Blatchford et al. 2002 for details). This gives a conversion for French of 1 SD unit = 1.0 years of progress and for Maths 1 SD unit = 0.65 years of progress. Thus for Maths the adjusted gender difference (boys – girls) would be just under 1 month and for French (girls – boys) would be just under 3 months. We would emphasise that these figures are, however, very approximate and subject to verification.

Likewise, at the start of CP, using the global score, we have an approximate conversion of 1 SD unit = 1 year of progress. This implies that girls are about 1 month ahead of boys and mothers with BAC or university degree have children about 5 months ahead of those with no primary certificate, using an analysis with the global score as response. For comparison the between-school variance is about 24% of the total, similar to the adjusted percentages – see below, but we need to be careful since the global score is not comparable with the French and Mathematics scores at start of CE2. Furthermore, the allocation fo children to schools reflects a range of complex socio-demographic factors and this makes the between-school variation at start of CP very difficult to interpret.

As before, all the comparisons are adjusted for the other variables in the model. In particular adjustment for the initial global score. For variables with just 2 categories or continuous variables, we can compare directly with the estimated standard errors. We shall contrast the maths and French results for each variable. We note that these analyses are for those who do not repeat, and so effectively omit the very lowest achieving students.

Girls make less progress than boys for maths but greater progress for French. This result occurs in other data sets in other countries. The SES of head of household is not significant for Maths but is for French, although it is difficult to identify specific differences. The percentage of foreign students has negligible effect. Private schools perform less well for Maths with no significant difference for French. For both Maths and French explicit involvement is associated with poorer results but implicit involvement is associated with better results. When we exclude each of these from the model, the relationship for the other one does not substantially change. Also, these relationships remain when we do not adjust for the pretest score. Mother's education is highly significant for both Maths and French with a clear distinction between those with no qualification up to CAP-BEC and those with BAC or university. For Father's education there are significant differences with no qualification and primary only being distinguished from both BAC and university for Maths and French. There are significant differences depending on birth date, with those born in 1989/90 progressing less than the other categories for Maths and French. There is no significant birth order effect for Maths but for French the higher the birth order the worse the progress. Being in a ZEP is associated with less progress in both Maths and French. Support after school has no significant effect. Parental language has no significant effect. Enjoyment of nursery improves progress, especially for French. Years at nursery school is not associated with progress. Having multiple year groups is associated with greater progress for maths and French. Parental situation is associated with progress with a suggestion that having 1 real and one other parent is associated with less progress than having both real parents. An increase in the behaviour score at start of CP is associated with improved performance for both Maths and French with the steepest part of the relationship for those above average on the behaviour score scale. The percentage of variance at the school level is relatively high at 26% and 22% respectively for Maths and French.

One of the tentative findings from these results is that while there are similar effects for Maths and French the social background has a more important influence for French and the school characteristics for maths.

For both Maths and French there is a significant random slope associated with the Global score, but we do not pursue this and the results for the fixed part of the model are relatively unaffected by fitting this.

We have carried out some limited bivariate analyses with Maths and French as joint responses. The results in terms of correlations at pupil and school level are as follows:

Table 9. Correlations between Maths and French scores for different models. Model A is unadjusted, model B adjusts for global score, Model C adjusts for global score and SES head of household and model D additionally adjusts for mother's education.

Level	**A**	**B**	**C**	**D**
Pupil	0.74	0.55	0.54	0.54
School	0.89	0.89	0.86	0.85

We see that the correlation at the school level is consistently high. For the unadjusted scores there is a relatively high correlation for pupils, but only a moderate correlation in terms of progress. This supports the view that learning patterns for different curriculum subjects vary.

We have also carried out a joint analysis of the probability of not repeating, French and Mathematics scores. Thus we have two continuous responses and one binary response. At the pupil level we can estimate a correlation only between Mathematics and French since those who repeat do not have these scores. At the school level, however, we can estimate the full covariance matrix and this is given in Table 10, for the unadjusted model and the model fitting just the pre-test. The correlations are little changed when other predictors are fitted and the fixed coefficients are similar to those in the separate models.

Table 10. Covariance matrices (correlations off diagonal). A is unadjusted and B is adjusted for pre-est (quadratic). Variables in order: Mathematics, French, Not repeating.

	A			**B**		
Pupil	0.822			0.455		
	0.74	0.824		0.54	0.470	
School	0.178			0.152		
	0.90	0.177		0.87	0.137	
	0.45	0.46	0.427	0.41	0.39	0.869

We see similar results to Table 10 for Mathematics and French with only moderate positive correlation at the school level between French and Mathematics scores and not repeating.

Further analyses – May 2003

*Addition of behaviour score (*SCOFACTO) *at start of CP*

This score has been converted to have a standard Normal distribution. The general effect of introducing this variable is to reduce slightly the estimates for the other effects, but does not change any of the significance levels.

For the analysis of repeating (see Table 2) the coefficient estimate is 1.10 (s.e. = 0.09) which is highly significant with a higher behaviour score being associated with a higher probability of not repeating.

For the Maths analysis (Table 5) the coefficient is 0.231 (s.e. = 0.014) which is highly significant with a higher behaviour score associated with a higher Maths score. For the French analysis (Table 5) the coefficient is 0.192 (s.e. =0.014) which is also highly significant with a higher behaviour score associated with a higher French score. There is also some evidence for 'quadratic' effects but these do not change the general conclusion.

Analysis of repeating with separate components of CP global score

Table 11 presents the results of using the separate components of the CP global score. Of the 10 components only 5 were significant in the final analysis and these are presented in Table 11. They represent Maths (numerical task SC2B0), Temporal concepts (SC3B0), Writing (SC5A0), Reading (pre-reading SC6A0), and Maths (Numbers and geometrical figures SC6B0). All these scores are Normalised and have similar coefficient values with that for temporal concepts the smallest and that for numbers and geometrical figures the largest. The gender coefficient is much smaller and now non significant. Other coefficients change somewhat, namely parental language (but significance levels remain the same), mother working (now significant), birth order (now not significant) and years at nursery school (but significance levels remain the same).

Table 11. Response is arriving in CE2 without repeating. PQL2 model using logit link.

Fixed	Estimate	s.e.
Intercept	2.570	
SC2B0	0.394	0.089***
SC3B0	0.283	0.084***
SC5A0	0.369	0.093***
SC6A0	0.462	0.088***
SC6B0	0.513	0.080***
Female	0.104	0.122

Artisan-commercial	-0.951	0.544
Higher professional	-0.065	0.584
Teacher	0.496	0.942
mid-professional	-0.631	0.540
Employee	-0.982	0.531
Skilled	-1.018	0.520
Unskilled	-1.353	0.537
Unemployed	-0.987	0.553
1 parent	-0.058	0.225
1 real 1 other parent	-0.514	0.315
Other parental situation	0.000	0.000
Private school	-0.144	0.218
Explicit involvement	0.103	0.050
1 parent French	-0.209	0.191
Neither parent French	0.696	0.203
Has lunch at home	0.178	0.128
Mother Primary cert.	0.389	0.230
Mother BEPC	0.406	0.190
Mother CAP, BEC	0.526	0.127
Mother BAC	0.898	0.244
Mother univ	0.879	0.270
Mother working	0.309	0.136*
Birth order	-0.106	0.056
In ZEP year 1	-0.059	0.185
Number of pupils in class CP	-0.020	0.019
Child settled well at nursery	0.206	0.030***
No. of nursery years 1-2	-0.295	0.563
No. of nursery years 2-3	0.271	0.518
No. of nursery years 3+	0.483	0.529
Birth 2nd quarter 1991	-0.035	0.192

Birth 3rd quarter 1991	-0.039	0.183
Birth 4th quarter 1991	0.159	0.189
Birth 1989/90	0.741	0.450
Birth 1992	-1.121	0.556
Multiple year class CP	-0.305	0.167
Behaviour score	1.078	0.093***
Random		
Level 2 variance	0.820	0.173
For variables with 2 categories, or for continuous variables, we denote significance as follows: * P<0.05: ** P<0.01: *** P<0.001.		

Analysis of Maths and French with separate components of CP global score

Table 12 presents the results of using the separate components of the CP global score. Only oral comprehension (SC4B0) was not significant for either French or Maths. The additional variables reported in these analyses are therefore general knowledge (SC1A1), Knowledge of writing (SC1A2), reading (phonological tasks SC2A1) and reading (morphosyntactic tasks SC2A2). All these variables are Normalised.

For Maths the effects of settling at school and of having multiple year groups are now not significant. The smallest significant coefficient for CP variables is for reading (phonological tasks) of 0.030 and the largest is for Maths (numerical task) of 0.241. General knowledge and reading (morphosyntactic tasks) are not significant.

For French all the CP variable coefficients are significant with the smallest being 0.043 for general knowledge and the largest 0.128 for reading (pre-reading).

The significance levels for the other variables are unchanged in these analyses.

Table 12. Maths and French responses at start of CE2. Variance components model.

Fixed	Maths		French	
	Estimate	s.e.	Estimate	s.e.
Intercept	-0.588		-0.947	
SC1A1	0.021	0.013	0.043	0.013***
SC1A2	0.056	0.012***	0.064	0.013***

SC2A1	0.030	0.013*	0.050	0.013***
SC2A2	0.020	0.012	0.049	0.012***
SC2B0	0.241	0.015***	0.106	0.015***
SC3B0	0.136	0.014***	0.109	0.015***
SC5A0	0.130	0.016***	0.107	0.016***
SC6A0	0.044	0.016**	0.128	0.016***
SC6B0	0.087	0.015***	0.097	0.015***
Female	-0.095	0.020***	0.224	0.020***
Artisan-commercial	-0.037	0.068	0.017	0.069
Higher prof.	0.052	0.066	0.130	0.067
Teacher	-0.041	0.079	0.126	0.080
mid-professional	-0.008	0.063	0.105	0.064
Employee	-0.028	0.065	0.069	0.066
Skilled	-0.054	0.063	0.032	0.064
Unskilled	-0.119	0.072	-0.051	0.073
Unemployed	-0.038	0.082	0.134	0.083
In institution	-0.119	0.420	-0.149	0.428
Foreign	-0.087	0.066	-0.098	0.067
Private school	-0.088	0.038*	-0.039	0.037
Explicit involvement	-0.015	0.008*	-0.019	0.008**
Implicit involvement	0.029	0.014*	0.046	0.015***
Mother Primary Cert.	0.086	0.057	0.063	0.057
Mother BEPC	0.064	0.042	0.051	0.043
Mother CAP, BEC	0.078	0.037	0.075	0.038
Mother BAC	0.185	0.042	0.210	0.042
Mother university	0.239	0.043	0.238	0.043
Father Primary cert.	-0.092	0.056	0.004	0.057
Father BEPC	0.033	0.044	0.087	0.044
Father CAP, BEC	0.053	0.036	0.107	0.037
Father BAC	0.114	0.044	0.185	0.045

Father university	0.079	0.045	0.213	0.046
Birth 2nd quarter 1991	0.034	0.028	0.016	0.029
Birth 3rd quarter 1991	-0.003	0.029	-0.003	0.030
Birth 4th quarter 1991	-0.007	0.030	-0.008	0.031
Birth 1989/90	-0.680	0.126	-0.671	0.128
Birth 1992	0.105	0.084	0.056	0.085
Birth order	-0.010	0.011	-0.036	0.011***
In ZEP CE1	-0.158	0.046***	-0.102	0.045**
Not supported after school	-0.035	0.033	0.009	0.033
1 parent French	0.003	0.046	-0.007	0.047
Neither parent French	-0.037	0.048	-0.105	0.049
Settled well at nursery	0.005	0.006	0.021	0.006***
No. of nursery years 1-2	0.292	0.125	0.197	0.127
No. of nursery years 2-3	0.200	0.115	0.161	0.117
No. of nursery years >3	0.221	0.116	0.159	0.118
Multiple year groups	0.041	0.026	0.074	0.025**
1 parent	-0.113	0.052	-0.108	0.053
1 real 1 other parent	-0.106	0.072	-0.165	0.073
Other situation	0.084	0.338	0.020	0.344
Behaviour score CP	0.203	0.015	0.180	0.015
Random				
Level 2 variance	0.122	0.010	0.104	0.009
Level 1 variance	0.388	0.010	0.408	0.010
Percentage of variance at level 2	20%		20%	
For variables with 1 category, or for continuous variables, we denote significance as follows: * P<0.05: ** P<0.01: *** P<0.001.				

Components of Maths and French scores.

The intercorrelations between these scores are as follows in Table 13

Table 13. Correlations among French and Maths score components and CP global score. Variables Normalised

	scomp	scode	stext	scog	scom	scon	global
scomp	1.0000						
scode	0.6990	1.0000					
stext	0.5470	0.5202	1.0000				
scog	0.5784	0.4994	0.3951	1.0000			
scom	0.6564	0.6055	0.4484	0.5611	1.0000		
scon	0.6008	0.6220	0.4155	0.5341	0.6334	1.0000	
global	0.6103	0.5383	0.4290	0.5132	0.5538	0.5236	1.0000

These correlations are all very similar. This suggests that we should expect to find similar relationships for each response variable, but these may differ in some details.

CP global score as response

Table 14 shows the results of an analysis using the CP global score as the response.

Table 14. Response is Normalised CP global score.		
Fixed	Estimate	s.e.
Intercept	-0.990	
Female	0.080	0.020***
Artisan-commercial	-0.032	0.069
Higher professional	0.108	0.068
Teacher	0.223	0.083
mid-professional	0.029	0.066
Employee	-0.055	0.067
Skilled	-0.093	0.065
Unskilled	-0.181	0.072
Unemployed	-0.339	0.080
Institution	-0.373	0.403

1 parent	0.055	0.0.050
1 real 1 other parent	0.026	0.068
Other parental situation	-0.078	0.329
1 parent French	-0.006	0.046
Neither parent French	-0.259	0.039
Mother Primary cert.	0.044	0.054
Mother BEPC	0.256	0.040
Mother CAP, BEC	0.166	0.036
Mother BAC	0.393	0.040
Mother university	0.525	0.041
Father Primary cert.	0.131	0.053
Father BEPC	0.185	0.041
Father CAP, BEC	0.202	0.034
Father BAC	0.302	0.043
Father university	0.420	0.045
Birth order	-0.028	0.010**
Child settled well at nursery	0.093	0.006***
No. of nursery years 1-2	0.197	0.116
No. of nursery years 2-3	0.239	0.107
No. of nursery years 3+	0.298	0.108
Birth 2nd quarter 1991	-0.171	0.029
Birth 3rd quarter 1991	-0.285	0.029
Birth 4th quarter 1991	-0.478	0.030
Birth 1989/90	-0.735	0.106
Birth 1992	-0.239	0.084
Random		
Level 2 variance	0.134	0.011
Level 1 variance	0.566	0.012
For variables with 2 categories, or for continuous variables, we denote significance as follows: * P<0.05: ** P<0.01: *** P<0.001.		

The following are significance tests for Table 15.

Variable: overall χ^2	***Pairwise differences significant at 5% level***
SES of head: 79.5 (9 df) P<0.001	Agricultural *vs* (teacher, unskilled, unemployed); teacher *vs* (employee, skilled, unskilled, unemployed); higher professional *vs* (skilled, unskilled, unemployed).
Parental situation: 1.3 (3 df) P>0.05	
Parental language: 43.8 (2 df) P<0.001	(Both French, 1 parent French) *vs* both other;
Mother's education: 212.4 (5 df) P<0.001	(None, Primary), *vs* (BEPC, CAP-BEC, BAC, Univ); (BEPC, CAP-BEC, BAC) *vs* Univ.
Father's education: 92.0 (5 df) P<0.001	(None, Primary) *vs* Univ; (CAP-BEC, BEPC, BAC) *vs* univ
Years at Nursery school: 12.8 (3 df) P<0.01	
Birth date: 382.4 (5 df) P<0.001	1st quarter 1991 *vs* (2-4th quarters, 1989), 2nd quarter 1991 *vs* (3-4th quarters, 1989), 3rd quarter 1991 *vs* (4th quarter, 1989), 4th quarter 1991 *vs* 1989.

Note that parental situation is not significant, but parental language is. There are also important differences due to SES so that once the global score is adjusted for, the effect of SES (on French and Maths) is reduced. Similarly mothers and fathers education are more important. There is also now a significant effect of years at nursery school. The birth date effect does not seem to be easily explained. An analysis of the raw means for this variable shows a similar ordering of the mean global score. For comparison we give these means for the global CP score, French and Maths in Table 16.

Table 16. means (s.e.) for Global score at CP, French and Maths at CE2.

Birth date	Global score	Maths	French
1st quarter 1991	0.27 (0.02)	0.16 (0.02)	0.12 (0.02)
2nd quarter 1991	0.09 (0.02)	0.08 (0.02)	0.06 (0.02)
3rd quarter 1991	-0.09 (0.02)	-0.08 (0.02)	-0.06 (0.02)
4th quarter 1991	-0.22 (0.02)	-0.16 (0.02)	-0.11 (0.02)
1989/90	-0.99 (0.08)	-1.29 (0.13)	-1.34 (0.13)
1992	0.32 (0.09)	0.46 (0.10)	0.46 (0.10)

References

Goldstein, H. (2003). *Multilevel Statistical Models. Third edition.* London, Edward Arnold: