



**HAL**  
open science

# Regards croisés sur l'école primaire française : éléments d'évaluation macro et micro

Bruno Suchaut

► **To cite this version:**

Bruno Suchaut. Regards croisés sur l'école primaire française : éléments d'évaluation macro et micro. Education. Université de Bourgogne, 2007. tel-00199857

**HAL Id: tel-00199857**

**<https://theses.hal.science/tel-00199857>**

Submitted on 19 Dec 2007

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

**Université de Bourgogne**  
**Institut de recherche sur l'éducation**

**Regards croisés sur l'école primaire française**  
**Eléments d'évaluation macro et micro**

*Habilitation à Diriger les Recherches*  
*en Sciences de l'Education*

**Bruno Suchaut**

Soutenue le 29 novembre 2007 à Dijon

Composition du jury :

Pascal Bressoux : Professeur en Sciences de l'Education, Université Pierre Mendès-France,  
Grenoble

Thierry Chevaillier : Professeur en Sciences de l'Education, Université de Bourgogne

Marcel Crahay : Professeur en Sciences de l'Education, Université de Genève

Marc Demeuse : Professeur en Sciences de l'Education, Université de Mons-Hainaut

Michel Fayol : Professeur de Psychologie, Université Blaise Pascal, Clermont-Ferrand

Georges Solaux : Professeur en Sciences de l'Education, Université de Bourgogne

## Sommaire

Introduction.....	p.1
<b>Chapitre 1 : Un cadrage général sur l'efficacité et l'efficience :</b>	
<b>coûts et résultats de l'école primaire.....</b>	<b>p.12</b>
I	Le niveau moyen d'acquisition des élèves et l'efficience qualitative.....p.13
II	Coûts et financement de l'enseignement primaire en France.....p. 20
II.1	<i>Une perspective globale.....</i> p. 20
II.2	<i>La situation de l'enseignement primaire dans le secteur scolaire.....</i> p. 23
II.3	<i>Les facteurs explicatifs de l'évolution des coûts de l'école primaire.....</i> p. 27
II.4	<i>Les coûts unitaires dans une perspective comparative internationale.....</i> p. 32
<b>Chapitre 2</b>	
<b><i>Organisation scolaire et inégalités entre élèves.....</i></b> p 34	
I	Les problèmes associés à l'approche comparative.....p 36
II	Inégalités sociales de réussite et caractéristiques des systèmes éducatifs.....p 42
II.1	<i>Les Inégalités de performances.....</i> p 42
II.2	<i>Des différences de niveau moyen et d'inégalités sociales entre pays.....</i> p 49
II.3	<i>Contextes nationaux et inégalités sociales .....</i> p 52
II.4	<i>Sélectivité et différenciation au sein des systèmes éducatifs et inégalités .....</i> p 53
Conclusion.....	p 59
<b>Chapitre 3</b>	
<b><i>Le contexte de la classe : constitution et effets sur les élèves .....</i></b> p 60	
I	La fabrication du contexte : procédures de constitution des classes à l'école élémentaire.....p. 64
I.1	<i>La priorité aux intérêts individuels.....</i> p. 66
I.2	<i>Le poids des usages.....</i> p 68
I.3	<i>Le rôle du directeur d'école.....</i> p. 73

II	Efficacité pédagogique des classes à cours double .....	p. 77
II.1	<i>Les travaux antérieurs sur l'efficacité des classes à cours multiple</i> .....	p. 78
II.2	<i>Cours simple versus cours multiple : quelle efficacité globale?</i> .....	p. 80
II.3	<i>L'affectation des élèves dans les classes</i> .....	p. 86
III.	L'influence de la composition sociale.....	p. 94
III.1	<i>Des différences globales entre contextes</i> .....	p. 95
III.2	<i>Les modèles de progressions des élèves au niveau agrégé</i> .....	p. 100
III.3	<i>Les modèles multiniveau intégrant variables individuelles et contextuelles</i> .....	p. 103
III.4	<i>Les effets d'interaction</i> .....	p. 106
III.5	<i>L'exploration des effets de contexte</i> .....	p. 109
	Conclusion .....	p. 114
 <b>Chapitre 4</b>		
	<b><i>Les apprentissages des élèves : structure et évolution</i></b> .....	p. 115
I.	L'évolution globale des acquisitions à l'école élémentaire .....	p. 116
II.	Structure et évolution des compétences au cycle III.....	p. 122
II.1	<i>Les compétences à l'entrée au CE2</i> .....	p. 122
II.1.1	<i>Une approche basée sur les relations statistiques entre items</i> .....	p. 125
II.1.2	<i>L'analyse en variables latentes</i> .....	p. 128
II.1.3	<i>Structure relationnelle et hiérarchique des compétences</i> .....	p. 132
II.1.4	<i>Conclusion sur les acquisitions à l'entrée au cycle III</i> .....	p. 137
II.2	<i>Evolution des compétences au cycle III</i> .....	p. 138
III	Les liens entre capacités cognitives, compétences scolaires et origine sociale.....	p. 145
III.1	<i>Mémoire de travail et capacités cognitives</i> .....	p. 146
III.2	<i>La description des données</i> .....	p. 148
III.3	<i>Le poids des capacités cognitives et de l'origine sociale sur les progressions scolaires</i> .....	p. 150
	<b>Conclusion générale</b> .....	p. 156
	Bibliographie.....	p. 160
	Liste des tableaux.....	p. 177
	Liste des graphiques.....	p. 181

## **Introduction**

*«Une HDR en sciences de l'éducation ne peut pas consister en une histoire de vie professionnelle ou en une autobiographie raisonnée. Elle doit illustrer les qualités de recherche du candidat attestées par des publications reconnues au sein de la section - articles, ouvrages, rapports de recherche - et par son engagement dans la vie universitaire de la 70<sup>e</sup> section (participation à des colloques, membre scientifique de revues publiant des recherches en ou sur l'éducation). Une HDR atteste que le candidat, depuis la soutenance de sa thèse, a approfondi, élargi et théorisé un domaine de recherche dont il devient l'une des figures emblématiques. Deux types d'HDR sont susceptibles d'être soutenues en sciences de l'éducation.*

*1. Lorsque le candidat possède un ensemble important d'articles dans un domaine nettement circonscrit, voire lorsqu'il a produit un ouvrage mettant en perspective ses recherches, l'HDR se présente fréquemment comme la synthèse et le développement théoriques de ses travaux.*

*2. Lorsque le candidat possède un ensemble de publications au contenu épars, l'HDR ne peut pas prendre la forme d'une note de synthèse de cette diversité. Elle s'apparente alors à un nouvel ouvrage, synthèse du domaine de recherches choisi par le candidat au sein de ses diverses investigations. L'étude conduite ne peut cependant faire l'impasse de publications importantes préexistantes à l'HDR»<sup>1</sup>.*

En référence à ces conseils énoncés par la 70<sup>ème</sup> section du C.N.U., et au regard de notre parcours de recherche depuis la publication de la thèse, notre Habilitation à Diriger les Recherches prendra la forme d'une mise en perspective de nos recherches les plus récentes avec la présentation de travaux originaux. Un fil conducteur peut en effet clairement se dégager de nos publications produites depuis la soutenance de la thèse : l'évaluation des systèmes éducatifs avec un accent mis particulièrement sur le système éducatif français et le fonctionnement de son école primaire.

---

<sup>1</sup> Extrait du document de la section n°70 de la Conférence Permanente du Conseil National des Universités : conseils généraux aux candidats à la qualification (<http://www.cpu.fr/Cp-cnu/Docs/section70.pdf>).

La thèse, soutenue en janvier 1996 (Suchaut, 1996a), portait sur l'utilisation du temps d'enseignement et ses effets sur les acquisitions des élèves. Deux dimensions principales ont été explorées dans ce travail. La première soulignait la grande variété des pratiques dans la gestion du temps d'un enseignant à l'autre et ces incidences sur les acquisitions des élèves (Suchaut, 1996b). La seconde dimension a concerné un aspect plus qualitatif du temps scolaire à travers une évaluation d'une expérimentation d'activités musicales en grande section de maternelle et la mise en évidence d'effets différés et transversaux de ces activités sur les progressions des élèves (Mingat, Suchaut, 1996).

Depuis plus d'une dizaine d'années, l'ensemble de nos travaux a concerné le fonctionnement des systèmes éducatifs avec des approches différentes explorées sur la base de plusieurs objets de recherche. De façon générale, deux perspectives d'analyse complémentaires ont été mobilisées. La perspective macro tout d'abord qui utilise l'approche comparative pour identifier des relations entre les contextes éducatifs nationaux et les résultats en termes de compétences des élèves (Duru-Bellat, Mons, Suchaut, 2003 ; Duru-Bellat, Suchaut, 2005a), mais aussi en termes de développement quantitatif et qualitatif des systèmes dans les pays en développement (Mingat, Suchaut, 2000). La perspective micro se centre quant à elle sur l'étude de la fonction de production de l'élève en mesurant les effets des différents facteurs liés à l'élève lui-même, à son environnement social et familial et à son contexte scolaire sur ses résultats en termes d'acquisitions et de carrières (Suchaut, 2003a). Cette perspective micro recouvre elle-même diverses dimensions qui ont été explorées à des degrés divers à travers plusieurs travaux.

En premier lieu, une série de publications s'est attachée à évaluer des dispositifs destinés à apporter une aide aux apprentissages des élèves, soit au sein des établissements par le biais d'actions nationales (Danner, Duru-Bellat, Le Bastard-Landrier, Suchaut, 2001) ou décidées localement (Piquée, Suchaut, 2004), soit en dehors du temps scolaire avec des dispositifs d'accompagnement (Piquée, Suchaut, 2002) ou plus simplement l'action des parents en matière d'aide scolaire (Jarousse, Leroy-Audouin, Suchaut, 1999). D'autres travaux se sont centrés de manière spécifique sur l'efficacité de certaines pratiques d'enseignement, principalement au niveau de l'école élémentaire (Suchaut, 1998 ; Le Bastard-Landrier, Suchaut, 2000).

En second lieu, l'exploration du contexte scolaire a fait l'objet de plusieurs recherches en se centrant sur deux dimensions. La première s'attache à étudier les modes de groupements des élèves au sein des classes en examinant les stratégies et les choix réalisés par les enseignants pour constituer les classes (Leroy-Audouin, Suchaut, 2005, 2007a) et en mesurant les effets de ces modes de groupements sur les acquisitions des élèves (Leroy-Audouin, Suchaut, 2007b). La seconde dimension complète la précédente en se focalisant sur la composition

sociale des classes et ses incidences, toujours en termes d'acquisitions scolaires, sur des élèves d'origine sociale variée (Duru-Bellat, Le Bastard-Landrier, Piquée, Suchaut, 2004 ; Duru-Bellat, Suchaut, 2005b).

Enfin, des travaux récents et en cours de développement, explorent de manière plus approfondie la variable dépendante centrale qui gouverne toutes nos analyses, à savoir les acquisitions des élèves. Il s'agit de les envisager ici dans une logique dynamique en examinant comment les apprentissages des élèves se construisent dans le temps et comment ils se structurent entre eux à travers l'association de nombreuses compétences (Morlaix, Suchaut, 2007a, 2007b ; Barrouillet, Camos, Morlaix, Suchaut, 2007).

Ce panorama de nos travaux pose de suite plusieurs questions à la fois sur leurs fondements théoriques et méthodologiques et sur leur articulation, la première question étant de nature épistémologique. L'ensemble des travaux conduits se fonde sur l'utilisation du paradigme économique en matière d'évaluation (De Ketele, 1993). Ce paradigme a pour origine lointaine la théorie du capital humain (Becker, 1964) ; il s'agissait alors d'apprécier les effets de l'éducation sur les comportements privés et sociaux des agents et d'étudier les relations entre le système éducatif et le système de production. Mais ce sont les travaux relatifs à l'efficacité interne des systèmes qui ont permis de développer progressivement ce paradigme économique. Le rapport Coleman (Coleman, 1965) a initié tout un courant de recherches dans lesquels les inputs éducatifs (ressources matérielles, caractéristiques personnelles et familiales de l'élève, caractéristiques de l'enseignant, contexte pédagogique...) sont intégrés à des fonctions de production classiques permettant d'expliquer la qualité de l'instruction dispensée dans les écoles. Depuis cette époque, de nombreux travaux ont été produits en utilisant cette fonction de production ; en France c'est seulement à partir des années quatre-vingt que cette approche s'est développée, principalement grâce aux recherches conduites au sein de l'Irédu (Institut de recherche sur l'éducation).

La variabilité des acquisitions des élèves est la notion centrale de ce paradigme économique et deux concepts clés s'y rattachent : l'efficacité et l'équité. Si ces deux concepts guident effectivement l'ensemble de nos recherches, les outils méthodologiques utilisés dans les analyses contribuent également à ce que la discipline économique nourrisse au niveau théorique notre démarche. La sociologie de l'éducation est pour sa part très présente et dépasse la seule problématique de l'évaluation, notamment par l'étude quasi systématique dans nombre de nos travaux, des déterminants sociaux et scolaires de la réussite des élèves (Duru-Bellat, 2002). Nos recherches récentes, centrées sur les apprentissages, s'appuient quant à elles sur une autre discipline, à savoir la psychologie cognitive, qui permet de fournir un cadre théorique à nos analyses. Il est donc clair, que nous revendiquons une approche

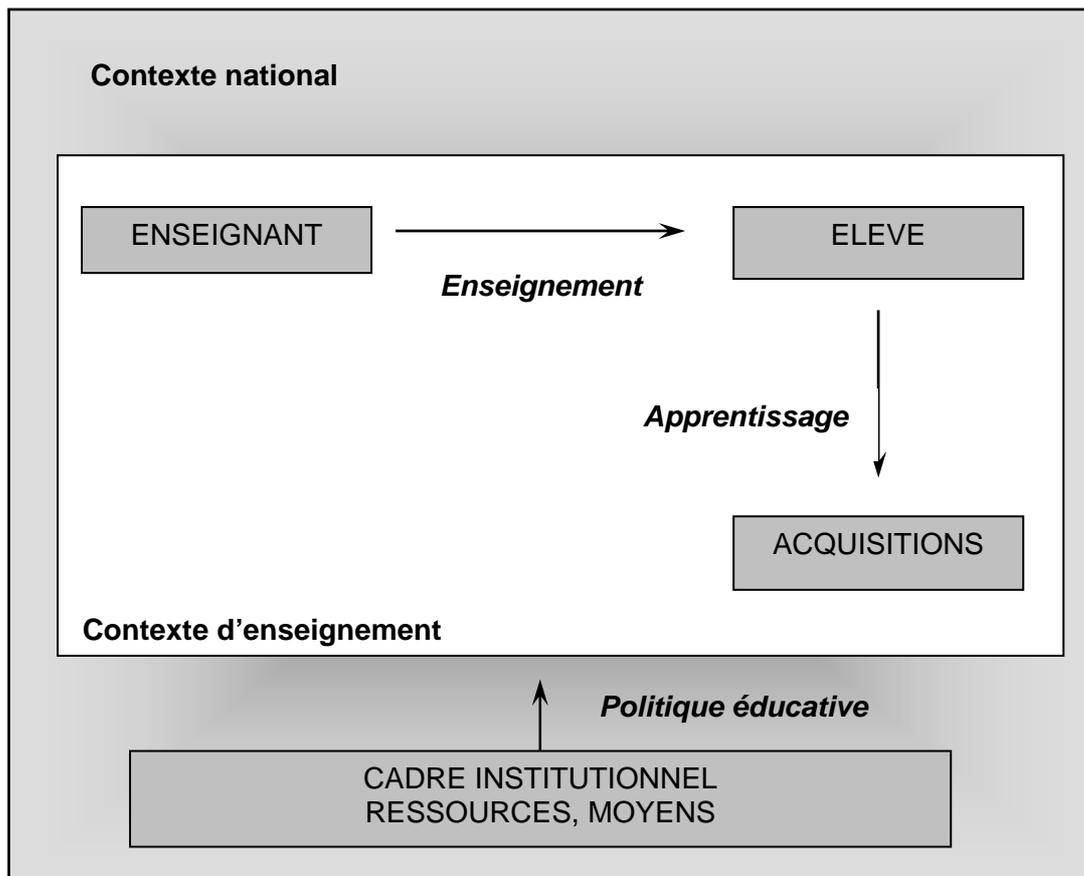
pluridisciplinaire dans nos travaux, celle-ci reflétant une partie de la diversité des champs théoriques des sciences de l'éducation dans le paysage français.

Un deuxième questionnement concerne l'approche méthodologique utilisée dans l'ensemble de nos travaux. Les évaluations réalisées dans le domaine de l'éducation relèvent traditionnellement de deux approches : l'une qualifiée d'interne qui donne une place prépondérante aux acteurs, l'autre qualifiée d'externe qui se focalise sur l'analyse objective d'éléments factuels en adoptant un regard extérieur aux situations observées (Duru-Bellat, Mingat, 1993, 1998a). La distinction entre les deux approches renvoie à une opposition plus large relative aux fonctions mêmes de l'évaluation (Ardoino, Berger, 1986). Dans un cas, on privilégie le vécu, le sens et la subjectivité de l'objet, alors que dans l'autre cas ce sont la rigueur, la mesure, et l'objectivité qui gouvernent la démarche (Vial, 1997). Il est manifeste que nous nous situons nettement dans le domaine de l'évaluation externe à laquelle sont associées systématiquement les dimensions quantitative et comparative. Une caractéristique majeure de cette approche est qu'elle nécessite, de fait, une simplification de la réalité, ceci pouvant par ailleurs être considéré comme un principe méthodologique de base de l'évaluation des systèmes éducatifs (Thélot, 1993).

A un niveau méthodologique plus précis, des problèmes se posent également sur la nature même des analyses statistiques qui sont mobilisées dans les travaux utilisant le paradigme économique. Si les modèles de régression multivariée sont un progrès considérable par rapport aux tableaux statistiques simples faisant intervenir des comparaisons de fréquences ou de moyennes, il n'en reste pas moins qu'il doivent faire l'objet d'une grande attention dans leur formalisation mathématique pour fournir des estimations robustes et fiables. Cela est particulièrement vrai en ce qui concerne les variables qui rendent compte de l'environnement scolaire au sens large. Il a été clairement montré que ces variables de contexte étaient mesurées imparfaitement quand elles sont traitées avec des modèles de régression simple par les moindres carrés ordinaires (Bressoux, Coustère, Leroy-Audouin, 1997 ; Bressoux, 2000), ce qui a été le cas dans les études françaises réalisées jusqu'au milieu des années quatre vingt dix. Nombre de résultats concernant des variables contextuelles ont ainsi pu être mesurés imparfaitement ce qui est d'autant plus regrettable que ces variables (nombre d'élèves dans la classe, niveau de qualification de l'enseignant, modes de groupements des élèves...), sont des variables centrales de politique éducative. Une solution à ces problèmes statistiques réside alors dans l'utilisation de modèles multiniveaux qui permettent de séparer dans les analyses les parts de variance des différents niveaux d'analyse mobilisés : élève, classe, école.... Ces modèles ont été systématiquement utilisés dans nos recherches depuis presque une dizaine d'années.

Le troisième point qui interroge nos travaux est la dimension politique liée à l'évaluation du système éducatif. En France, l'évaluation a connu un développement progressif depuis le début des années quatre-vingt, cette évolution découle, certes de la volonté politique d'améliorer la qualité des services éducatifs, mais elle est également liée aux recherches qui, depuis une trentaine d'années, apportent régulièrement des informations factuelles sur les résultats de l'école. L'évaluation est en effet devenue un axe de recherche à part entière en sciences sociales et une sociologie de l'évaluation a vu le jour. L'évaluation pose la question de la relation entre science et politique et il n'est pas toujours aisé de distinguer dans les productions ce qui relève de jugements fondés sur une idéologie de ce qui ressort des faits objectifs (Solaux, 2000a). Certains résultats relevant du concept d'efficacité peuvent même être manipulés par les acteurs en fonction de leurs représentations et priorités politiques ce qui tend à montrer que l'évaluation est un processus sociologiquement complexe et il ne suffit pas d'évaluer pour que des améliorations se produisent (Demaillly, 2000a). Une réflexion sur les résultats produits dans les études et recherches est donc nécessaire ; celle-ci concerne leur utilité, leur portée et leur utilisation, tant au niveau politique (Lapostolle, Suchaut, 2007), qu'au niveau des acteurs impliqués dans les actions ou dispositifs soumis aux évaluations (Genelot, Suchaut, 2000 ; Suchaut, 2004 ; Suchaut, 2005a). De manière plus précise, ces résultats nous interrogent sur l'efficacité et l'équité de la politique éducative en termes d'actions conduites à destination de publics spécifiques (Suchaut, 2005b).

A partir de recherches récentes et en cours, nous nous interrogerons dans ce document sur la qualité de notre système éducatif, et plus particulièrement de notre école primaire en mobilisant des analyses qui présentent comme originalité une complémentarité visible sur deux plans. Sur le plan méthodologique et conceptuel tout d'abord puisque plusieurs niveaux d'analyse statistique sont pris en compte : le pays, l'établissement, la classe et l'élève ; cela permet de percevoir une image globale du fonctionnement de l'école et de mieux comprendre comment les différentes facettes du contexte scolaire interviennent sur les apprentissages des élèves. La complémentarité se situe également sur le plan théorique puisque, comme nous l'avons déjà évoqué, plusieurs disciplines sont réunies pour l'étude d'un même objet : le fonctionnement du système éducatif. Le schéma suivant formalise notre logique de recherche en indiquant l'ensemble des éléments associés aux acquisitions des élèves, ceux-ci étant considérés comme l'indicateur principal de la qualité de l'école.



L'évaluation intervient ici à tous les niveaux : de l'approche macroscopique par une analyse globale de la situation du système éducatif, à une approche microscopique qui se focalise sur les apprentissages des élèves. Au degré le plus global, le contexte national du pays (économique, social et culturel) est le premier facteur qui contribue à expliquer le niveau de développement quantitatif du système mais aussi le niveau d'acquisition de ses élèves. Au-delà de ce contexte général du pays, les modalités d'organisation des activités d'enseignement à travers les orientations politiques jouent un rôle non négligeable sur les performances moyennes et sur les inégalités scolaires (Duru-Bellat, Mons, Suchaut, 2003). C'est grâce aux comparaisons internationales, dont les vertus heuristiques en sciences humaines ne sont plus à démontrer, qu'il devient possible d'évaluer l'influence de ce qui, de fait, ne varie pas au sein d'un pays donné. L'approche comparative introduit des sources de variation au niveau des facteurs contextuels (richesse économique, structures sociales...) mais aussi, et cela est central dans notre approche, au niveau de l'organisation du système scolaire (structure, organisation et régulation).

Le cadre institutionnel et budgétaire exerce une influence sur le mode d'organisation des activités d'enseignement. Il s'agit ici d'identifier comment les instructions officielles, les réglementations mais aussi les moyens financiers et humains, se traduisent concrètement sur le

terrain. Ces relations, qui concernent les mécanismes de gestion, de pilotage et d'évaluation des acteurs et des structures (les classes, les écoles) ne sont pas toujours facilement identifiables, du fait d'une certaine opacité du fonctionnement de notre système qui rend, en outre, sa régulation difficile (Duru-Bellat, Mingat, 1998b). L'organisation centralisée du système éducatif français rend a priori peu probable, dans l'enseignement public au moins, toute variété. Pourtant, au-delà de l'apparente unité qui existe dans les textes officiels, tant du point de vue des contenus d'enseignement dispensés que des structures, plusieurs travaux ont montré qu'il existait une profonde diversité sur le plan des curricula et des pratiques pédagogiques in situ, ayant pour conséquence des inégalités de réussite et de carrière scolaires au niveau des élèves (Attali, Bressoux, 2002).

Un second niveau explore l'environnement dans lequel se construisent au quotidien les apprentissages des élèves. Le contexte d'enseignement s'appréhende, à l'échelle de l'établissement scolaire ou au niveau de la classe et de l'enseignant (Bressoux, 1994). Ces effets de contexte, mis en évidence en France dans les années quatre-vingt (Mingat, 1983, 1984 ; Duru-Bellat, Mingat, 1988) témoignent de l'influence potentielle de nombreux facteurs pouvant être répartis en plusieurs catégories. Il existe tout d'abord des variables directement liées à la politique éducative et des ressources financières associées qui rendent compte de la manière dont est organisé le contexte d'un point de vue structurel, c'est par exemple le cas du nombre d'élèves dans la classe (Meuret, 2001). La seconde catégorie regroupe des variables concernant la composition du groupe classe et sa mixité. Cette dernière peut s'exprimer au niveau social (tonalité sociale du groupe d'élèves), académique (niveau scolaire des élèves), mais aussi structurelle (groupements d'élèves d'âge et de sections différent), cela contribuant à constituer des contextes d'apprentissages plus ou moins hétérogènes (Duru-Bellat, Mingat, 1997 ; Duru-Bellat, Suchaut, 2005b ; Leroy-Audouin, Suchaut, 2005). Un troisième ensemble d'éléments traduit l'influence des situations pédagogiques et didactiques au sein de la classe. De multiples aspects peuvent être envisagés dans cet ensemble, mais les recherches françaises, peu nombreuses, se sont limitées seulement à quelques-uns (Attali, Bressoux, 2002) dont les plus explorés sont la planification de l'enseignement (Gauthier, 1997), la gestion du temps scolaire (Altet, Bressoux, Bru, Lambert, 1994, 1996 ; Suchaut, 1996 ; Aubriet-Morlaix, 1999 ; Arnoux, 2004) et l'implication des élèves (Altet, Bressoux, Bru, Lambert, 1994, 1996). Le contexte d'apprentissage de l'élève dépend aussi du comportement social de l'enseignant, le groupe classe constitue un terrain de recherche riche pour la psychologie sociale qui commence à livrer des éléments utiles à l'action, notamment par l'étude des situations de comparaison sociale fréquentes en milieu scolaire (Toczek, Martinot, 2004), du jugement évaluatif de l'enseignant (Bressoux, Pansu, 2003). La psychologie cognitive insiste aussi sur l'importance du lien entre l'organisation des relations sociales en classe et les acquisitions des élèves et notamment sur le fait que les enseignants doivent multiplier les « feed-back » en vue de faire prendre conscience aux élèves des procédures

qu'ils utilisent (Fayol, 1989). Plus largement, l'étude des interactions maître-élèves et des comportements des acteurs demande une complémentarité d'approches disciplinaires pour mieux en saisir la complexité (Altet, 1994).

Le dernier aspect du contexte scolaire concerne l'influence des variables relatives aux caractéristiques personnelles à l'enseignant relevant de facteurs individuels comme sa motivation, son engagement professionnel ou encore ses capacités relationnelles. Ce domaine est largement inexploré dans le paysage français où l'on se limite à faire des hypothèses sur l'importance de ces variables dans l'explication de l'effet-maître (Mingat, 1991). Il est donc important que l'on puisse rechercher, ce qui dans le comportement de l'enseignant, conduit à cette efficacité en observant de façon méthodique son travail et surtout les relations qu'il établit avec ses élèves en privilégiant une conception interactive de l'effet-maître (Bressoux, 2002).

L'observation du contexte est une phase délicate pour toute recherche empirique car les activités scolaires sont par nature très complexes. A priori, tout ce qui se déroule lors d'une situation d'enseignement peut être digne d'observation et d'analyse : l'activité du maître, celle des élèves, les supports didactiques, l'organisation effective du travail scolaire.... Le chercheur devra donc effectuer des choix pour ne se limiter qu'à certains aspects du contexte d'apprentissage et cela explique pourquoi les travaux ne ciblent que des dimensions précises. La mesure de l'efficacité pédagogique est alors associée à de nombreux facteurs qu'il est difficile d'appréhender, d'autant plus que des travaux ont mis en avant la grande variabilité des pratiques d'enseignement pour un même enseignant (Bru, 1991 ; Crahay, 1989). Il est donc d'autant plus utile de disposer d'un grand nombre de recherches sur les facteurs d'efficacité du contexte scolaire et sur les pratiques d'enseignement pour pouvoir dégager les résultats les plus significatifs et nourrir la prise de décision politique ; à ce titre, la recherche française, en référence aux pays anglo-saxons est particulièrement déficitaire et elle ne fournit que des pistes très limitées en matière d'action.

Enfin, le schéma précédent matérialise la relation entre l'élève et ses acquisitions ; on dépasse ici la conception instrumentale des acquisitions des élèves adoptée dans les travaux relevant du paradigme « processus-produits » (Doyle, 1986). Cette relation constitue le niveau d'analyse le plus fin puisqu'il se focalise sur les mécanismes et la structure des acquisitions des élèves. Ce champ de recherche peut être considéré comme novateur dans la mesure où l'approche économique en évaluation considère les acquisitions des élèves uniquement comme une variable de résultat qui produit de la variance, sans interrogation profonde sur leur nature. Ces acquisitions scolaires renvoient à des connaissances et des compétences très variées qui ne jouent pas le même rôle dans la scolarité des élèves. Les apprentissages ne se conçoivent pas de façon isolée mais selon une logique d'interdépendance, il est alors utile de rechercher les liens qui les unissent pour mieux comprendre comment les acquisitions des élèves se construisent.

Celles-ci s'inscrivent également dans un processus temporel et il est essentiel de connaître les principes généraux qui gouvernent cette évolution (Morlaix, Suchaut, 2007a).

Les différents éléments exposés dans le schéma précédent constituent la structure de notre problématique, il s'agit de porter un large regard sur l'évaluation de l'école à travers les recherches empiriques récentes. Alors que la culture de l'évaluation est bien présente aux différents niveaux de responsabilité du système, il apparaît légitime de s'interroger sur les éléments sur lesquels la politique éducative peut s'appuyer pour relever le double défi auquel notre système est actuellement confronté. Ce défi concerne principalement deux concepts chers aux économistes et aux sociologues de l'éducation : efficacité et équité. Sur le plan des compétences des élèves, les récentes évaluations internationales (PIRLS<sup>2</sup>, 2001 ; PISA<sup>3</sup>, 2000, 2003) situent la France à une position moyenne en comparaison aux autres pays développés. Au niveau national, les évaluations nationales à l'entrée en sixième révèlent chaque année qu'une proportion d'élèves non négligeable, de l'ordre de 10 à 15% environ, ne maîtrise pas les compétences minimales requises pour qu'une scolarité correcte au collège soit assurée et qu'une qualification soit garantie à la sortie du système puisque 160 000 jeunes quittent l'école en échec avéré. Ces constats récents (Forestier, Thélot, Emin, 2007) interrogent sur la pertinence de la politique éducative des vingt cinq dernières années pendant lesquelles l'effort budgétaire pour le secteur éducatif s'est accru sans que les objectifs qualitatifs et quantitatifs soient totalement atteints. De nombreuses mesures institutionnelles ont pourtant vu le jour afin de lutter contre l'échec scolaire et favoriser l'égalité des chances entre les élèves dont le plus marquant est probablement la mise en place des zones d'éducation prioritaire. Bien que la production scientifique reste encore très insuffisante dans notre pays pour porter des jugements irrévocables, les évaluations disponibles concernant ces dispositifs donnent lieu très souvent à des conclusions en demi-teinte qui amènent à s'interroger sur l'efficacité des actions au regard des moyens qui sont mobilisés (Meuret, 1994 ; Moisan, Simon, 1997 ; Suchaut, 2005b).

La multiplicité des dispositifs peut aussi donner l'impression que la politique éducative manque de cohérence pour cibler les priorités et qu'une certaine errance se manifeste quant il s'agit d'apporter des réponses. La volonté de faire acquérir une formation de qualité à tous, indifféremment des caractéristiques sociales, scolaires et culturelles se heurte à des déterminismes parfois marqués quand on observe les carrières scolaires des élèves (Duru-Bellat, 2002). Une des clés utiles à la compréhension de cet échec est inhérente au fonctionnement interne du système et fait intervenir la dimension du pilotage. Un exemple incontestablement significatif de ces dernières décennies est la politique des cycles à l'école primaire intégrée à la loi d'orientation de 1989 qui, sans réformer les structures, s'est appuyées sur l'existant en proposant un mode de fonctionnement qui devait conduire à un

---

<sup>2</sup> PIRLS : Progress in International Reading Literacy Study

<sup>3</sup> PISA : Programme for International Student Assessment

changement profond des pratiques pédagogiques en vigueur dans les écoles. Or, près de vingt ans plus tard, on ne peut que confirmer les conjectures précédentes et constater l'échec de cette politique volontaire et ambitieuse, tant en ce qui concerne sa mise en place effective qu'au niveau de la production de résultats tangibles (Ferrier, 1998). Cette difficulté à traduire la volonté politique en actions concrètes et durables est sans doute un indice qui demande à être analysé pour comprendre les dysfonctionnements de notre système, repérer les leviers d'action possibles et faciliter son évaluation (Solaux, 2000b ; Lapostolle, Suchaut, 2007).

Notre démarche présente une difficulté majeure due à la nature même de notre objet de recherche. Celui-ci, bien que limité aux premiers cycles du système scolaire, est potentiellement très vaste. Il ne s'agit certainement pas de dresser un bilan exhaustif des résultats de l'école, mais plutôt de montrer comment la recherche en éducation, sur la base de travaux d'évaluation récents ou originaux, peut participer à dégager des pistes de réflexion utiles à la politique éducative. Le pari est bien de mettre en perspectives des résultats produits à des niveaux d'observation et d'analyse différents pour faire émerger les vraies questions auxquelles notre système est confronté aujourd'hui. L'efficacité, l'efficience et l'équité de notre école seront examinées sous l'angle de plusieurs questions, traitées soit de façon centrale, soit de manière plus implicite, dans toutes les recherches qui seront mobilisées dans ce document. L'organisation, la structure et la régulation du système, la mixité sociale, le rôle des acteurs, la construction des apprentissages des élèves, sont autant de sujets qui, quand on les articulent, donnent une image relativement complète des enjeux majeurs de la politique éducative actuelle.

La structure de notre document suit la logique des constats et réflexions précédentes en articulant des analyses macro et micro qui permettent d'identifier les principaux facteurs responsables des différences d'acquisitions entre les élèves. Dans un premier chapitre nous adopterons une perspective macro en mobilisant des indicateurs globaux de coûts et de résultats de notre système éducatif. L'approche comparative sera utilisée, sur la base de données nationales et internationales pour mieux mettre en valeur les forces et les faiblesses de notre école primaire. Dans le second chapitre, nous examinerons comment l'organisation globale du système éducatif peut participer à l'explication des différences entre élèves, notamment sur le plan des inégalités sociales de réussite. Dans le troisième chapitre, nous nous centrerons sur un aspect de l'environnement scolaire qui traduit la variété du contexte d'apprentissage chez les élèves, à savoir la composition du groupe classe dans sa dimension sociale et pédagogique. Après avoir rappelé les principes de constitution des classes dans les écoles, il s'agira plus précisément d'étudier comment la structure sociale et scolaire du groupe d'élèves influe sur les progressions scolaires. Dans un quatrième chapitre, sur la base d'analyses longitudinales, nous nous focaliserons sur les apprentissages des élèves en étudiant comment ceux-ci se construisent et se structurent au cours de l'école primaire. En guise de

conclusion, nous prolongerons les constats établis dans les chapitres précédents en dégagant les éléments qui nous semblent aujourd'hui être les plus urgents à prendre en compte en termes de politique éducative. Les questions relatives au pilotage du système, à sa gestion pédagogique seront abordées, sans omettre son évaluation, celle-ci pouvant être considérée comme le mécanisme de régulation prioritaire.

*Chapitre 1*  
*Un cadrage général sur l'efficacité et l'efficience :*  
*coûts et résultats de l'école primaire*

Ce chapitre a pour objectif de fournir un cadrage général sur la situation de notre école primaire en présentant quelques éléments sur son efficacité et son efficience. Nous nous situons ici dans une perspective d'évaluation globale en mobilisant des données agrégées au niveau du pays avec les limites inhérentes à ce genre d'approche, notamment celle de masquer les disparités territoriales, sociales et sexuées au sein du système. L'interrogation sur le fonctionnement de notre école, sur ses perspectives d'amélioration ne peut s'apprécier et s'analyser qu'en référence à ces données générales qui concernent autant les résultats observés chez les élèves que les ressources mobilisées pour atteindre les objectifs du système éducatif. Les données mobilisées dans ce chapitre, certes imparfaites, correspondent néanmoins à une collecte spécifique correspondant aux années les plus récentes. Deux points seront successivement abordés. Le premier apporte des informations sur le niveau des élèves et sur l'efficience qualitative de l'école primaire, le second présente une description et une analyse des coûts alloués à l'école primaire. Une double dimension comparative est utilisée pour rendre compte des évolutions spatiales et temporelles des principaux indicateurs sélectionnés.

## **I Le niveau moyen d'acquisition des élèves et l'efficience qualitative**

Les comparaisons internationales fournissent des éléments précieux pour apprécier l'efficacité de l'école primaire française. Notre pays a participé à plusieurs programmes d'évaluation des acquis des élèves depuis une quarantaine d'année ; la participation française n'est toutefois pas systématique puisque seulement 13 présences sur 27 enquêtes internationales sont répertoriées dans un récent rapport (Bottani, Vrignaud, 2005). Outre cette présence irrégulière française, une difficulté supplémentaire surgit pour établir des comparaisons temporelles puisque les niveaux scolaires, les pays et les disciplines évaluées varient d'une enquête à l'autre. Nous avons ciblé pour notre questionnement deux enquêtes initiées par l'I.E.A.<sup>4</sup>, à savoir l'enquête nommée « Reading Literacy » de 1990-1991 comprenant 32 pays et l'enquête PIRLS de 2001 comprenant 35 pays. Ces deux études ont l'intérêt de cibler la même dimension des acquis des élèves, à savoir la compréhension en lecture et le même niveau scolaire, soit la 4<sup>ème</sup> année de l'élémentaire (élèves âgés de 9 ans).

La mise en relation des deux enquêtes permet de mesurer l'évolution des performances relatives de notre pays sur une période de 10 ans. Il n'est pas possible, compte tenu d'une part de la participation irrégulière de la France aux enquêtes et, d'autre part des âges des élèves ciblés par ces études, d'obtenir la même comparaison pour les mathématiques. Le tableau

---

<sup>4</sup> International Association for the Evaluation of Educational Achievement.

suivant présente les résultats de ces deux enquêtes en mettant en évidence la position de la France à la fois en termes de score et de classement.

Il apparaît clairement que le classement relatif de la France est moins bon en 2001 (18<sup>ème</sup> sur 35 pays) qu'en 1990 (4<sup>ème</sup> sur 24 pays), même si l'écart en terme de score est pratiquement le même, soit environ 25 points au-dessus de la moyenne. En 1990, la France n'était devancée que par la Finlande (score de 569), les Etats-Unis (score de 547) et la Suède (score de 539). Si l'on ne prend en compte que les pays communs aux deux enquêtes, soit 16 pays, le constat est identique ; la France est en 3<sup>ème</sup> position en 1990 et en 11<sup>ème</sup> en 2001. Le score moyen des élèves français est en outre légèrement inférieur à la moyenne des scores de ces 16 pays (score moyen de 529).

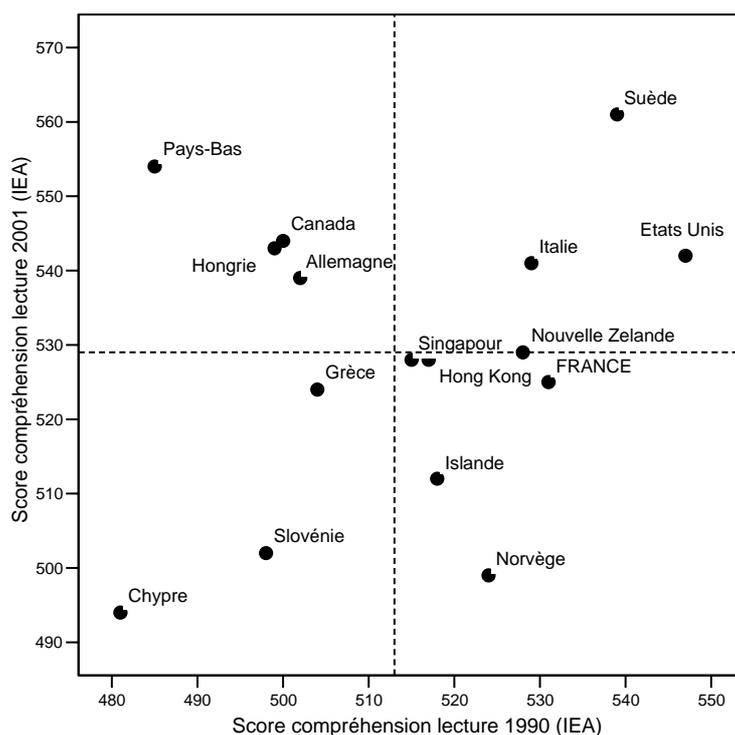
*Tableau 1 : Comparaison des performances de la France dans les enquêtes internationales en compréhension en lecture*

	Reading Literacy 1990	PIRLS 2001
Nombre de pays	24	35
Score mini sur l'ensemble des pays	394	327
Score maxi sur l'ensemble des pays	569	561
Score moyen sur l'ensemble des pays	507	500
Ecart-type sur l'ensemble des pays	33	60
<b>Score de la France</b>	<b>531</b>	<b>525</b>
<b>Position de la France</b>	<b>4<sup>ème</sup></b>	<b>18<sup>ème</sup></b>
<b>Ecart de la France à la moyenne</b>	<b>+ 24 points</b>	<b>+ 25 points</b>

Le graphique suivant permet de visualiser l'évolution temporelle pour les pays ayant participé aux deux enquêtes sur la compréhension en lecture<sup>5</sup>.

---

<sup>5</sup> Les lignes pointillées représentent les valeurs moyennes des deux indicateurs.



*Graphique 1 : Evolution du niveau moyen des élèves de 4<sup>ème</sup> année primaire en compréhension en lecture entre 1990 et 2001*

En termes de dispersion des résultats, la France se situe plutôt parmi les pays qui se caractérisent par une homogénéité des performances des élèves. Si l'on classe les pays en fonction de leur degré d'homogénéité sur la base de l'écart-type des scores, notre pays occupe en effet la 7<sup>ème</sup> place en 1990 (sur 24 pays) et la 10<sup>ème</sup> place en 2001 (sur 35 pays). L'enquête PIRLS permet également de disposer d'un indicateur qui identifie la proportion d'élèves les plus en difficulté dans le domaine de la lecture-compréhension<sup>6</sup>. La France affiche un pourcentage de 10% à cet indicateur alors que la moyenne est de 22,5% sur l'ensemble des pays mais seulement de 12% pour les pays les plus riches.

Ces résultats globaux sur l'école primaire peuvent être complétés par ceux issus des évaluations qui ciblent des élèves plus âgés et des disciplines différentes. Pour avoir une perspective longitudinale suffisamment longue, nous pouvons mobiliser les données de l'enquête TIMSS<sup>7</sup> de 1995 (élèves âgés de 13 ans) et l'enquête PISA de 2003 (élèves âgés de 15 ans). Les résultats diffèrent sensiblement selon la discipline considérée. En Sciences, la France se situait à la 28<sup>ème</sup> place (sur 41 pays) en 1995 et à la 13<sup>ème</sup> place en 2003 (sur 41 pays également). En mathématiques, le classement de la France correspondait à la 13<sup>ème</sup> place en

<sup>6</sup> Il s'agit des élèves qui ont obtenu un score inférieur à 435 points.

<sup>7</sup> Trends in International Mathematics and Science Study.

1995 et à 16<sup>ème</sup> place en 2003. Il est certes difficile de comparer cette évolution dans les classements (l'âge des élèves et les pays ne sont pas exactement les mêmes et la nature des compétences testées peut également différer) ; les écarts sont néanmoins tels que l'on peut supposer que les résultats des élèves français se seraient améliorés en sciences alors qu'ils seraient stables en mathématiques, voire qu'ils auraient connu une légère dégradation relative<sup>8</sup>.

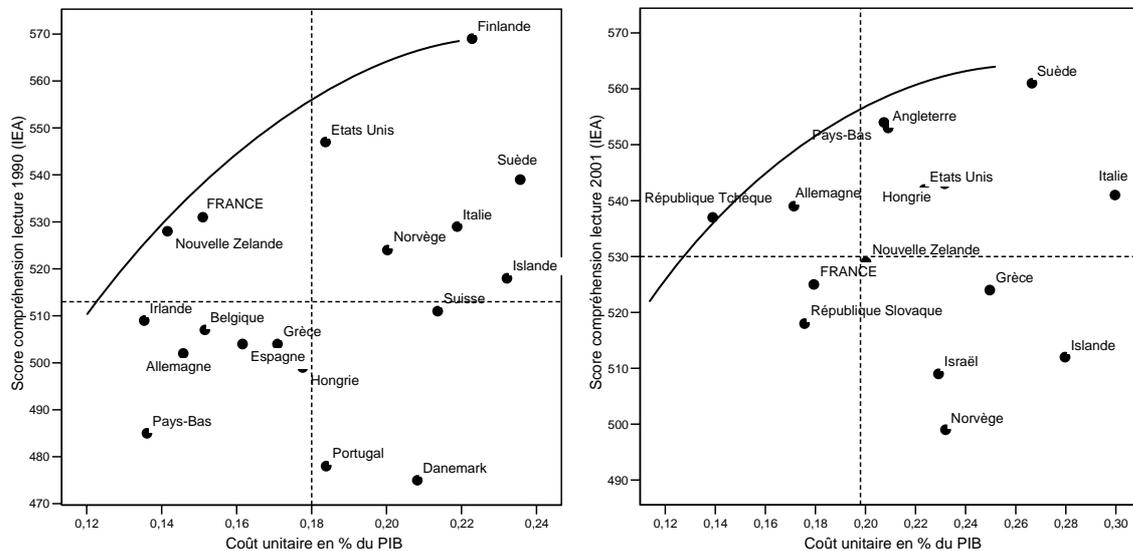
Les résultats de l'école primaire doivent également s'interpréter en fonction des ressources mises à dispositions des systèmes éducatifs concernant ce cycle d'enseignement ; la question de l'efficience qualitative est alors abordée. Il est d'usage d'utiliser les coûts unitaires pour rendre compte des ressources allouées à l'école ; ces coûts unitaires, pour être directement comparables d'un pays à l'autre demandent à être exprimés de manière relative et non pas absolue puisque les pays n'ont pas le même niveau de développement économique. Les coûts sont en effet d'autant plus élevés que les pays ont un PIB également élevé et, pour tenir compte de ce mécanisme, les coûts unitaires peuvent être exprimés en pourcentage de PIB par habitant. Une double perspective comparative (temporelle et spatiale) a été mobilisée pour traiter cette question de l'efficience qualitative de notre école primaire. Les graphiques suivants présentent pour les 16 pays communs aux enquêtes de 1990 (Reading Literacy) et 2001 (PIRLS) l'évolution des relations entre les coûts et les résultats des élèves. Les données collectées sur les coûts permettent de cibler des années équivalentes à celles qui correspondent aux enquêtes sur les acquis des élèves, soit 1990 et 2001. Les graphiques mettent successivement en relation les coûts unitaires de 1990 avec les scores moyens des pays de l'enquête Reading Literacy et les coûts unitaires de 2001 avec les scores moyens de l'enquête PIRLS. Les lignes en pointillés sur les graphiques représentent les valeurs moyennes pour les deux indicateurs (coûts et scores). Pour les pays pour lesquels nous disposons des résultats aux deux enquêtes, les coûts unitaires relatifs varient au début des années 1990 de 0,14 à 0,24 unités de PIB par habitant, la valeur moyenne étant de 0,18.

En 2001, les coûts varient de 0,14 à 0,30 avec une valeur moyenne de 0,22. La France affiche une valeur de 0,15 en 1994 et de 0,18 en 2001 ; aux deux périodes considérées, et compte tenu de son niveau de richesse, notre pays dépense donc plutôt moins pour l'école primaire que d'autres pays (Finlande, Suède, Italie, Islande notamment). Certains pays, comme l'Allemagne, la Nouvelle Zélande, l'Irlande ou les Pays-Bas présentent quant à eux des coûts unitaires plus réduits que ceux relevés pour la France. La lecture des graphiques est riche

---

<sup>8</sup> L'écart de la France à la moyenne de l'ensemble des pays est de +38 points pour les élèves de 8ème année testés dans TIMSS et de seulement +11 points en ce qui concerne les élèves de 15 ans testés dans PISA (la moyenne des deux tests étant fixée à 500).

d'enseignements quand on se focalise sur l'efficacité des systèmes éducatifs et à son évolution dans le temps.



Graphiques 2 et 3 : Evolution de l'efficacité qualitative de l'école primaire de 1990 à 2001

Certains pays obtiennent de bons résultats compte tenu des ressources mobilisées, ils s'approchent des frontières d'efficacité matérialisées sur les graphiques par les courbes en trait plein. Les pays qui peuvent ici être considérés comme efficaces sont l'Irlande, la Nouvelle-Zélande, les Etats-Unis et la France pour l'année 1990 et la République Tchèque, les Pays-Bas, l'Angleterre et la Suède pour l'année 2001. A l'inverse, d'autres pays obtiennent des résultats faibles au regard des coûts engagés pour leur école primaire et ils s'éloignent fortement de la frontière d'efficacité ; c'est le cas du Portugal, de l'Islande, de la Norvège, du Danemark ou bien encore d'Israël. Il est également instructif de comparer la situation d'un même système éducatif à dix années d'intervalle. On remarque ainsi que certains pays comme l'Allemagne réalisent des gains d'efficacité significatifs sur cette période. Ce n'est pas le cas de la France puisqu'en dix années son classement n'a pas évolué favorablement alors que les coûts unitaires ont augmenté.

Même si les données mobilisées ne sont pas parfaites<sup>9</sup>, les comparaisons internationales laissent penser que l'efficacité de l'école primaire française ne s'est pas améliorée au cours de cette dernière décennie. De façon complémentaire, on peut sérieusement s'interroger sur l'efficacité actuelle de l'école primaire, celle-ci s'étant visiblement dégradée au fil du temps.

<sup>9</sup> Les pays participants aux enquêtes ne sont pas toujours les mêmes, les disciplines testées et les niveaux scolaires ciblés sont également variables, si bien qu'il est très difficile d'avoir, pour un même pays des mesures comparables au fil des années.

Les analyses comparatives fournissent une image relative du niveau moyen des acquisitions des élèves en le situant par rapport au niveau des élèves des autres pays, mais il n'est pas possible de savoir précisément si les acquisitions ont réellement diminué. Pour répondre à cette question il est nécessaire de disposer d'une mesure des performances établie sur une même échelle de mesure. Depuis plus de quinze ans, la D.E.P.<sup>10</sup> a mis en place des évaluations nationales qui fournissent chaque année une mesure du niveau de connaissances des élèves à l'école primaire. Ces évaluations institutionnelles sont organisées depuis 1989 en début d'année scolaire dans les classes de CE2 et de 6<sup>ème</sup>. L'objectif de ces évaluations est de « *permettre d'apprécier les réussites et les difficultés éventuelles de chaque élève considéré individuellement, à un moment précis de la scolarité.* »<sup>11</sup>. Il pourrait être alors envisageable d'utiliser les résultats de ces évaluations diagnostiques pour étudier l'évolution temporelle des acquisitions en français et en mathématiques. Ceci serait toutefois inapproprié et pourrait donner lieu à des conclusions trompeuses. En effet, d'une année à l'autre, le contenu des tests varie, ces différences de contenu s'accompagnant d'une variation de la difficulté des items. Il n'est donc pas surprenant que les pourcentages moyens de réussite aux évaluations nationales varient parfois fortement d'une année sur l'autre, comme le montre le graphique suivant avec l'évolution des résultats en français en CE2 entre 1996 et 2005<sup>12</sup>. Des écarts sont présents pour les scores globaux moyens mais aussi pour les différents champs évalués.

La moyenne des scores globaux oscille entre 60,5% (en 2001) et 73,5% (en 2004) avec un chiffre moyen de 67,3%. Ces chiffres dépendent donc fortement de la nature des évaluations et les variations constatées d'une année sur l'autre ne renvoient pas forcément à une baisse ou une hausse du niveau des élèves, comme c'est par exemple le cas pour les scores en production d'écrits entre 1996 et 1998. La question de l'évolution des performances dans une période récente peut néanmoins être traitée de façon partielle par une étude réalisée par la D.P.D.<sup>13</sup>

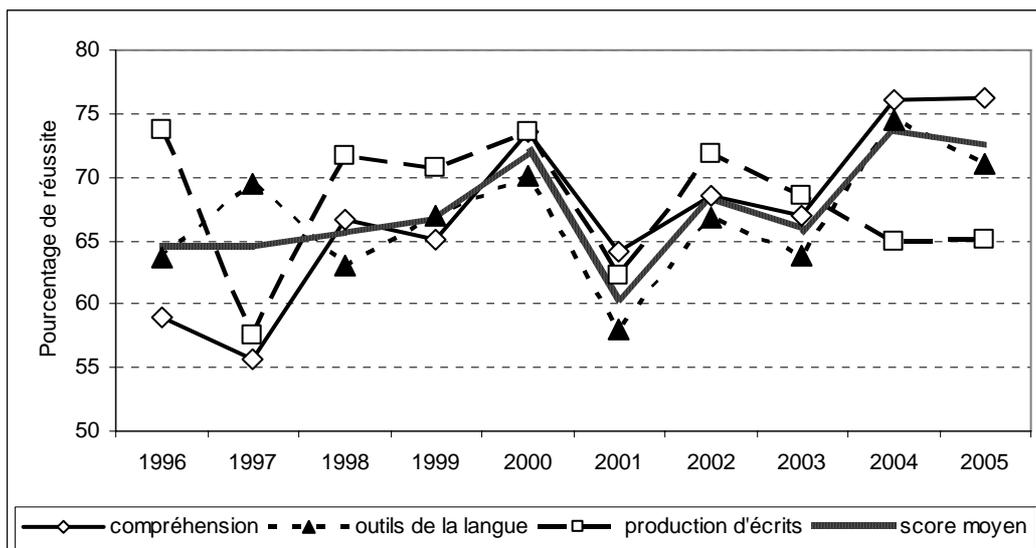
---

<sup>10</sup> DEP : Direction de l'Evaluation et de la Prospective.

<sup>11</sup> Note de service 2002-105 du 30.04.2002, B.O. N°19 du 9 mai 2002, 096 du 24-6-2005.

<sup>12</sup> Pour l'année 2005, le domaine habituellement nommé « outils de la langue » est scindé en deux champs « reconnaissance des mots » et « écriture et orthographe ». Nous avons reporté sur le graphique les scores correspondant à l'ensemble des items des deux champs.

<sup>13</sup> Direction de la Programmation et du Développement (ancienne dénomination de la D.E.P.).



*Graphique 4 : Evolution temporelle des résultats des évaluations nationales de français au CE2 entre 1996 et 2005*

Pour évaluer l'évolution des performances des élèves à la fin de l'école primaire, le Ministère de l'Education nationale a réalisé une étude comparative à dix années d'intervalles (1987-1997) en ciblant le domaine de la compréhension en lecture (MEN-DPD, 1998). Un échantillon représentatif d'élèves fréquentant la classe de CM2 en 1997 a été soumis à un test identique à celui qui avait été administré dix ans plus tôt à un autre échantillon d'élèves de même niveau scolaire, le CM2. De façon globale, l'étude met en évidence que la capacité des élèves à comprendre un texte est, en 1997, identique à celle de leurs camarades testés dix ans auparavant. Les scores moyens sont de 65,6% en 1987 et de 66,1% en 1997. Qui plus est, la dispersion des résultats est aussi comparable, puisque l'écart type des résultats est de 14,8 pour 1987 et de 15,3 pour 1997.

Au-delà de ce constat général, l'étude a également permis de mettre en évidence quelques résultats instructifs. Aux deux périodes considérées, environ 70% des questions relevant de la compréhension immédiate (simple prélevé d'informations) sont réussies. A l'inverse, les questions mettant en jeu la compétence « construction d'informations et de significations sur un texte » ne sont réussies que dans 50% des cas, tant en 1987 qu'en 1997. La comparaison temporelle a aussi porté sur l'effet des caractéristiques des élèves sur leurs performances. Le sexe n'influence que très faiblement les résultats au test puisque les performances des garçons et des filles sont similaires en 1987 et 1997. L'âge est un facteur qui influence fortement les performances à l'épreuve : en 1987 et 1997, ce sont les élèves en avance qui obtiennent les meilleures performances et les élèves en retard scolaire les plus faibles ; les élèves « à l'heure » affichant des scores intermédiaires.

Comme à tous les niveaux scolaires, l'origine sociale pèse également sur les résultats des élèves et une hiérarchie bien connue est mise en évidence. Aux deux moments de l'évaluation, ce sont toujours les enfants de cadres ou de personnes exerçant une profession intellectuelle supérieure qui obtiennent les résultats les plus élevés. Viennent ensuite respectivement, les enfants dont le chef de famille exerce une profession intermédiaire, les enfants d'agriculteurs, d'employés et d'artisans ou commerçants, les enfants dont le chef de famille est ouvrier ou sans profession et, enfin, les enfants de retraités. L'orientation en fin de CM2 est un facteur qui permet de différencier nettement les élèves. En 1987 et 1997, les élèves promus en sixième obtiennent des performances nettement supérieures à celles des élèves doublants : 66,5 % contre 54,2 % en 1987 et 66,9 % contre 50,0 % en 1997. Ce qui est particulièrement intéressant dans ces derniers chiffres c'est que l'on note une nette détérioration des résultats pour les élèves redoublants sur une période de dix années. L'étude montre aussi que la baisse des résultats affecte particulièrement les compétences de haut niveau (compréhension logique notamment) alors qu'elle est moins prononcée pour les compétences de plus bas niveau (compréhension immédiate).

Pour conclure sur ce descriptif général, et bien que les données disponibles ne soient pas parfaites, l'examen de la qualité des acquisitions des élèves français dans une perspective comparative ne permet pas de conclure à une amélioration du niveau global. En ce qui concerne l'école primaire, la France est plutôt moins bien classée dans les enquêtes internationales qu'auparavant. Pour les élèves plus âgés (15 ans), l'enquête PISA nous apprend que notre pays se situe dans la moyenne internationale, mais la comparaison entre 2000 et 2003 nous indique également que la proportion de jeunes les moins performants (situés dans le groupe le plus faible dans l'échelle des scores de PISA), augmente en passant de 4,2% à 6,3%. L'ensemble de ces constats sur la qualité et l'efficacité de notre école primaire, nous amène naturellement à étudier de manière plus spécifique les coûts alloués au fonctionnement de l'école.

## II Coûts et financement de l'enseignement primaire en France

On examinera dans un premier temps la situation de la France de façon globale sur le plan de l'efficacité quantitative. Dans un second temps, une comparaison des coûts du primaire à ceux des autres niveaux d'enseignement sera présentée et commentée. Dans un troisième point une analyse détaillée des coûts unitaires du primaire sera réalisée en identifiant les facteurs explicatifs des évolutions constatées au cours des trente dernières années. Enfin, les coûts de l'enseignement primaire seront interprétés dans une perspective de comparaisons internationales.

### II.1 Une perspective globale

En 2003, la France affecte 11,0 % de ses dépenses publiques totales et 5,9 % de son Produit Intérieur Brut au financement public de l'éducation (tableau 2).

Tableau 2 : Comparaisons internationales des dépenses publiques d'éducation (2003)

Pays	Dépenses publiques d'éducation			% Financement final (primaire et secondaire)			
	% des dépenses publiques	% du PIB	E.V.S. <sup>14</sup> (Années)	Central	Régional	Local	Total
<b>France</b>	<b>11,0</b>	<b>5,9</b>	<b>16,8</b>	<b>74</b>	<b>14</b>	<b>12</b>	<b>100</b>
Allemagne	9,7	4,7	17,4	10	69	21	100
Belgique	11,8	6,1	19,6	15	80	5	100
Danemark	15,1	8,3	19,0	29	12	59	100
Etats-Unis	15,2	5,7	16,9	1	1	99	100
Italie	9,9	4,9	17,0	81	6	14	100
Japon	10,7	3,7	17,2	1	81	18	100
Norvège	15,7	7,6	18,4	10	29	62	100
Royaume-Uni	11,9	5,4	20,7	25	-	75	100
Suède	12,8	7,5	20,3	-	-	-	-
Suisse	13,0	6,0	16,8	-	57	43	100
Moyenne OCDE	13,3	5,5	17,4	33	28	39	100
Moyenne U.E.(19 pays)	11,2	5,5	17,7				

Source : « Regards sur l'éducation, 2006 », OCDE 2006: Tableaux B 4.1, C 1.1, B 4.3a (web)

Pour identifier l'intensité de l'effort public pour le secteur éducatif, il est préférable de prendre en compte le volume des ressources publiques d'éducation en proportion du PIB plutôt que la part des dépenses publiques pour le secteur au sein des dépenses publiques globales. En effet, l'importance des dépenses varie d'un pays à l'autre (le poids des dépenses publiques étant, en France, parmi les plus forts des pays de l'OCDE). Le niveau des dépenses

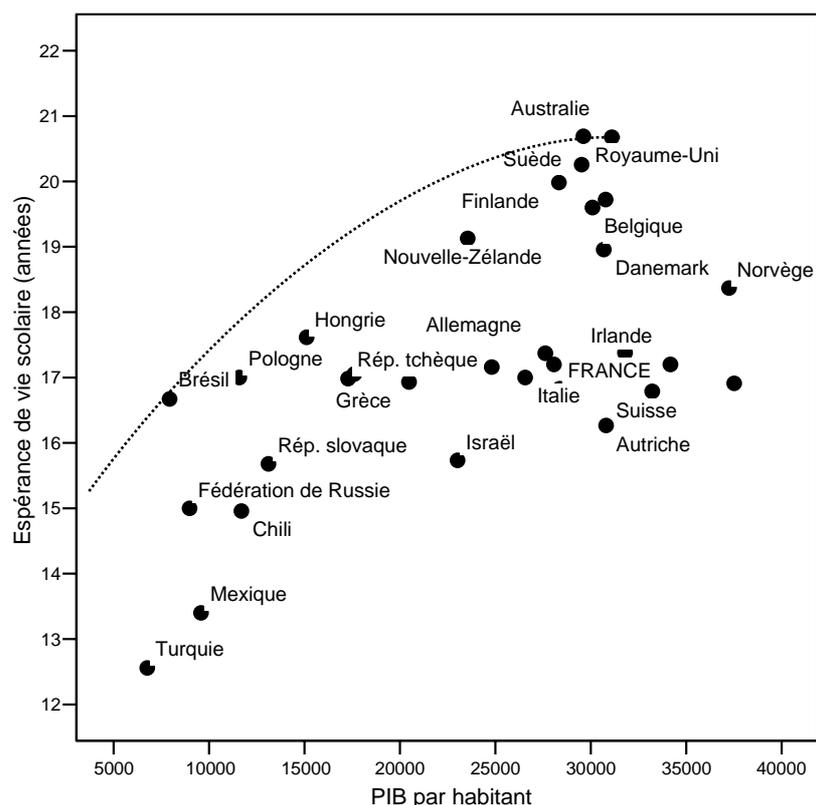
<sup>14</sup> E.V.S. : Il s'agit d'une estimation du nombre d'années de scolarisation et de formation continue (à l'exclusion des programmes destinés aux enfants de moins de 5 ans). Le tableau mentionne les valeurs correspondant à l'année 2004.

publiques d'éducation est en France un peu plus élevé (5,9 % du PIB) que la moyenne des pays de l'OCDE ou de l'Union Européenne (5,5 %). Cette moyenne est toutefois assortie d'une assez forte dispersion : des pays comme l'Allemagne (4,7 %), l'Italie (4,9 %) ou plus encore le Japon (3,7 %) ont des niveaux de dépenses publiques pour l'éducation sensiblement inférieurs au chiffre observé pour la France, alors que les pays scandinaves présentent des chiffres plus élevés (entre 7,5% et 8,3 %).

En contrepartie des dépenses engagées, les systèmes éducatifs produisent des résultats qui sont de différentes natures. En premier lieu, quand les jeunes sont encore scolarisés, les résultats peuvent se mesurer, soit de façon quantitative (avec des taux ou des durées moyennes de scolarisation), soit de façon qualitative par le niveau des apprentissages comme précédemment. En second lieu, on peut aussi apprécier les résultats des investissements financiers quand les jeunes ont quitté l'école, soit en mesurant la contribution de l'éducation reçue à la croissance économique (ou à la cohésion sociale), soit en examinant dans quelle mesure les jeunes les plus éduqués ont des comportements sociaux différents des autres.

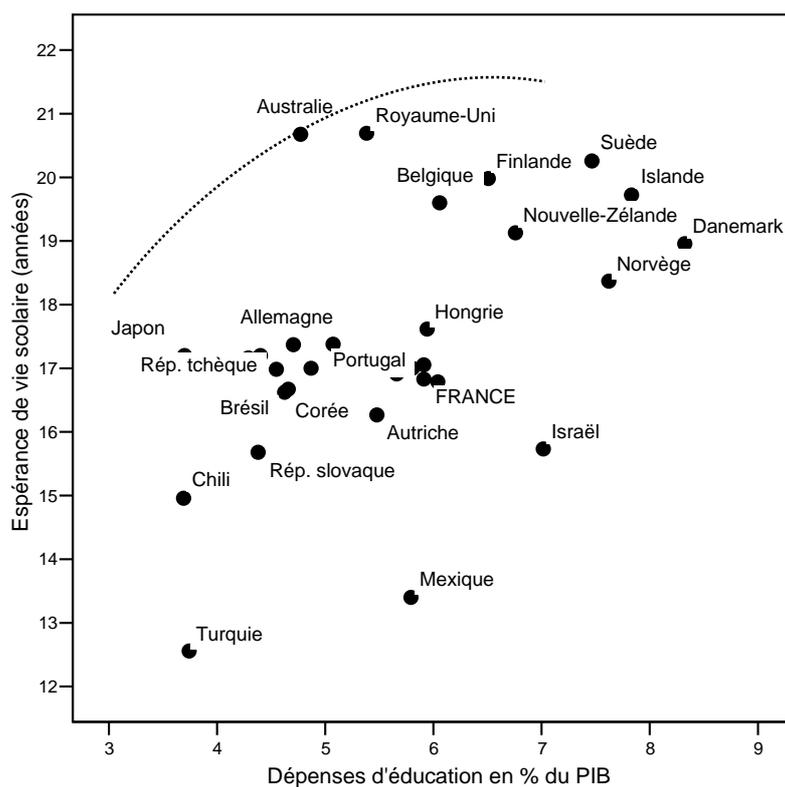
Nous nous limitons volontairement ici au fonctionnement interne du système éducatif en cibler le développement quantitatif global des systèmes sur la base de l'Espérance de Vie Scolaire (E.V.S.). D'un pays de l'OCDE à l'autre, cet indicateur varie de 12,6 années (Turquie) à 20,7 années (Australie et Royaume-Uni) ; la valeur pour la France est de 16,8 années (la valeur moyenne des pays de l'UE est de 17,7 années). Le graphique ci-après, illustre pour l'année 2003 la relation entre l'E.V.S. et le niveau de développement économique des pays (exprimé en PIB par habitant en dollars des Etats-Unis et en parité de pouvoir d'achat).

Les données de ce graphique montrent qu'il y a bien une liaison croissante entre l'espérance de vie scolaire et le développement économique des pays, mais qu'il existe aussi une dispersion assez forte de l'E.V.S. lorsqu'on se situe au même niveau de PIB par habitant. La courbe en pointillés sur le graphique représente une pseudo frontière d'efficacité, c'est-à-dire le niveau maximum de l'E.V.S. pour chacun des niveaux de PIB par habitant. Parmi les pays proches de cette frontière, on peut citer le Brésil et la Pologne pour les pays à revenu intermédiaire et l'Australie, la Nouvelle-Zélande et la Suède pour les pays les plus développés. La France n'est pas très bien classée non plus sur le plan de l'efficacité quantitative, elle affiche en effet une E.V.S. de 16,8 années alors que sa valeur pourrait atteindre environ 20 années. Le constat en matière d'efficacité quantitative correspond au précédent sur l'efficacité qualitative, à savoir que la France occupe une place peu enviable sur le plan international.



Graphique 5 : Espérance de Vie Scolaire selon le PIB par habitant (USD en PPA), 2003

Il est maintenant intéressant d'examiner dans quelle mesure ce développement quantitatif de l'éducation dans les différents pays dépend du niveau de leur financement public pour le secteur (graphique 6). Selon ces données, le volume des dépenses publiques pour l'éducation exerce certes un rôle pour prédire le développement quantitatif des systèmes, mais la relation globale reste relativement modeste ( $R^2 = 0,23$ ). On observe en effet que des pays ayant un même niveau de dépenses publiques pour l'éducation ont des développements quantitatifs variés. Ainsi, si on prend le cas d'Israël, ce pays est caractérisé par une E.V.S. de 15,7 années et une dépense publique d'éducation qui correspond à 7 % du PIB. La situation d'Israël peut être comparée à celles de la Finlande et de la Suède, pays qui présentent un niveau de dépense publique équivalent mais qui offrent une E.V.S. supérieure à 20 années d'études. A l'inverse, la Slovaquie affiche une E.V.S. équivalente à celle d'Israël (15,7 années) mais ses dépenses publiques ne correspondent qu'à 4,4 % du PIB du pays. Ceci témoigne d'une assez grande variété du niveau d'efficacité des pays dans l'usage des ressources publiques consacrées à leur secteur scolaire. La situation de la France est de nouveau moyenne (voire médiocre) dans la mesure où l'efficacité maximum suggère une E.V.S. supérieure à 20 années compte tenu de son niveau de mobilisation des ressources. A titre de comparaison, l'Allemagne a une E.V.S. comparable à celle de la France mais réalise ce résultat en mobilisant 20 % de crédits publics pour l'éducation en moins.



Graphique 6 : *Espérance de vie scolaire selon la part des dépenses publiques d'éducation dans le PIB, 2003*

Concernant l'organisation globale du système, il est également intéressant d'observer la structure de son financement en distinguant ce qui vient du niveau central ou national de ce qui trouve son origine au niveau déconcentré (régional ou local). A cet égard, la situation des différents pays est très variée (tableau 2). Alors que dans un certain nombre d'entre eux (Allemagne, Etats-Unis, Finlande, Japon, Norvège) l'éducation primaire et secondaire est financée de façon très majoritaire aux échelons administratifs décentralisés, ce n'est pas le cas de l'Italie et de l'Irlande où plus de 80 % du financement final est assuré par le niveau central; ce n'est pas non plus le cas de la France où les trois quarts du financement du primaire et secondaire proviennent du niveau central. Il existe bien en France des transferts financiers entre le niveau central et les échelons décentralisés (régions ou départements) mais ceux-ci apparaissent pour des montants relativement modestes.

## II.2 La situation de l'enseignement primaire dans le secteur scolaire

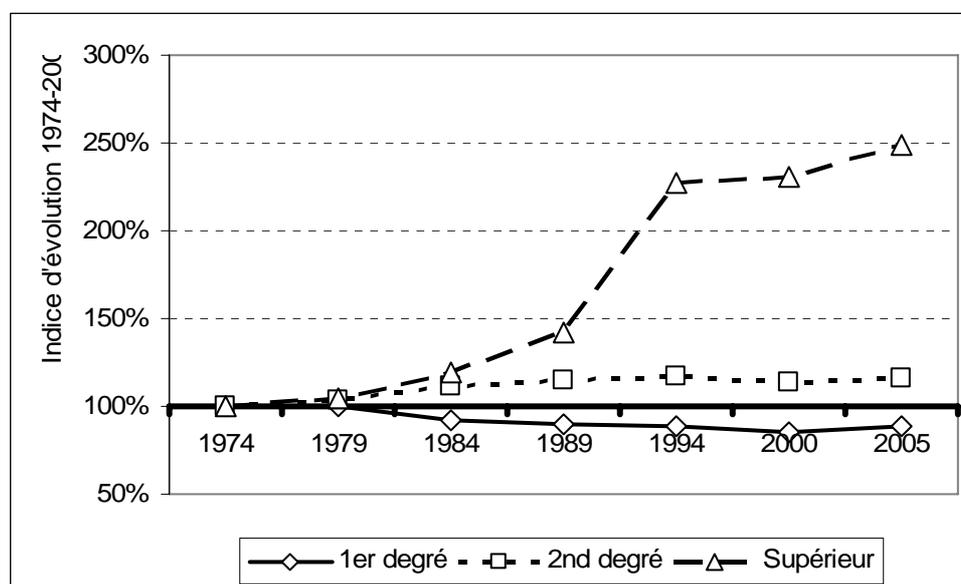
Avant d'examiner plus en détail les aspects financiers, il importe de décrire l'évolution des effectifs aux différents niveaux d'enseignement. Le Tableau 3 propose cette évolution depuis 1974.

Tableau 3 : Effectifs d'élèves (en milliers) par niveaux d'enseignement entre 1974 et 2005

	1974	1984	1994	2000	2005	Ratio 2005/1974
Premier degré	7 411,4	6 804,5	6 561,7	6 301,3	6 585,5	0,89
Second degré	4 798,7	5 349	5 617,3	5 443,8	5 540,3	1,15
Supérieur	913,5	1 086,3	2 074,6	2 105,6	2 269,8	2,48

Source : « Repères et références statistiques sur les enseignements et la formation » : Editions 1975, 1985, 1995, 2001, 2006.

Compte tenu des aspects démographiques et des dynamiques propres des différents niveaux d'enseignement, les évolutions entre 1974 et 2005 sont relativement contrastées. Ainsi, le niveau primaire voit-il ces effectifs se réduire de plus d'un million entre 1974 et 2000 (baisse de 15 %) et on assiste ensuite à une légère augmentation des effectifs (+ 284 000 entre 2000 et 2005). En revanche, pour le cycle secondaire, on remarque une assez forte augmentation des effectifs entre 1974 et 1994 (+ 820 000 élèves représentant une augmentation de 17 %), puis une stabilisation est observée. Dans l'enseignement supérieur, la dynamique de croissance est beaucoup plus forte, les effectifs doublant entre 1984 et 1994 (+ 1 million d'étudiants) pour plus ou moins se stabiliser par la suite (+ 200 000 entre 1994 et 2005). Le Graphique 7 ci-après illustre ces évolutions.



Graphique 7 : Evolution (base 100 en 1974) des effectifs par ordre d'enseignement (1974-2005)

D'une certaine manière, cette dynamique des effectifs a des conséquences dans la dimension financière. Le Tableau 4, ci-après, décrit l'évolution entre 1974 et 2005 de la Dépense Intérieure d'Education (D.I.E.<sup>15</sup>) selon les niveaux d'enseignement. Outre les dépenses

<sup>15</sup> DIE : effort consenti par la collectivité nationale pour le fonctionnement et le développement du système

clairement identifiées par niveau d'enseignement, le tableau reprend aussi les dépenses pour la formation continue (non ventilées par niveaux) ainsi que certaines autres dépenses globales pour le système (également non ventilées par niveaux). Entre 1974 et 2005, la D.I.E. en monnaie courante passe de 84,4 milliards de francs (correspondant à 12,9 milliards d'euros) à 117,9 milliards d'euros, soit une multiplication par un facteur 9,1 sur la période. La valeur de la monnaie évidemment évoluée au fil des ans, il est préférable de raisonner en monnaie constante, ce que proposent les données du tableau suivant.

*Tableau 4: Structure de la DIE par niveaux d'enseignement (millions d'euros prix 2005)*

	1974		1984		1994		2000		2005		Ratio 2005/1974
	M €	%	M €	%	M €	%	M €	%	M €	%	
<b>Premier degré</b>	<b>16790</b>	<b>30,8</b>	<b>20339</b>	<b>28,0</b>	<b>27430</b>	<b>27,0</b>	<b>31319</b>	<b>27,0</b>	<b>32 085</b>	<b>27,2</b>	<b>1,91</b>
Second degré	23501	43,1	33219	45,4	46313	45,1	52744	45,4	52 515	44,5	2,23
Supérieur	7932	14,5	11209	15,3	17238	16,8	19518	16,8	20 593	17,5	2,60
Formation continue	5672	10,4	7520	10,3	10481	10,2	11129	9,6	10 786	9,1	1,90
Non ventilé	653	1,2	870	1,2	1223	1,2	1353	1,2	1 949	1,7	2,98
<b>Total courant</b>	<b>54548</b>	<b>100,0</b>	<b>73157</b>	<b>100,0</b>	<b>102684</b>	<b>100,0</b>	<b>116062</b>	<b>100,0</b>	<b>117 928</b>	<b>100,0</b>	<b>2,16</b>

Source : Calculs basés sur les comptes de l'éducation : 1974 - 2005

En monnaie constante (prix de 2005), le volume des ressources pour le secteur passe de 54,5 milliards d'euros en 1974 à 117,9 milliards d'euros en 2005, soit une augmentation par un facteur de 2,16. Dans ce contexte, la part du primaire diminue de 30,8 % à 27,2 % mais cette diminution concerne principalement les dix premières années de la période. Le volume des ressources pour le premier degré progresse donc à un rythme un peu inférieur à celui de l'ensemble de la DIE (facteur d'augmentation de 1,91 contre 2,16 pour l'ensemble). Les ressources pour le niveau secondaire progressent un peu plus vite que pour l'ensemble mais c'est surtout l'enseignement supérieur qui voit ces ressources progresser de façon sensible (la part de l'enseignement supérieur passe de 14,5 % en 1974 à 17,5 % en 2005, le coefficient de progression étant de 2,60 sur l'ensemble de la période). Il importe toutefois de mettre ces évolutions en perspective avec celles des effectifs scolarisés (tableau 5).

*Tableau 5 : Evolution comparée des ressources et des effectifs par niveau d'études (1974-2005)*

	Rapport entre 2005 et 1974		
	Effectifs (a)	DIE (b)	Rapport (b/a)
Primaire	0,89	1,91	2,15
Secondaire	1,15	2,23	1,94
Supérieur	2,48	2,60	1,05

Au regard de ces chiffres, la situation favorable de l'enseignement supérieur observée précédemment est maintenant contredite par ces informations. A ce niveau d'études, la croissance des financements (en valeur monétaire constante) est du même ordre de grandeur que celle des effectifs, alors que ce n'est pas le cas pour le secondaire et plus encore pour le primaire. Pour ce dernier niveau les effectifs globaux baissent de 11 % sur la période, tandis que les ressources en termes réels augmentent de 94 %. Alors que la situation budgétaire de l'enseignement primaire paraissait défavorable sur la base des données agrégées, la mise en regard avec les effectifs suggère que ce niveau d'études a en fait été clairement favorisé. Une façon plus directe de mettre en regard les dépenses et les effectifs consiste à mesurer des coûts unitaires de scolarisation (tableau 6). Ceux-ci sont habituellement calculés sur la base des seules dépenses courantes qui représentent, en 2005, 91,5 % de la dépense totale dans le premier degré, 89,0 % dans le second et 89,9 % dans le supérieur.

*Tableau 6 : Coûts unitaires de scolarisation (en euros de 2005) par niveau d'études après ventilation des dépenses non affectées (1974 – 2005)*

Niveaux d'études	1974	1984	1994	2000	2005	Ratio 2005/1974
1er degré	2 431	3 440	4 252	5 003	5 463	2,25
2nd degré	4 368	6 045	7 676	8 876	10 620	2,43
Supérieur	9 068	10 265	8 676	9 708	10 171	1,12

*Source : Calculs basés sur les comptes de l'éducation : 1974 - 2005*

On retrouve ici la dégradation relative de la situation de l'enseignement supérieur : alors que son coût unitaire de fonctionnement était plus le double de celui du second degré en 1974, il se trouve être inférieur à ce dernier pour l'année 2005. On pourra certes arguer qu'il est normal que le coût marginal de développement soit plus faible que le coût moyen dans l'enseignement supérieur eu égard à l'existence d'économies d'échelle à ce niveau d'études; mais ce point n'est évidemment pas de nature à expliquer l'ampleur des évolutions respectives du coût unitaire dans le secondaire et le supérieur. Dans l'enseignement primaire, le coût unitaire de fonctionnement passe, en valeur monétaire de 2005, de 2 431 euros en 1974 à 5 463 euros en 2005 (une augmentation par un facteur de 2,25).

### *II.3 Les facteurs explicatifs de l'évolution des coûts de l'école primaire*

Jusqu'à présent, nous avons considéré le premier degré et son financement de façon globale. Nous pouvons maintenant distinguer le préélémentaire (école maternelle) et l'élémentaire (du CP au CM2) d'une part, le financement de l'Education Nationale et celui des collectivités territoriales, d'autre part. Le Tableau 7 reprend les dépenses des administrations publiques (Education nationale et collectivités territoriales, soit les communes, les départements et les régions). Les dépenses des ménages ne sont pas prises en compte mais il est estimé que celles-ci représentaient environ 1,6 milliards d'euros en 2005. Par ailleurs, les dépenses des différents niveaux de collectivités territoriales sont agrégées sachant que, dans le premier degré, la très grande majorité de celles-ci sont le fait des communes (les contributions des départements et des régions sont très réduites).

Une première observation est que le financement du premier degré est partagé entre l'Etat et les collectivités territoriales et que ces dernières assument un rôle croissant à cet égard. Ainsi, en 1984 (période de la première loi de décentralisation), les collectivités territoriales assuraient 34 % de la dépense totale, alors que ce chiffre a augmenté progressivement pour atteindre 43,5 % en 2005 (l'Etat assure donc 56,5 % du total). Si la quasi-totalité des dépenses d'investissement (dont les constructions scolaires et l'équipement des écoles) est assurée par les collectivités territoriales (les communes), celles-ci assurent aussi une part importante et croissante des dépenses courantes. Les communes prenaient en charge 29,8 % des dépenses courantes en 1984 et cette proportion est estimée à 37,8 % en 2005.

Le tableau permet de retrouver la forte augmentation, dans le temps et en valeur monétaire constante, du niveau des ressources publiques par élève dans le primaire. Il permet aussi de séparer les dépenses agrégées et les coûts unitaires entre la maternelle et l'élémentaire. Il ressort que depuis le milieu des années 90, la maternelle et l'élémentaire ont des coûts unitaires très proches. Alors que dans les années 80, l'enseignement élémentaire était environ 20 % plus coûteux que le préélémentaire.

Tableau 7: Dépenses des administrations publiques du premier degré (courantes et capital)  
par type de financeur (euros de 2005, 1974-2005)

	1974	1984			1994			2000			2005		
	Total	Mat.	Elém.	Total	Mat.	Elém.	Total	Mat.	Elém.	Total	Mat.	Elém.	Total
Dépenses courantes	11 902	4 848	10 051	15 804	6 993	10 904	18 977	8 349	12 523	22 011	10 623	15 371	27 113
Education Nationale	8 289	3 139	7 203	11 094	4 348	7 248	12 496	5 122	8 109	14 152	6 371	9 583	16 870
Collectivités territoriales	3 613	1 709	2 848	4 711	2 645	3 657	6 481	3 228	4 414	7 859	4 253	5 788	10 243
Dépenses en capital													
Collectivités territoriales	1 819	186	845	1 036	221	1 143	1 368	284	1 495	1 786	434	2 283	2 727
<b>Total</b>	<b>13 721</b>	<b>5 034</b>	<b>10 896</b>	<b>16 841</b>	<b>7 213</b>	<b>12 047</b>	<b>20 345</b>	<b>8 634</b>	<b>14 018</b>	<b>23 797</b>	<b>11 057</b>	<b>17 654</b>	<b>29 840</b>

% Education Nationale dans total	60,4	62,3	66,1	65,9	60,3	60,2	61,4	59,3	57,8	59,5	57,6	54,3	56,5
% Collectivités territoriales dans total	39,6	37,7	33,9	34,1	39,7	39,8	38,6	40,7	42,2	40,5	42,4	45,7	43,5
% Collectivités territoriales dans les dépenses courantes	30,4	35,3	28,3	29,8	37,8	33,5	34,2	38,7	35,2	35,7	40,0	37,7	37,8

<b>Effectifs</b>	<b>7 411,4</b>	<b>2 461,4</b>	<b>4 233,4</b>	<b>6 804,5</b>	<b>2 548,5</b>	<b>3 943,2</b>	<b>6 561,7</b>	<b>2 416,8</b>	<b>3 828,3</b>	<b>6 301,3</b>	<b>2 609,5</b>	<b>3 924,6</b>	<b>6 585,5</b>
<b>C.U. de fonctionnement</b>	<b>1 606</b>	<b>1 970</b>	<b>2 374</b>	<b>2 323</b>	<b>2 744</b>	<b>2 765</b>	<b>2 892</b>	<b>3 455</b>	<b>3 271</b>	<b>3 493</b>	<b>4 071</b>	<b>3 917</b>	<b>4 117</b>

Source : Calculs basés sur les comptes de l'éducation : 1974 - 2005

Il importe à présent d'essayer d'identifier les éléments qui expliquent la forte augmentation des coûts unitaires dans le primaire, ceci est effectivement important pour rendre compte de l'inefficience constatée précédemment. Rappelons qu'en valeur monétaire constante (euros de 2005), le coût unitaire de fonctionnement passe de 1 606 euros de 2005 en 1974 à 4 117 euros pour l'année 2005, soit une multiplication par facteur de 2,6. Deux raisons principales peuvent expliquer les évolutions constatées. En premier lieu, il est possible que le contexte d'enseignement se soit globalement amélioré, par exemple en réduisant le rapport élèves-maître, ou bien en utilisant davantage de personnels d'appui (pédagogique ou technique), ou bien encore en augmentant la disponibilité de moyens d'enseignement pour les élèves et les enseignants. Ce serait donc davantage de facteurs de production scolaire qui seraient mobilisés. En second lieu, le coût de ces facteurs de production scolaire a pu augmenter, soit pour des raisons spécifiques à l'école (on emploie des enseignants plus diplômés et donc plus rémunérés), soit pour des raisons plus générales tenant notamment à l'augmentation du pouvoir d'achat des salariés dans le pays, les enseignants ne faisant évidemment pas exception à ce mouvement d'ensemble. Le tableau 8 permet d'obtenir une estimation de la décomposition des principaux effets ayant influé sur l'évolution du coût unitaire de fonctionnement de l'enseignement du premier degré.

*Tableau 8 : Synthèse du poids des facteurs ayant influencé l'évolution des coûts unitaires dans le premier degré, 1974 – 2005*

	1974	1984	1994	2000	2005	Ratio 2005/1974
<b>Coût unitaire de fonctionnement (euros 2005)</b>	<b>1 606</b>	<b>2 323</b>	<b>2 892</b>	<b>3 493</b>	<b>4 117</b>	<b>2,56</b>
Dépenses salariales unitaires (euros 2005)	1 193	1 838	2 366	2 910	3 219	2,70
Dépenses salariales enseignement	1 126	1 730	2 192	2 711	2 956	2,63
Dépenses salariales annexes	67	108	174	199	263	3,90
Autres dépenses unitaires (euros 2005)	413	485	526	583	898	2,17
Rapport élèves-maître du 1 <sup>er</sup> degré public	25,5	18,7	18,7	18,0	16,8	1,52
Proportion des professeurs des écoles	0	0	11	45	85	-
Coefficient d'impact salarial de la revalorisation	100	100	102	108	115	1,15
Coefficient de gain de pouvoir d'achat	100	113	140	157	150	1,50
PIB/habitant (euros 2005)	16 061	18 528	22 743	25 763	27 408	1,71

L'évolution constatée du coût unitaire de fonctionnement de 1 606 euros par élève en 1974 à 4117 euros en 2005 (en euros constants) demande à ne pas être interprété de façon globale car tous les indicateurs mentionnés, d'une part ne connaissent pas nécessairement la même dynamique sur la période et, d'autre part, ne sont pas influencés par les mêmes facteurs. En ce qui concerne tout d'abord les dépenses non salariales<sup>16</sup>, elles évoluent de 413 euros par élève en 1974 à 898 euros en 2005, soit une multiplication par un facteur de 2,17 sur cette période. Pour cet agrégat, l'interprétation est immédiate : les élèves ont bénéficié d'un volume sensiblement plus élevés de biens et services, tant dans la dimension pédagogique que de ce qui lui est périphérique pour assurer le fonctionnement des services.

Les dépenses salariales connaissent une évolution de 1 193 euros en 1974 à 3 219 Euros en 2005. Il est dans notre questionnement tout à fait pertinent de distinguer dans ces dépenses, d'une part les coûts associés aux personnels enseignants et, d'autre part les coûts relatifs aux personnels qui assurent les services annexes. L'évolution concernant le coût unitaire salarial spécifique des services annexes se traduit par une variation de 67 euros en 1974 à 263 euros en 2005, une multiplication par un facteur 3,90 sur la période. Cette évolution résulte a priori, à la fois d'une augmentation du rapport entre le nombre de ces personnels et celui des élèves (davantage de services assurés) et d'une augmentation du niveau moyen de salaire de ces personnels lorsque celui-ci est exprimé en valeurs monétaires constantes (gain de pouvoir d'achat).

La part principale de l'évolution globale du coût unitaire revient aux personnels d'enseignement. Il est estimé que l'évolution spécifique des dépenses salariales d'enseignement correspond à des valeurs de 1 126 euros en 1974 et de 2 956 euros en 2005,

<sup>16</sup> Il s'agit des dépenses de transport scolaire, de livres et de fournitures, de cantines etc...

soit une multiplication par un facteur de 2,63 sur la période considérée. Cette augmentation globale de la masse salariale unitaire de ces personnels résulte, d'une part de variations dans le rapport élèves-maître (il est passé de 25,5 en 1974 à 16,8 en 2005<sup>17</sup>, manifestant un coefficient spécifique d'augmentation du coût unitaire de 1,52<sup>18</sup>) et, d'autre part, de l'augmentation du niveau moyen de salaire des personnels concernés. L'évolution du niveau moyen de rémunération de ces personnels résulte à son tour de deux composantes, l'une spécifique concernant la création du corps des professeurs des écoles en 1990<sup>19</sup>, et l'autre plus générale qui correspond aux gains de pouvoir d'achat des salariés de la fonction publique.

La première composante peut être évaluée sur la base de la proportion des professeurs des écoles au sein des enseignants du premier degré (elle est estimée à 11 % en 1994 pour atteindre 45 % en 2000 et 85 % en 2005). On doit aussi tenir compte des niveaux salariaux respectifs à ces deux corps de fonctionnaires (instituteurs et professeurs des écoles<sup>20</sup>). Les chiffres de notre estimation montrent que l'impact spécifique de la revalorisation des enseignants du premier degré sur les coûts unitaires serait de l'ordre de 15 % en 2005.

La seconde composante est a priori difficile à évaluer de façon directe. Nous avons procédé de façon indirecte en adoptant un raisonnement par déduction. L'amélioration du rapport élèves-maître implique une hausse du coût unitaire de fonctionnement de 52 % et la revalorisation une augmentation de 15 %, ces deux facteurs combinés entraînent alors un accroissement du coût unitaire de 75,3 % sur la période 1974-2005 (soit un facteur d'augmentation de 1,75). Comme notre estimation du facteur global d'augmentation du coût unitaire salarial d'enseignement est de 2,63, on en déduit que le facteur relatif au gain de pouvoir d'achat sur la même période est de 1,48 (2,63 – 1,75), manifestant un gain de 50 % sur l'ensemble de la période.

Ce constat doit être interprété avec une double perspective. La première consiste à rapprocher l'évolution du niveau de rémunération des enseignants (indépendamment de la revalorisation) de l'évolution générale de la richesse du pays (mesurée par le PIB par habitant). On observe alors que le gain de PIB par habitant (en valeur monétaire constante) est de 71 % sur la période tandis que la progression du pouvoir d'achat des enseignants (des fonctionnaires en

---

<sup>17</sup> La diminution forte de cette statistique entre 1974 et 1984 provient pour une part essentielle de l'amélioration des conditions d'encadrement dans le préscolaire (il y avait, en moyenne, 40 enfants par enseignant en 1974).

<sup>18</sup> Cette augmentation, et donc le poids de facteur, aurait été sensiblement moins importante si on avait utilisé les années 1984-2005, plutôt que les années 1974-2005, comme période de référence pour conduire cette analyse.

<sup>19</sup> Décret n° 90-680 du 1er août 1990.

<sup>20</sup> Ce dernier point est une estimation car nous ne connaissons pas précisément la structure des âges des deux corps.

général) n'aurait été que de 50 % sur la période. La seconde perspective renvoie à la question de l'évolution entre 1974 et 2005. Les données estimées dans l'avant dernière ligne du tableau précédent suggèrent que la progression n'a pas été linéaire et que si la période allant de 1974 à 2000 a été globalement favorable (gain de pouvoir d'achat des enseignants, même si celui-ci est inférieur à l'évolution de la richesse nationale), ce n'est pas le cas de la période allant de 2000 à 2005 qui manifeste probablement une baisse du pouvoir d'achat des enseignants du premier degré.

Nous pouvons à présent revenir à l'analyse de l'évolution des dépenses salariales annexes. Si on applique le coefficient de gain de pouvoir d'achat des enseignants (hors impact de la création du corps des professeurs des écoles) aux personnels ayant des fonctions annexes à l'enseignement, on aboutit alors à ce que le facteur global de 3,90 résulte d'une augmentation salariale de 50 % (gain de pouvoir d'achat) et d'une augmentation en quantité par un facteur 2,40. Selon ces estimations, cela suggère qu'il y a eu une forte augmentation du nombre de ces personnels, notamment depuis 1994 et surtout depuis l'année 2000. Il faut bien sur considérer les chiffres du tableau précédent comme des ordres de grandeur mais qui permettent néanmoins d'expliquer la très forte augmentation du coût unitaire de l'enseignement du premier degré en France au cours des trente dernières années.

Le tableau suivant, à titre de synthèse, propose une répartition des 2 511 euros d'augmentation du coût unitaire (en euros de 2005) entre 1974 et 2005 en fonction des différents facteurs. On distingue ce qui tient à des améliorations du contexte de scolarisation (davantage de biens et services, meilleur rapport élèves-maître, meilleur niveau de formation des enseignants, ..) de ce qui est lié à des augmentations de prix (de pouvoir d'achat des personnels).

*Tableau 9 : Décomposition de l'augmentation du coût unitaire de scolarisation du premier degré entre 1974 et 2005*

	Ecart entre 2005 et 1974	Contexte de scolarisation	Augmentation des prix
<b>Coût unitaire de fonctionnement public (euros 2005)</b>	<b>2 511</b>	<b>1685</b>	<b>826</b>
Autres dépenses unitaires (euros 2005)	485	485	
Dépenses salariales annexes (euros 2005)	196	138	58
Dépenses salariales enseignement (euros 2005)	1830		
Rapport élèves-maître du 1 <sup>er</sup> degré public		795	
Impact salarial de la revalorisation		268	
Gain de pouvoir d'achat			767

Les aspects d'amélioration du contexte d'enseignement comptent pour environ les deux tiers (1 685 / 2 511) de l'augmentation globale du coût unitaire entre 1974 et 2005, alors que les

augmentations de prix comptent pour le tiers restant. A l'intérieur des facteurs d'amélioration du contexte scolaire, la réduction du rapport élèves-maître (795 euros de 2005) a le poids le plus important suivi par l'augmentation des dépenses non salariales (485 euros); vient ensuite, mais avec un poids inférieur (268 euros), l'impact de la réforme visant à la transformation progressive du corps des instituteurs en professeurs des écoles. La question reste évidemment posée de savoir dans quelle mesure ces améliorations du contexte d'enseignement ont correspondu à des gains d'efficacité, c'est-à-dire à des améliorations visibles dans les résultats du système et notamment en ce qui concerne le niveau moyen d'acquisition des élèves. A ce sujet, le début de ce chapitre fournit une réponse partielle à cette question. Le dernier point abordé dans ce chapitre mobilise une perspective comparative internationale pour avoir une idée encore plus large des coûts de notre école primaire.

#### ***II. 4 Les coûts unitaires dans une perspective comparative internationale***

Ici également, la mise en perspective des statistiques d'un système éducatif d'un pays donné avec celles d'autres pays est toujours intéressante mais sujette à des difficultés méthodologiques. En effet, il n'existe pas une manière unique de procéder et pour donner davantage de robustesse à l'analyse, nous utiliserons de façon complémentaire les trois méthodes les plus couramment employées dans la littérature sur le sujet. La première consiste à mesurer la dépense unitaire dans chacun des pays et à l'exprimer dans une unité monétaire commune en prenant soin de faire l'évaluation en termes de parité de pouvoir d'achat (PPA) pour éviter les questions liées à la variabilité des taux de change ; on compare alors la valeur du coût unitaire dans un pays donné avec celle observée dans les autres pays.

Les chiffres obtenus par la première méthode restent caractérisés par une relation assez forte avec le niveau de richesse des différents pays (fut-elle évaluée en parité de pouvoir d'achat), la seconde méthode diminue l'influence de cette relation avec le niveau de richesse des différents pays en proposant une estimation du coût unitaire qui contrôle l'effet de ce facteur contextuel «parasite». Sur le plan pratique, cela revient à estimer, pour un pays donné, le niveau de coût unitaire qu'a, en moyenne, un pays de même niveau de PIB par habitant; on compare alors les valeurs, observées et estimées, pour un même pays. La troisième méthode consiste à calculer directement le rapport entre la valeur du coût unitaire de scolarisation observé dans un pays et celle de son PIB par habitant ; ce rapport, qui est une mesure du coût unitaire exprimé en unités de PIB par habitant, est alors la base de comparaison entre pays<sup>21</sup>.

Le tableau suivant propose les éléments de comparaison des coûts unitaires de scolarisation aux niveaux du préscolaire, de l'élémentaire (primaire dans la classification internationale) et

---

<sup>21</sup> C'est cet indicateur qui a été utilisé précédemment pour apprécier l'efficacité qualitative des pays.

du secondaire dans les pays de l'OCDE et de l'Union Européenne. Les données concernent l'année 2003 et les pays choisis sont les mêmes que ceux pris en compte auparavant dans le tableau 2<sup>22</sup>.

Tableau 10 : Eléments de comparaison des coûts unitaires de scolarisation (2003)

	Pré primaire			Primaire			Secondaire			Rapport Secondaire / Primaire
	USD en PPA	USD en PPA Estimés	En PIB/tête	USD En PPA	USD en PPA Estimés	En PIB/tête	USD en PPA	USD en PPA Estimés	En PIB/tête	
<b>France</b>	<b>4 744</b>	<b>4 860</b>	<b>16,7</b>	<b>4 939</b>	<b>5 916</b>	<b>17,4</b>	<b>8 653</b>	<b>7 554</b>	<b>30,5</b>	<b>1,75</b>
Allemagne	4 865	4 768	17,6	4 624	5 753	16,7	7 173	7 332	26,0	1,55
Belgique	4 663	5 071	15,5	6 180	6 288	20,5	7 708	8 058	25,6	1,25
Danemark	4 824	5 144	15,7	7 814	6 415	25,5	8 183	8 230	26,7	1,05
Etats-Unis	7 755	5 983	20,7	8 305	7 895	22,1	9 590	10 236	25,6	1,15
Italie	6 116	4 638	23,0	7 366	5 524	27,7	7 938	7 022	29,9	1,08
Japon	3 766	4 823	13,4	6 350	5 851	22,6	7 283	7 465	25,9	1,15
Norvège	3 895	5 950	10,5	7 977	7 836	21,4	10 919	10 156	29,3	1,37
Royaume-Uni	7 153	5 012	24,2	5 851	6 184	19,8	7 290	7 917	24,6	1,25
Suède	4 091	5 002	13,9	7 291	6 165	24,7	7 662	7 891	26,0	1,05
Suisse	3 558	5 456	10,7	8 131	6 965	24,5	12 209	8 976	36,8	1,50
Moyenne OCDE	4 508		18,4	5 450		20,5	6 962		26,0	1,28
Moyenne UE-19	4 589		19,4	5 399		20,4	6 961		26,0	1,29
<b>France/moyenne UE</b>	<b>1,03</b>		<b>0,86</b>	<b>0,91</b>		<b>0,85</b>	<b>1,24</b>		<b>1,17</b>	<b>1,36</b>
<b>France / sa valeur estimée</b>		<b>0,98</b>			<b>0,83</b>			<b>1,15</b>		

Les deux dernières lignes du tableau (chiffres en gras) rendent compte de la situation de la France selon les trois méthodes mentionnées et pour chacun des niveaux d'enseignements. D'un point de vue strictement statistique, la seconde méthode est probablement la plus pertinente (USD en PPA estimés). Selon cette méthode, les coûts unitaires de scolarisation de la France seraient très proches de la moyenne de l'Union européenne au niveau préscolaire (rapport de 0,98 entre valeur observée et estimée) et seraient à un niveau plutôt faible dans l'enseignement primaire (rapport de 0,83 manifestant un écart de 17 % en dessous de la référence internationale). A l'inverse, les coûts seraient relativement élevés au niveau secondaire, le rapport entre valeurs observées et estimées étant alors de 1,15 (la valeur observée en France se situe 15 % au-dessus de la référence comparative internationale).

La situation respective des coûts unitaires dans le primaire et le secondaire est d'une certaine façon confirmée lorsqu'on calcule le rapport des coûts unitaires de ces deux niveaux d'études (colonne de droite dans le tableau); ce rapport vaut en effet 1,75 pour la France contre seulement 1,36 pour la moyenne de l'Union Européenne, aucun pays n'ayant en fait une

<sup>22</sup> Les données de base concernant ce tableau sont extraites de « Regards sur l'Éducation », OCDE, 2006.

valeur plus élevée que la France pour ce rapport au sein des pays de l'Union. On notera enfin que ces conclusions ne sont pas contredites par l'utilisation des deux autres méthodes, les chiffres obtenus suggérant une appréciation relativement homogène selon l'une ou l'autre des trois méthodes utilisées.

Les données internationales et nationales analysées et commentées dans ce premier chapitre montrent que le niveau moyen des élèves français stagne. La position de la France dans le contexte international s'est même plutôt dégradée au cours de ces dix dernières années dans le domaine de la langue écrite. Quand on relie ce niveau d'acquisition avec les ressources allouées à l'école, on observe là encore une situation peu favorable de la France dans le contexte international. Cela se traduit par une faible efficience, à la fois qualitative (relation entre les coûts et le niveau d'acquisition des élèves), et quantitative (relation entre les coûts et la durée moyenne de scolarisation). Pourtant, l'augmentation des coûts unitaires a plutôt avantage l'école primaire eu égard aux accroissements des effectifs d'élèves aux différents niveaux d'enseignement. Les conditions d'enseignement se sont donc progressivement améliorées au cours de ces trente dernières années (meilleurs taux d'encadrement, augmentation des moyens, revalorisation des salaires...) sans que le niveau d'acquisition des élèves connaisse une progression notable. Une première constatation générale est donc l'incapacité de notre école primaire à transformer efficacement les ressources en résultats.

Pour poursuivre notre logique d'analyse, il convient à présent d'examiner, toujours avec une perspective macro, les facteurs qui peuvent expliquer les performances des élèves et les inégalités scolaires et sociales de réussite. Nous nous centrerons dans le chapitre suivant sur l'organisation globale de notre système éducatif en utilisant à nouveau les comparaisons internationales.

*Chapitre 2*  
*Organisation scolaire et inégalités entre élèves*

L'objectif de ce chapitre est d'étudier comment l'organisation générale du système éducatif peut participer à expliquer les inégalités sociales de réussite. Il s'agit d'examiner, en complément du chapitre précédent, comment le système éducatif français se positionne dans ce domaine par rapport à des organisations scolaires variées mises en places dans les autres pays. Les comparaisons internationales constituent ici un outil indispensable pour déceler l'influence de facteurs de contexte que ne peuvent prendre en compte les travaux élaborés à partir de données exclusivement nationales. Les constructions théoriques fondées sur ce type de données se limitent alors à une « sociologie locale » qui ne permet pas d'explorer le fonctionnement concret de notre institution scolaire et de le confronter à celui des pays comparables.

Si les comparaisons internationales en matière d'inégalités sociales se sont multipliées au cours de cette dernière décennie, elles se sont surtout centrées sur le déroulement des carrières scolaires et sur le concept de mobilité sociale, plus que sur des mesures précises des acquisitions des élèves. Dans cette perspective, la place des inégalités de compétences ou d'acquis scolaires est plutôt minimisée. Ainsi, il a été montré que dans la genèse des inégalités sociales de carrières scolaires, les inégalités d'acquis ne pèsent que pour environ la moitié (Erikson, Jonsson, 1996), et que dans la reproduction des inégalités sociales d'une génération à l'autre, les inégalités de carrière professionnelle tout au long de la vie pèsent davantage que les inégalités tenant à l'obtention d'un niveau donné de formation initiale (Goux, Maurin, 1997).

Cette approche, dominante dans la sociologie européenne, considère de fait la formation avant tout de manière instrumentale, comme vecteur d'accès à une position sociale et dans sa rentabilisation sur le marché du travail. Une alternative moins répandue est de s'intéresser à ce qu'acquièrent les élèves, comme un bien en soi, indépendamment de la façon dont ces acquis se traduiront en carrière scolaire ou en statut social. Les données des enquêtes internationales sont alors précieuses, puisqu'elles entendent appréhender le niveau de compétences des élèves à la fois pour la compréhension de concepts fondamentaux et pour la capacité à les utiliser dans leur vie quotidienne, comme c'est le cas dans le programme PISA. C'est cette dernière enquête que nous allons mobiliser pour traiter notre questionnement. L'âge des enquêtés (15 ans) permet d'obtenir une image globale de l'expérience scolaire accumulée depuis l'école maternelle jusqu'à la fin du premier cycle de l'enseignement secondaire. Au-delà d'une exploitation spécifique des données disponibles dans l'enquête PISA, seront mobilisées des informations supplémentaires, accessibles dans les banques de données internationales ou collectées spécifiquement par d'autres voies (Duru-Bellat, Mons, Suchaut, 2003).), l'objectif étant de rapporter les inégalités constatées aux caractéristiques variées des contextes, cet aspect étant d'ailleurs très peu développé dans les rapports officiels sur PISA (OCDE, 2001a, 2001b, 2001c, 2003) des différents pays ayant participé à l'enquête.

L'enquête PISA explore trois dimensions des compétences des élèves, à savoir la compréhension de l'écrit, la culture mathématique et la culture scientifique. Nous n'utiliserons pour notre part que les résultats relatifs à la première dimension (compréhension de l'écrit) sachant que la corrélation entre les performances moyennes des élèves des différents pays est très élevée puisque supérieure à +0,95 (Duru-Bellat, Mons, Suchaut, 2003). Autrement dit, le classement des pays est très peu variable selon que l'on considère telle ou telle dimension évaluée. Par ailleurs, l'enquête PISA à laquelle nous faisons référence ciblait en priorité les compétences en compréhension de l'écrit. Avant de présenter les résultats spécifiques à notre questionnement, il est utile de s'arrêter auparavant sur les intérêts et les limites des comparaisons internationales en éducation, bien que celles-ci aient déjà été mobilisées dans le chapitre précédent. La mise en perspective des caractéristiques des systèmes éducatifs avec les résultats observés au niveau des élèves pose des problèmes particuliers qu'il convient de cerner et qui sont présents dans toute approche comparative.

## **I Les problèmes associés à l'approche comparative**

Le premier problème sur lequel butent les comparaisons est de trouver un critère raisonnablement comparable à l'aune duquel comparer les effets des politiques (entendues au sens large, en incluant les choix structurels, les dispositifs pédagogiques, les programmes...). En l'occurrence, les données du programme PISA s'avèrent précieuses, qui fournissent, sur la base d'exercices communs, des mesures des compétences manifestées par les jeunes de 15 ans dans un nombre important de pays. Un autre intérêt de ces données est qu'elles portent sur l'ensemble des jeunes scolarisés à cet âge, quelle que soit la filière, ce qui n'est pas toujours le cas dans les travaux comparatifs antérieurs ; par exemple, dans l'enquête TIMSS (1994), il s'agissait d'observer les connaissances en mathématiques et sciences des élèves scolarisés au niveau 4<sup>ème</sup> et non les élèves d'un âge donné (ce qui avantagait la France et les pays pratiquant le redoublement par rapport aux pays qui ne le pratiquent pas). Evidemment, ces particularités de l'enquête PISA ont leur revers. Portant sur les élèves scolarisés à 15 ans, elles vont concerner un pourcentage d'une génération variable selon les pays, proche de 100% dans les pays de l'OCDE les plus riches, mais bien plus faible dans des pays comme la Thaïlande, le Brésil et le Mexique, avec de multiples problèmes afférents. Un autre inconvénient est que les données de l'enquête PISA ne seront pas parfaitement comparables avec les enquêtes internationales antérieures, comme TIMSS, puisqu'il s'agit ici de compétences non directement liées aux programmes.

A cela s'ajoutent les difficultés habituelles afférentes à toute mesure des acquis des élèves, inévitablement imparfaite : toute la littérature docimologique montre qu'on ne sait pas mesurer sans marge d'erreur les acquis des élèves. Il est banal de constater que toute performance manifeste est contingente au support et au mode de questionnement qui a servi à l'évaluer, dans une certaine mesure, toute la question étant d'évaluer ce « dans une certaine mesure ». Il faut évidemment compter aussi avec les erreurs d'échantillonnage et donc les marges de fiabilité des estimations chiffrées, et par conséquent se montrer prudent avant de conclure à l'existence d'écarts significatifs entre pays, d'autant plus que certaines particularités techniques de l'enquête peuvent elles aussi produire des biais non négligeables (Monseur, 2006).

Pour revenir à la question, centrale, de ce qui est mesuré, il faut rappeler qu'une compétence n'est pas directement observable, ce qui contraint à une définition opérationnelle comme celle qui est classiquement attribuée à Binet au début du XX<sup>ème</sup> siècle : « l'intelligence, c'est ce que mesure mon test »... Cela n'empêche pas d'intégrer des considérations théoriques (en l'occurrence sur ce qu'est l'acte de lire) dans la définition des compétences en cause. Il faut néanmoins conclure que la question de savoir ce que serait une compétence mesurée sans biais (sans influence systématique de l'appareil de mesure) est des plus ouvertes... Il reste qu'il n'est pas toujours aisé de passer des programmes d'enseignement à des objectifs opérationnels, puis à la ou aux compétences que chacun d'entre eux mobilise ; certes le fait de dépasser une approche immédiate par connaissances (comme le font d'ailleurs les évaluations nationales en France) et de se focaliser sur des comportements observables est en soi intéressant ; et, dans les comparaisons internationales, on peut admettre que quel que soit le pays, l'école forme à des situations professionnelles ou plus largement sociales assez comparables qui mobilisent des compétences communes.

Mais cette approche des compétences peut-elle se prétendre sans « biais » ? Cette notion de biais est couramment invoquée, notamment en France, pour pointer le fait que certains groupes seraient systématiquement lésés du fait du caractère particulier de la mesure, celle-ci valorisant des savoirs ou savoir-faire que les élèves ont eu des opportunités inégales d'apprendre, et n'appréhendant pas par conséquent la compétence « réelle ». Ainsi, dans la sociologie des inégalités garçons-filles, on se demande, face à des résultats inférieurs des filles en sciences naturelles, si le fait que les épreuves exigent des manipulations au microscope, où les filles sont moins à l'aise, n'induit pas un biais de sexe. Cette question peut également être posée face aux performances seulement moyennes des élèves français à PISA : ne faut-il pas invoquer des « biais culturels » étant le fait d'une moindre familiarité avec un certain type d'exercice ou de « format » de question ? On écartera, en amont, la thèse, qui révèle une certaine conception des finalités de l'école, selon laquelle il serait illégitime (en France) de s'intéresser, au-delà de la connaissance des programmes et de la

réussite aux examens, à la capacité à maîtriser des tâches de la vie courante qui peuvent en être relativement éloignées (capacité elle-même sous-tendue vraisemblablement par la maîtrise de concepts fondamentaux)... Mais il est clair que chaque forme d'évaluations incorpore nécessairement certaines attitudes, savoirs et savoir-faire spécifiques.

Concernant PISA, on a souvent souligné que cette évaluation était largement imprégnée de la culture anglo-saxonne ; mais techniquement, toutes les précautions ont été prises (les matériaux d'évaluation ont été préparés à la fois en anglais et en français, et traduits par la suite dans les langues des pays participants, à partir de ces deux sources, dans la plupart des cas, le format des items a été diversifié autant que possible...), et on note par ailleurs qu'à la fois les pays dont les élèves sont le plus performants appartiennent à des aires culturelles très différentes (Corée, Finlande, Canada) et que les élèves des pays culturellement proches peuvent manifester des performances relativement inégales (les USA et le Canada par exemple).

La question de la réalité de biais culturels concernant spécifiquement la France reste néanmoins ouverte, non sans rapport avec les performances moyennes des élèves de notre pays ; si les publications du Ministère de l'Education invoquent régulièrement ce type de biais, la recherche peine à les identifier. Les chercheurs responsables du dispositif PISA (OCDE, 2001b) ont en effet ré-estimé les performances des élèves à la seule aune d'items considérés par les pays eux-mêmes comme particulièrement bien adaptés à leurs élèves, et cela ne modifie que fort peu la position de la France (qui passe de la 14<sup>ème</sup> à la 12<sup>ème</sup> position). Ceci constitue à l'évidence un argument fort à l'encontre de la thèse selon laquelle les résultats de l'enquête PISA seraient disqualifiés par l'existence de biais culturels.

Une question tout aussi importante mais moins souvent posée est de savoir, si biais il y a, s'ils sont jugés secondaires (ce ne sont que des biais), ou importants par rapport aux objectifs visés par les différents systèmes éducatifs. Si par exemple la manipulation d'un microscope est jugée seconde, on supprimera des évaluations le recours à cette technique pour supprimer les biais de sexe ; si par contre ce savoir-faire est jugé important, on s'attaquera à compenser le « handicap » des filles. De même pour les « biais culturels » : le constat de performances systématiquement inférieures des élèves français à certains sub-tests débouchera soit sur une simple dénonciation du caractère biaisé de l'exercice en cause, soit sur une réflexion sur les voies à suivre pour compenser cette faiblesse selon que ce que révèle l'exercice en question sera jugé accessoire ou important.

En ce qui nous concerne, la question essentielle est de savoir si les scores obtenus restent néanmoins raisonnablement comparables d'un pays à l'autre. On peut estimer que la réponse est positive, au niveau de « focale » où nous nous situons. De plus, même si le niveau moyen

peut être (légèrement) affecté à la hausse ou à la baisse par tel ou tel biais, dans chacun des pays, il est peu probable (mais ceci ne peut être exclu) que cela affecte de manière systématique la dispersion de ces scores notamment selon l'origine sociale, perspective centrale dans le présent travail.

Il sera même, dans la perspective qui est la nôtre, particulièrement justifié de travailler sur les scores moyens (en compréhension de l'écrit, en mathématiques...), qui minimisent les erreurs attachées à chaque sub-test, même s'il est clair que derrière tout score moyen il y a un ensemble de scores élémentaires susceptibles de varier de manière pas forcément cohérente. La comparaison conduite ici se fait à un niveau relativement agrégé, et vise avant tout, non pas à établir un palmarès (perspective en elle-même peu heuristique), mais à établir des relations entre tel score (et telle dispersion de ces scores en fonction de variables comme l'origine sociale) et telles caractéristiques des systèmes. Si dans une perspective de classement entre les pays, classement censé fondé sur les compétences « réelles » des élèves, les problèmes de traduction sont considérés comme des biais (Robin et Rocher, 2002), la question se pose différemment à propos du fait que les programmes et les cursus varient : si tant est que l'on s'accorde sur les compétences que devraient maîtriser les jeunes, alors cette variation des programmes est elle-même intéressante et elle interroge. La question des biais devient (encore) moins prégnante dès lors que l'on entend partir des différences (de moyennes et de dispersion) entre pays avant tout pour les expliquer par des facteurs sociaux, tout particulièrement par l'organisation des systèmes.

L'exercice consistant à mettre en relation ces mesures de scores moyens et de dispersion et les caractéristiques structurelles des systèmes (avec en filigrane l'objectif d'éclairer un certain nombre de questions de politiques éducatives) reste au demeurant plein d'écueils (ni plus ni moins cependant que celui consistant à expliquer les scores nationaux par les spécificités « culturelles » des exercices). Les différences entre pays, qu'il s'agisse de différences de scores moyens ou de différences dans les distributions, ne livrent pas, de manière immédiate et fidèle, un reflet des différences d'efficacité des systèmes (ou d'efficacité de tel ou tel paramètre du système), et elles exigent tout un travail d'interprétation ; elle résultent, de manière jointe et cumulée, chez les jeunes de 15 ans, des curricula mis en œuvre, mais aussi des structures éducatives ou des relations maîtres-élèves qui prévalent, des inégalités sociales globales qui font du public d'élèves une population elle-même plus ou moins diversifiée, etc. En d'autres termes, elles ne reflètent pas seulement, ce que les élèves ont acquis grâce à l'école mais aussi ce qu'ils ont appris par leur famille, leurs pairs, les médias, la vie quotidienne...

La recherche de corrélations entre compétences manifestées à 15 ans et modes d'organisation des systèmes pourra certes être guidée par certains résultats de la recherche : on dispose de « bonnes raisons » d'attendre plus d'inégalités sociales là où les structures sont différenciées de manière précoce, là où le temps scolaire alloué est limité, là où les contenus de formation sont proches de la culture classique... Mais dans la réalité, ces diverses caractéristiques des systèmes sont vraisemblablement liées et forment système. Un écueil classique des comparaisons internationales est de donner un sens à un élément isolé de son contexte. En particulier, il est clair qu'on peut contester la légitimité de corrélations rapprochant terme à terme telle performance et telle caractéristique du système éducatif. Même si tous les comparatistes n'adhèrent pas à la « méthode sociétale » (Maurice, Sellier, Silvestre, 1982), chacun s'accorde pour questionner la pertinence de la comparaison d'éléments isolés extraits de leur contexte, alors que sont vraisemblablement à l'œuvre des ensembles de traits constituant des cohérences sociétales spécifiques à chaque pays, structures globales sur lesquelles devrait porter la comparaison. Néanmoins, l'objectif des comparaisons reste bien de dégager des « lois » ayant une portée « supra-nationale » ; en l'occurrence, ici, des relations entre tel mode d'organisation du système et tel « produit » chez les élèves.

Il reste que, à partir de l'observation de corrélations, les imputations causales ne pourront être qu'extrêmement prudentes, bien qu'elles soient aussi répandues que teintées d'idéologie dans les travaux comparatifs ; c'est le cas quand l'OCDE (2002) rapproche en suggérant, malgré des précautions rhétoriques, une relation de cause à effet, les bonnes performances de certains pays et le fait que les établissements y jouissent d'une certaine autonomie ; ou encore quand Damoiselet, Lévi-Garboua (2001) concluent au « déclin apparent des niveaux éducatifs » (tel qu'évalué par des experts) comme « conséquence du report dans le temps de la différenciation par les capacités » alors que seule une corrélation est visible. Il est donc prudent de ne pas conclure trop vite à une relation causale entre tel facteur et tel phénomène, la démarche la plus sûre étant sans doute, en construisant des typologies, de rechercher quels sont les ensembles de facteurs qui jouent de concert.

Il nous reste à aborder un ensemble d'autres difficultés, elles aussi très classiques dans les travaux comparatifs. Ainsi, une difficulté de méthode courante concerne la qualité des données qu'il va être possible de mobiliser, quant aux caractéristiques des systèmes. Nous reviendrons sur la question, triviale mais élémentaire, de la disponibilité des informations dans les sources existantes : pour des raisons parfois peu claires, certaines données sont manquantes dans les fichiers de l'UNESCO ou de l'OCDE ; les compléter en enrichissant les bases existantes n'est alors pas dépourvu d'enjeu scientifique et revêt un intérêt autonome, même si cela conduit à « mixer » ainsi des données d'origine différente et peut-être imparfaitement comparables. Plus délicate est la question de l'écart qui peut exister entre une caractéristique formelle et le fonctionnement réel d'un système ; ainsi, au vu des données

administratives existantes, un système peut être formellement indifférencié au niveau de l'école moyenne, et fonctionner de fait avec des classes de niveau. Une solution pourrait être de croiser, chaque fois que cela est possible, les informations formelles avec, soit les résultats de la recherche sociologique, soit les opinions d'experts ; mais en aucun cas on ne saurait retenir comme unique source ces opinions, comme le font pourtant certains travaux comparatifs (Damoiselet, Lévi-Garboua, 2001).

Dans tout travail comparatif, on n'échappe pas à un arbitrage entre le « souhaitable » et le « faisable », à la question de la disponibilité et de la qualité des données : faut-il abandonner toute idée de comparaison sous ce prétexte ou au contraire, tenter de collecter avec prudence des données en les vérifiant (même si ces données sont imparfaites). Sachant de plus que l'imperfection est plus ou moins marquée selon le type d'indicateurs, avec du même coup le risque que les modèles économétriques donnent plus de poids aux variables bien mesurées par rapport aux variables moins bien mesurées.

Enfin, en éducation comparée comme dans d'autres domaines des sciences sociales, les résultats seront contingents à l'échantillon retenu, à son degré de variété notamment. Une caractéristique universellement répandue dans l'échantillon de pays étudiés ne pourra de ce fait voir son influence estimée ; de même, si on mêle dans une même analyse des facteurs qui varient peu et d'autres qui varient fortement, d'un pays à l'autre, alors, le poids des seconds s'exprimera davantage que celui des premiers. En d'autres termes, non seulement on n'évalue que ce qui varie, mais on l'évaluera d'autant mieux, si l'on peut dire, que cela varie. Certaines relations pourront ainsi, nous en verrons des exemples, être « tirées » par quelques pays en position extrême ; ainsi, si l'on inclut dans l'échantillon de pays étudiés, des pays très riches et des pays très pauvres, la richesse du pays pourra s'avérer influente sur le fonctionnement du système (ou les scores des élèves), alors qu'elle ne le serait pas sur un échantillon de pays plus homogènes à cet égard. De même, on ne sera pas surpris de voir que les facteurs de contexte associés significativement aux performances des élèves pourront être différents dans les pays pauvres et dans les pays riches (Mingat, Suchaut, 2000), puisque nombre de facteurs jouent seulement s'ils n'atteignent pas un niveau minimum (tel niveau de formation des maîtres, tel nombre de livres par élève). A cet égard, l'ajout dans les données dites « PISA Plus » (OCDE, 2003) d'une dizaine de pays de niveau de richesse et d'horizon culturel très différents (Albanie, Indonésie, Israël...) est très heuristique. Pour autant, l'échantillon de pays ainsi élargi (soit 40 pays) ne reflète pas la variété des pays qui existe à l'échelle de la planète : en particulier, il n'inclut aucun pays d'Afrique ou du Moyen-Orient, ni de nombreux pays asiatiques (Inde, Chine, Pakistan, Malaisie, Bhoutan...) ; il est néanmoins représentatif des pays industrialisés et émergents.

## **II Inégalités sociales de réussite et caractéristiques des systèmes éducatifs**

Les inégalités sociales de compétences, dans les différents pays, se situent sur un « fond de carte » structuré par l'ampleur des inégalités entre élèves. Dans un pays dont les élèves ont des performances très homogènes, les inégalités tenant à l'origine sociale de ces derniers auront moins de latitude pour s'exprimer. Une première étape est donc de décrire les performances moyennes des pays et la diversité de leurs résultats.

### ***II.1 Les Inégalités de performances***

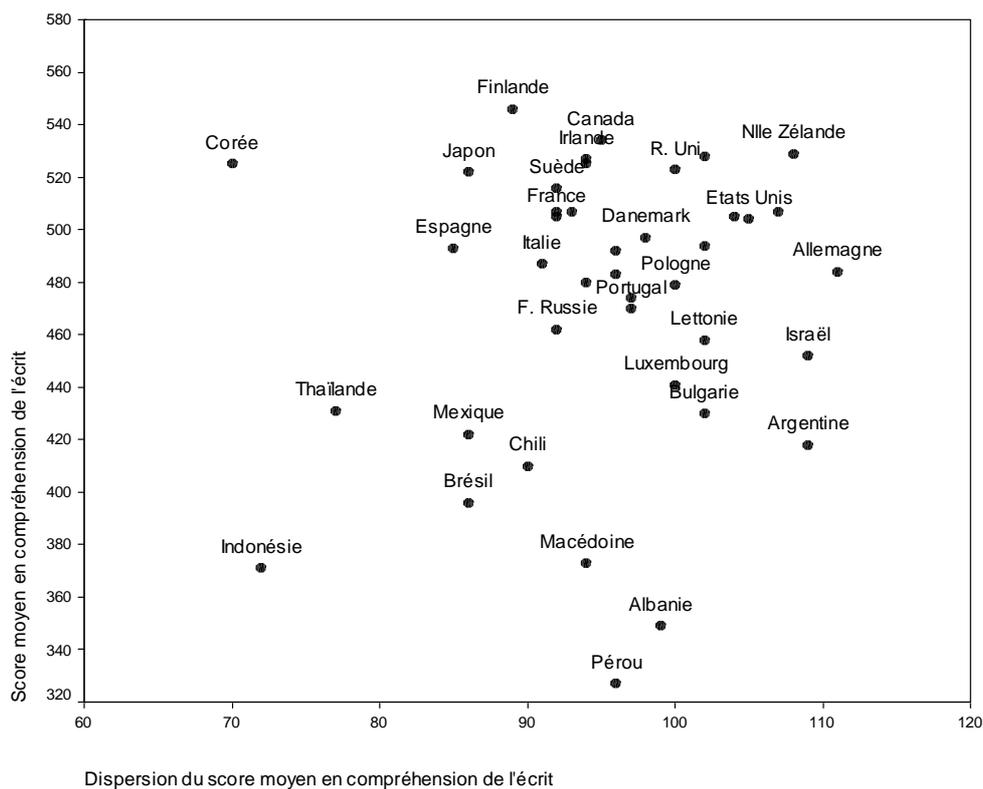
Sur l'ensemble des pays, la distribution des scores des élèves est exprimée dans une échelle arbitraire avec une moyenne de 500 et un écart-type de 100 comme c'est le cas dans la plupart des enquêtes internationales. En ce qui concerne la compréhension de l'écrit, les scores moyens s'échelonnent de 327 pour le Pérou à 546 pour la Finlande (cet écart correspond donc à 2,1 écart-type dans la distribution des scores entre élèves). En mathématiques, les scores moyens vont de 292 pour le Pérou à 560 pour Hong Kong (soit 268 points d'écart) ; et enfin, dans le domaine des sciences, on note un score moyen de 333 pour le Pérou et de 552 pour la Corée. En France, les scores moyens sont de 505 en compréhension de l'écrit, de 517 en mathématiques et de 500 en sciences, soit une position générale proche de la moyenne des pays de l'OCDE. Comme nous l'avons déjà mentionné, les classements des pays enquêtés dans PISA sont très proches quelle que soit la discipline évaluée, mais il existe toutefois quelques pays qui affichent des performances variables selon la discipline considérée, notamment lorsque l'on raisonne en terme de scores et non plus de rangs. C'est le cas notamment des pays comme le Japon ou la Corée qui sont légèrement moins performants en compréhension de l'écrit qu'en mathématiques ou en sciences (en écrit, ils ne se situent qu'autour des sixième et neuvième rangs). On pourrait également citer le cas du Danemark relativement plus faible en sciences (score de 481) qu'en mathématiques (score de 514).

Si ces différences de moyennes fournissent une première idée des différences globales de performances entre pays, elles demeurent limitées pour deux raisons. La première est liée à l'ampleur des différences entre pays : mis à part les cas extrêmes, elles sont parfois minimes et non significatives sur le plan statistique, ce qui rend alors circonspect quant à tout classement entre pays. Ainsi, pour la compréhension de l'écrit, des pays comme la France (score moyen de 505), le Danemark (score de 497), la Suisse (score de 494), les Etats-Unis (score de 504) peuvent être considérés comme équivalents sur le plan des performances. Les erreurs d'échantillonnage au sein de chaque pays sont parfois très importantes, ce qui conduit à considérer par exemple que le Japon (score moyen de 522) n'a pas des résultats significativement différents de ceux de la France.

La deuxième raison est liée à la variété des résultats au sein de chacun des pays, variété qui peut être considérable ; on est alors amené à analyser simultanément les performances moyennes et la dispersion de ces performances. Dans certains cas, cette dispersion intra-pays correspond en partie à des différences d'organisation des systèmes éducatifs selon les régions, les communautés et les états. Un exemple caractéristique à cet égard est la Belgique avec la communauté flamande et la communauté française, dont les résultats moyens sont assez éloignés. Ainsi, en compréhension de l'écrit, les scores sont respectivement de 532 et de 476 pour les deux communautés<sup>23</sup>, sachant que dans le même temps les deux systèmes sont différents. Au niveau de l'analyse où nous nous situons, il reste légitime de considérer la Belgique comme un seul pays au même titre que d'autres qui partagent à des degrés divers ces deux caractéristiques, le plus souvent en fonction des régions ou des états comme l'Australie, l'Allemagne... Le graphique suivant présente pour l'ensemble des pays considérés les deux indicateurs de performances en compréhension de l'écrit (moyenne et écart-type). On observe tout d'abord des pays pour lesquels les performances sont à la fois élevées en moyenne et relativement homogènes, c'est le cas de la Finlande et du Japon et encore bien plus nettement de la Corée qui est isolée en haut et à gauche du graphique. A l'inverse, dans d'autres pays, les performances sont à la fois faibles et dispersées comme en Argentine. Il existe aussi des pays dont les performances sont à la fois faibles et peu dispersées comme l'Indonésie.

---

<sup>23</sup> La communauté française de Belgique a d'ailleurs produit un rapport spécifique analysant les écarts entre les deux communautés (Lafontaine et alii, 2001).



Graphique 8 : Niveau moyen et dispersion en compréhension de l'écrit

Enfin, on remarque aussi des pays dont les élèves sont à la fois performants et inégaux (notamment la Nouvelle Zélande ou le Royaume-Uni). On peut aussi souligner que pour un même niveau moyen de performances, les inégalités de résultats entre élèves sont plus ou moins fortes d'un pays à l'autre : entre la Corée et la Nouvelle-Zélande pour les pays performants, entre l'Espagne et l'Allemagne pour les pays proches de la moyenne. Pour condenser en un indice unique les indicateurs de performances moyennes et d'inégalités entre élèves, on peut calculer un coefficient de variation (écart-type / moyenne) : plus cet indicateur est faible, plus le pays est à la fois performant et homogène : c'est le cas de la Corée, du Japon et de la Finlande avec des valeurs allant de 0,13 à 0,16. Dans quelques cas, un indicateur faible peut traduire des performances médiocres assorties d'une dispersion très réduite (Thaïlande, Indonésie). Plus cet indicateur est fort, plus le pays est de niveau médiocre et hétérogène, c'est le cas de l'Allemagne, du Luxembourg, de la Lettonie et du Brésil (avec des valeurs comprises entre 0,22 et 0,23), mais plus encore de l'Albanie (0,28) et du Pérou (0,29). Pour la France, le coefficient de variation vaut 0,18, soit un chiffre assez faible, égal à celui relevé au Canada, en Islande ou en Suède.

En compréhension de l'écrit (mais ce le cas également pour les deux autres disciplines évaluées), on est frappé par la situation particulière des pays les plus pauvres de l'échantillon : ces pays se démarquent fortement des autres par des performances nettement plus faibles. On

aura l'occasion plus loin dans ce texte d'aborder à nouveau cette question, mais on peut d'ores et déjà signaler que la prise en compte de ces pays affecte sensiblement les relations statistiques observées sur l'ensemble de l'échantillon (et qui ont fait l'objet du premier rapport PISA de 2001). Ainsi, si on conduit l'analyse uniquement sur les pays les plus riches de l'échantillon<sup>24</sup>, on note une corrélation négative entre le niveau moyen et la dispersion des scores (le coefficient de corrélation vaut -0,29), cela signifie, notamment dans les matières scientifiques, que les pays les plus performants ont également tendance, en moyenne, à être ceux dont les élèves ont les scores les moins dispersés. Or, cette relation statistique disparaît quand l'analyse porte sur l'ensemble des pays de l'échantillon PISA. Ceci s'explique par le fait que les pays les plus pauvres de l'échantillon ont des performances à la fois faibles et plutôt moins dispersées que les pays riches. Au total, il paraît établi que la notion d'arbitrage efficacité-équité communément avancée par les économistes (qui voudrait que l'on « paye » plus d'efficacité par davantage d'inégalités) n'est pas vérifiée ici. Certes, il ne s'agit pas ici véritablement d'efficacité du système, mais la relation entre performances moyennes et dispersion est clairement négative sur les pays riches du moins.

Une autre approche pour apprécier les performances est d'adopter une perspective critériée en se basant sur différents niveaux de compétences. PISA a identifié dans le domaine de l'écrit six seuils de score qui correspondent à des degrés hiérarchisés de maîtrise de la langue. Nous nous centrerons pour notre part sur des indicateurs rendant compte des groupes extrêmes avec le pourcentage d'élèves qui se situe en dessous du niveau 1 de compétence<sup>25</sup> et le pourcentage au-dessus du niveau 5<sup>26</sup>, soit respectivement les élèves en « difficulté » (pour utiliser la lecture comme un outil) et « l'élite » scolaire dans chaque pays. Globalement ces deux indicateurs sont liés de façon négative (corrélation de -0,66) ce qui signifie que le nombre d'élèves en difficulté est d'autant plus élevé que le nombre d'élèves faisant partie de l'élite est faible. Il existe quand même certaines spécificités au sein de l'échantillon comme on peut le voir sur le graphique suivant.

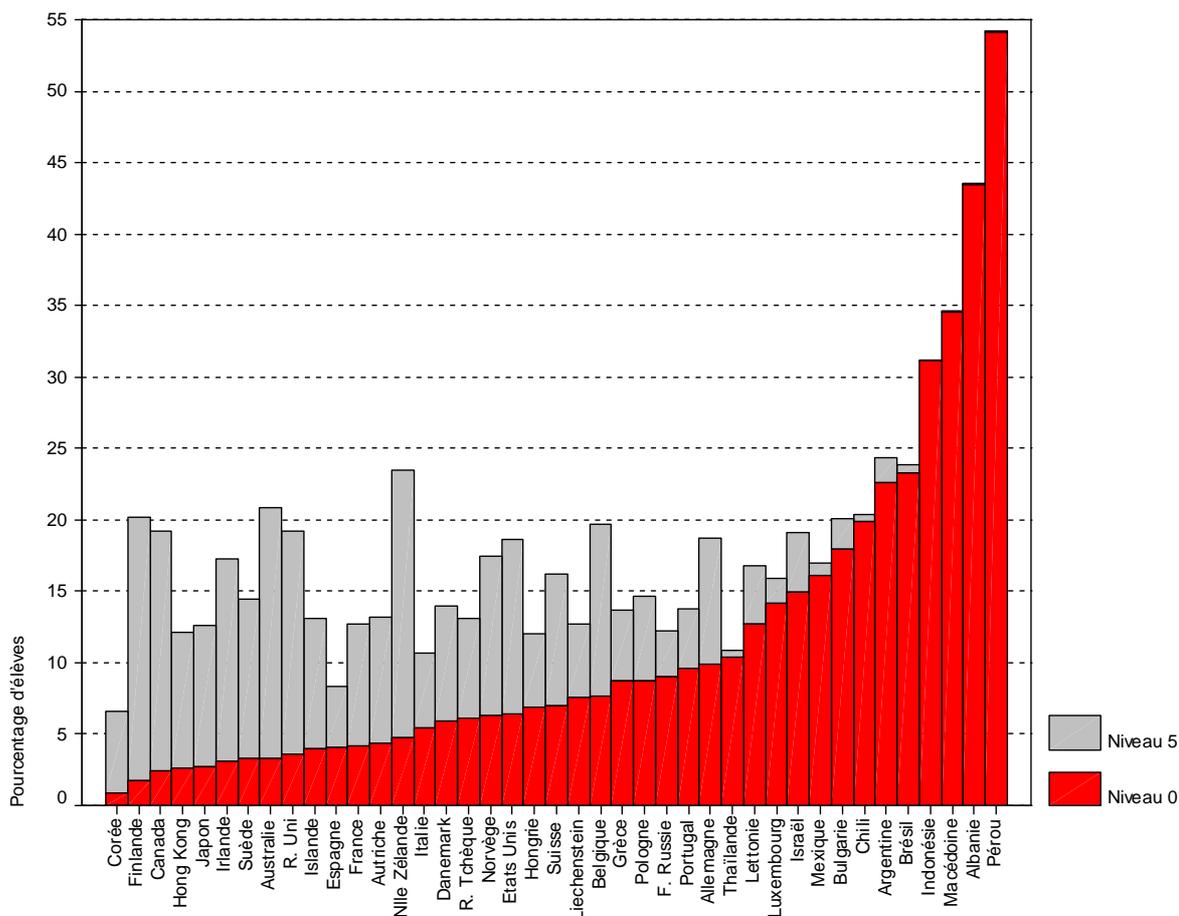
---

<sup>24</sup> Australie, Autriche, Belgique, Canada, République Tchèque, Danemark, Finlande, France, Allemagne, Grèce Hong-Kong, Hongrie, Islande, Irlande, Israël, Italie, Japon, Corée, Lettonie, Liechtenstein, Luxembourg, Nouvelle Zélande, Norvège, Pologne, Portugal, Fédération de Russie, Espagne, Suède, Suisse, Royaume Uni, Etats-Unis.

<sup>25</sup> Les élèves qui atteignent seulement ce niveau (score inférieur à 335) ne sont pas capables de mettre couramment en oeuvre les connaissances et compétences les plus élémentaires. Les élèves qui accèdent au niveau sont ceux uniquement capables d'effectuer les tâches de lecture les moins complexes des épreuves PISA, notamment : localiser un fragment unique d'informations, identifier le thème principal d'un texte ou établir une relation simple entre des connaissances de la vie courante.

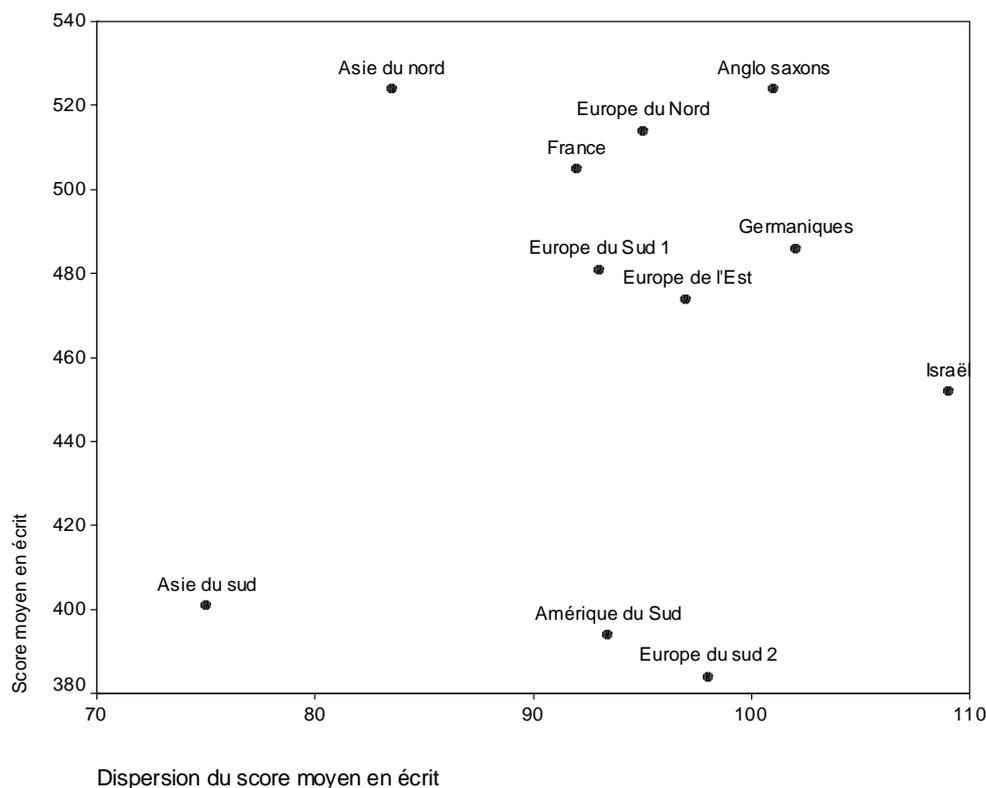
<sup>26</sup> Il s'agit des élèves ayant un score supérieur à 625 qui sont donc capables de mener à bien des tâches de lecture complexes (traiter des informations difficiles à trouver dans des textes non familiers, procéder à des évaluations critiques de ces informations....)

En moyenne, sur l'ensemble des pays, les pourcentages d'élèves en dessous du niveau 1 et au-dessus du niveau 5 sont respectivement de 11 et de 7. La Corée manifeste encore sous cet angle un resserrement très net des scores autour de la moyenne avec à la fois peu d'« élites » (5,7%) et quasiment pas d'élèves en difficulté (0,9%). La Belgique est dans une situation inverse avec à la fois un nombre d'élèves excellents et d'élèves faibles supérieurs à la moyenne (respectivement 12% et 7,7%). On peut aussi opposer des pays comme la Finlande et le Luxembourg, le premier pays cité présentant très peu d'élèves faibles (1,7%) et un fort pourcentage d'élèves au-dessus du niveau 5 (18,5%), le second pays, pour sa part, ayant une élite limitée en nombre (1,7%) alors que les élèves dotés d'un faible niveau sont plus nombreux (14,2%). Quant à la France, elle affiche une position moyenne pour la proportion d'élèves excellents (8,5%) et est en 11<sup>ème</sup> position quant au nombre d'élèves présentant des difficultés (4,2%). On remarque enfin la position particulière des pays les plus pauvres qui n'ont parfois aucun élève au-dessus du niveau 5 (Indonésie, Macédoine, Albanie et Pérou). Une chose est claire, certains pays parviennent à faire en sorte qu'aucun élève ne présente un niveau très faible en compréhension de l'écrit, ce qui atteste, si besoin en était, que les grandes difficultés scolaires n'ont rien de fatal ; cela indique également qu'a priori tous les pays pourraient atteindre ce niveau minimum de compétence, même si la suite de l'analyse montrera qu'à cet égard certaines conditions socio-économiques ou pédagogiques s'avèrent plus ou moins favorables.



Graphique 9: Pourcentages d'élèves en dessous du niveau 1 et au-dessus du niveau 5 de compétences en compréhension de l'écrit

Ces résultats qui concernent les extrémités de la distribution doivent bien sûr être rapportés aux scores moyens qu'ils affectent nécessairement. Ainsi, la Corée, avec une élite peu nombreuse, atteint néanmoins un score élevé, ce qui indique qu'elle compte beaucoup d'élèves aux niveaux intermédiaires (des niveaux 2, 3 et 4). La Nouvelle Zélande, avec un niveau moyen comparable a un pourcentage d'élèves forts bien supérieur à la Corée et un pourcentage d'élèves faibles également plus important. Pour un même niveau moyen, la structure des performances entre élèves peut donc être sensiblement différente d'un pays à l'autre. Ceci peut traduire des politiques spécifiques, visant à limiter le plus possible le poids des plus faibles ou, à l'autre bout du spectre, à développer une élite. Toujours est-il que cette diversité des pays présents dans l'enquête PISA incite à examiner les résultats par grandes zones économiques, géographiques et culturelles. On voit alors apparaître des « modèles » de performances distincts comme le montre le graphique 10.



*Graphique 10 : Niveau moyen et dispersion en compréhension de l'écrit par grandes régions*

Les pays asiatiques du Nord (ont en fait été regroupés les pays riches de cette zone) se distinguent des autres par des performances à la fois élevées et homogènes alors que les pays anglo-saxons (Royaume-Uni, Nouvelle Zélande, Australie, Etats-Unis...) ont des performances également élevées mais beaucoup plus hétérogènes. Sans anticiper sur la suite de l'analyse, on peut noter que ces pays sont souvent des pays d'immigration, ce qui peut expliquer pour une part l'hétérogénéité des performances de leurs élèves (cette remarque vaut aussi d'ailleurs pour Israël, qu'il s'est avéré impossible de regrouper avec un groupe de pays). Les pays nordiques ne sont, sur ces deux critères de moyenne et de dispersion, pas très éloignés des pays anglo-saxons tout en ayant des performances plus homogènes. Les pays de l'Europe de l'Est, germaniques et les pays les plus riches de l'Europe du Sud ne sont pas très différents en terme de performances moyennes (globalement moins élevées que les pays précédents), mais on peut souligner la plus forte dispersion des performances des élèves des pays germaniques. Les pays d'Amérique du Sud sont très atypiques dans cet ensemble, avec en particulier des performances moyennes très faibles. Ils ne sont pas très différents des pays les plus pauvres d'Europe du Sud (notés Europe du Sud 2 sur le graphique) : Macédoine, Bulgarie et Albanie. Les pays les plus pauvres d'Asie (notés Asie du Sud sur le graphique) c'est-à-dire Thaïlande et Indonésie sont de niveau moyen comparable aux deux groupes

précédents mais se distinguent par la relative homogénéité de leurs performances. Nous avons à dessein isolé la France qui, de fait, est relativement proche des pays d'Europe du Nord.

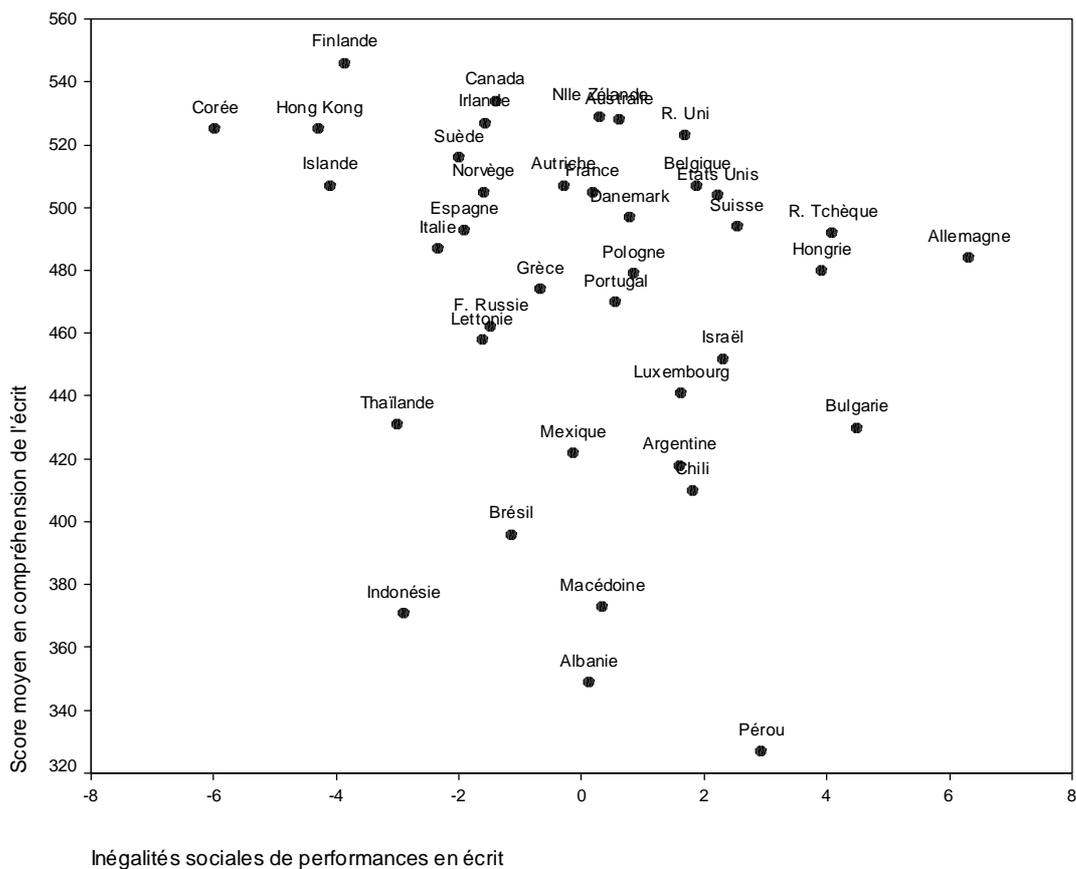
Cet exercice de comparaison (de score moyen, de dispersion) a néanmoins des limites : environ un dixième de la variance totale des performances des élèves dans PISA s'explique par l'appartenance à un pays, et peut donc être rapproché des différences de moyennes entre pays, elles-mêmes explicables tout d'abord par les contextes économiques globaux, mais aussi potentiellement par des facteurs d'organisation scolaire. Cela signifie, soit que ces facteurs structurels (ou plus largement les diverses caractéristiques des pays) n'ont que peu d'importance sur ce que les élèves finissent pas maîtriser comme compétences (à 15 ans), soit, et c'est sans doute plutôt le cas, que tous les systèmes des pays de l'OCDE pris en compte dans PISA partagent les mêmes caractéristiques qui s'avèrent essentielles en matière d'efficacité. Seule la prise en compte de pays (encore) plus différents, dans leur système éducatif ou leur richesse (etc.), permettrait de faire apparaître le rôle de tel ou tel facteur structurel et valider ce constat du dixième (qui est donc valable seulement sur l'échantillon de pays disponible dans PISA, pays qui ne représentent pas toute la variété existante à l'échelle de la planète). Certains travaux montrent par ailleurs que les facteurs d'organisation scolaire pèsent davantage sur les performances des élèves dans les pays pauvres (Heyneman, 1986).

## *II.2 Des différences de niveau moyen et d'inégalités sociales entre pays*

En ce qui concerne les inégalités sociales de performance, les variations entre pays sont également importantes. L'enquête PISA fournit certains indicateurs qui cernent la richesse matérielle ou culturelle des milieux familiaux et permettent d'évaluer l'impact du milieu social ainsi défini sur les performances des élèves, au sein des différents pays. Ce que nous cherchons à expliquer ici, c'est la variation, selon les pays, de la force de la relation entre le milieu socio-culturel de l'élève et ses résultats au test. Ces indicateurs ne sont pas indépendants les uns des autres et les classements sont en fait assez proches pour beaucoup de pays quel que soit l'indicateur retenu. Un premier indicateur du milieu social de l'élève est la variable « ISEI » qui exprime sous une forme continue la qualification professionnelle des parents; il est alors possible de comparer le niveau moyen des élèves dont le responsable familial appartient, soit au quartile inférieur ou au quartile supérieur ; on dispose ainsi d'un premier indicateur des inégalités sociales de réussite dans le pays. Une spécification complémentaire de l'influence d'ISEI sur les performances est le pourcentage de variance des résultats en écrit qui est expliqué par les variations d'ISEI. Un autre indicateur de la position sociale des parents des élèves est de nature composite : il tient compte, non seulement de ISEI, mais aussi du niveau d'éducation du parent le plus instruit, du niveau de vie de la famille (mesuré à partir des biens matériels possédés), du montant des biens culturels disponibles au foyer. Cette variable, nommée « IESCS » dans les données PISA, peut être

mise en relation avec la performance moyenne ; on peut alors calculer une droite de régression dont la pente est une mesure, comparable d'un pays à l'autre, de l'ampleur des inégalités sociales de performances. Enfin, un dernier indicateur de l'influence du milieu socio-culturel de l'élève est l'écart de performance selon que la mère de l'élève a un niveau d'éducation primaire ou secondaire inférieur (ISCED 1 et 2) ou au contraire un niveau de l'enseignement supérieur (ISCED 5 et 6). Afin de donner plus de consistance à la mesure des inégalités sociales de réussite, nous avons construit un indicateur synthétique (variable nommée « INEGA »), celui-ci s'exprime dans une échelle centrée réduite (moyenne de 0, écart-type de 1), et il a été calculé en effectuant la moyenne des indicateurs après que ces derniers aient été standardisés dans une même échelle.

Sur la base de cet indicateur global, examinons s'il existe une relation entre les inégalités sociales de performances en écrit et le niveau moyen des différents pays ; on pourrait en effet faire l'hypothèse qu'il y aurait concurrence entre des politiques visant soit à améliorer l'efficacité moyenne soit à réduire les inégalités sociales (ce qu'on désigne classiquement, chez les économistes, par l'arbitrage « efficacité/équité »). Le graphique ci-après montre que la variété de l'ampleur des inégalités sociales est faiblement associée au niveau moyen de performances, mais pas dans le sens attendu : la relation est en effet négative ( $r = - 0,26$ ) mais faible, à la limite de la significativité (seuil d'erreur de 11%).



Graphique 11 : Indice global d'inégalités sociales et niveau moyen de performance des pays

En d'autres termes, les pays les plus inégalitaires socialement sont aussi, un peu plus souvent, les pays où les élèves sont les plus faibles ; réciproquement, les pays les plus égalitaires socialement ont aussi en moyenne les élèves les plus performants. Toutefois, nombre de pays ont des niveaux d'inégalités comparables avec des performances variées. Du côté des pays socialement égalitaires, on peut opposer Finlande et Thaïlande dont les performances sont très différentes. De même, pour les pays plus moyens en termes d'égalité, le Danemark et l'Albanie et du côté des pays socialement plus inégalitaires, la Hongrie et le Pérou. Réciproquement, pour un même niveau de performances, on observe un niveau d'inégalités sociales très variable : Allemagne (socialement inégalitaires) versus Italie (qui l'est beaucoup moins), Royaume Uni versus Corée, Bulgarie versus Thaïlande.... On peut pointer la position de la France qui se situe de manière moyenne à la fois en termes d'inégalités sociales de performances et de niveau moyen de performances.

La lecture du graphique fait aussi apparaître de ressemblances assez fortes entre pays d'une même région. Ainsi les pays d'Asie du Nord présentent des performances élevées ainsi qu'un faible niveau d'inégalités sociales ; dans les pays d'Asie du Sud, l'effet de l'origine sociale

sur les performances est également assez ténu, mais le niveau moyen est nettement plus faible. En ce qui concerne les pays occidentaux, on remarque la position des pays anglo-saxons dont les performances moyennes sont assez élevées mais assez fortement marquées par l'origine sociale. Quant aux pays européens les plus pauvres (Macédoine, Albanie), ils sont très proches des pays d'Amérique du Sud avec à la fois des performances médiocres et des inégalités sociales assez importantes. On soulignera enfin qu'un certain nombre de pays pauvres (Indonésie, Thaïlande) ont des inégalités sociales de performances modérées, ce qui doit être considéré avec précaution dans la mesure où la scolarisation n'est pas universelle dans ces pays. Il est donc probable que les élèves qui sont encore scolarisés dans ces pays sont sélectionnés socialement et que, dans ce cas, les inégalités sociales de réussite vont y être plutôt minimisées. Mais la variété constatée, dans l'ampleur des inégalités sociales, à niveau de scolarisation donnée, convainc de l'intérêt d'une analyse autonome des facteurs susceptibles d'expliquer cette variété.

Un autre constat fort est l'existence d'une relation positive consistante entre l'importance des inégalités sociales de performances et la dispersion des résultats en compréhension de l'écrit : la corrélation entre la variable « INEGA » et la dispersion des résultats en écrit est de + 0,67. Cela signifie que les pays dont les performances des élèves sont hétérogènes sont aussi ceux où les inégalités sociales de performances sont importantes ; réciproquement, les pays dont les élèves ont des performances homogènes sont aussi ceux où les inégalités sociales ont le moins la possibilité de s'exprimer. Par ailleurs, il existe une relation modérée entre le poids des élèves les plus faibles (élèves dont le score en écrit est inférieur à 335) et l'importance des inégalités sociales ( $r = +0,27$ , significatif au seuil de 10%) : les pays où ces dernières sont les plus marquées ont aussi tendance à être ceux dans lesquels le pourcentage d'élèves faibles est le plus élevé ; réciproquement, quand les pays parviennent à réduire le pourcentage d'élèves très faibles, les inégalités sociales de réussite ont tendance à être moins marquées

### ***II.3 Contextes nationaux et inégalités sociales***

Pour tenter d'expliquer les différences dans l'ampleur des inégalités sociales de performances d'un pays à l'autre, avant même de s'intéresser au rôle spécifique éventuel des caractéristiques des systèmes éducatifs des pays, il convient de tenter d'évaluer l'impact du contexte global des pays

Il existe bien sûr une relation statistique entre la performance moyenne des élèves et la richesse économique du pays telle que mesurée par le P.I.B. par tête : plus le pays est riche, plus les scores sont élevés. Cela dit, pour un même niveau de richesse, les pays s'avèrent inégaux en termes de performances moyennes : ainsi les pays de l'Europe de l'Est, dont le

P.I.B. par tête est assez proche de celui des pays d'Amérique du Sud ont des élèves nettement plus performants que ces derniers ; interviennent donc, à côté de la richesse stricto sensu, des facteurs historiques et culturels. Notons par ailleurs que le P.I.B. est également lié à la dispersion des résultats entre les élèves : les performances des élèves s'avèrent d'autant plus hétérogènes que les pays sont riches, la relation étant toutefois modérée.

En ce qui concerne l'ampleur des inégalités sociales, il n'y a par contre aucune liaison significative avec la richesse du pays (la corrélation est nulle). De même, il n'existe aucune relation entre l'importance des inégalités sociales de réussite scolaire et les inégalités de revenus dans la population d'adultes (mesurées par l'indice de Gini). Cela dit, certains pays sont effectivement à la fois peu inégalitaires eu égard à l'indicateur de Gini et eu égard aux performances des élèves (Japon, Corée, Finlande). Et l'on ne trouve aucun pays qui serait inégalitaire en termes de revenus et où il y aurait peu d'inégalités sociales de performances scolaires ; les pays les plus inégalitaires au niveau des revenus (pays d'Amérique latine) étant des pays où l'impact de l'origine sociale sur les performances est relativement moyen.

Ce qui est le plus frappant, c'est la variété de l'influence de l'origine sociale pour des pays qui se situent dans un intervalle très réduit du point de vue de l'indice de Gini. Ce résultat indique que certains systèmes parviennent sans doute à compenser, mieux que d'autres, les inégalités qui existent dans la société. Notons enfin que tout comme le PIB n'apparaissait pas corrélé avec l'ampleur des inégalités sociales, les dépenses ou la part du budget allouées au système éducatif ne le sont pas davantage.

#### ***II.4 Sélectivité et différenciation au sein des systèmes éducatifs et inégalités***

Il existe une littérature abondante sur l'impact de la structure du système et du calendrier de la différenciation des itinéraires scolaires sur l'ampleur des inégalités entre élèves et notamment les inégalités sociales de carrières scolaires (Duru-Bellat, 2002, 2003a). Ces dernières sont fortement affectées par le degré d'ouverture des systèmes : les inégalités sont plus fortes dans les pays qui sélectionnent précocement et, réciproquement, une voie sûre pour démocratiser les carrières scolaires consiste à atténuer ou différer la sélection. On sait également que l'existence de paliers d'orientation précoce, l'absence de filières, le poids des familles dans les processus d'orientation (etc.) sont des paramètres institutionnels qui affectent les inégalités sociales. Récemment, certains chercheurs ont mobilisé dans cette perspective les études internationales de l'IEA (notamment l'enquête TIMSS menée en 1995 sur les connaissances en mathématiques et sciences des élèves de 13-14 ans) ; ils montrent, entre autres résultats (Vandenberghe, 2001 ; Zachary, 2001), que pour un niveau donné de performance, l'ampleur des inégalités sociales peut varier du simple au double selon les pays, et que, parmi les facteurs susceptibles d'expliquer cette variété, les pays qui privilégient un tronc commun long

semblent mieux placés que ceux qui ont opté pour une différenciation précoce et marquée entre filières de formation. D'autres caractéristiques institutionnelles, comme le fait de prévoir des redoublements ou l'existence d'une ségrégation entre écoles, sont également des facteurs d'inégalité sociale de résultats.

Nos résultats convergent avec cette littérature. Que l'on prenne en compte l'existence de filières à 14 ans ou la longueur du tronc commun, il se confirme qu'une sélection/orientation précoce, non seulement n'améliore pas les performances des élèves, mais de plus ne permet pas de dégager une élite plus fournie (le contrôle par le P.I.B. dans le modèle ne contredit pas ces conclusions). Plus nettes sont les relations entre ces indicateurs de différenciation des systèmes et l'ampleur des inégalités sociales. Ces dernières s'avèrent plus élevées dans les pays à filières et cette relation est statistiquement significative ; l'existence de filières à 14 ans, à elle seule, explique 12% des variations de l'indicateur d'inégalités sociales. La relative faiblesse de cette relation s'explique par le fait que parmi les pays qui ont des filières avant l'âge de 14 ans, on trouve simultanément l'Allemagne, qui présente un fort degré d'inégalités sociales de performances et des pays comme l'Espagne et l'Italie où celles-ci sont plus modérées et comparables à des pays sans filières tels que la Norvège et la Suède. Cette relation mérite d'autant plus d'être soulignée que, comme nous l'avons vu, l'existence de filières précoces n'est pas liée significativement au niveau moyen des performances. De manière convergente, si l'on prend en compte la longueur du tronc commun, on relève une corrélation significative avec l'ampleur des inégalités sociales de réussite : plus le tronc commun est long, moins les inégalités sont importantes et inversement, celles-ci tendent à être plus marquées quand le tronc commun est bref ( $r = -0,40$ ).

Selon les systèmes, les élèves sont plus ou moins regroupés dans les mêmes établissements selon leurs caractéristiques scolaires ou sociales (ceci peut notamment résulter de la possibilité laissée aux parents de choisir l'école de leur enfant) ; on peut appréhender la ségrégation scolaire entre établissements par la part de la variance des scores des élèves expliquée par l'établissement fréquenté. On constate que les résultats sont d'autant plus faibles que la variance inter établissements est plus élevée ( $r = -0,35$ ) ; ils sont également d'autant plus dispersés ( $r = +0,39$ ), mais ces deux relations sont d'intensité modérée. Il s'avère aussi que les inégalités sociales de réussite entre élèves ont tendance à être plus marquées dans les pays où il existe une forte ségrégation entre établissements ( $r = +0,66$ ) ; c'est le cas dans les pays germaniques ou d'Europe centrale qui s'opposent ainsi aux pays du Nord et aux pays asiatiques ou encore à des pays comme l'Espagne ou le Canada dont les établissements sont assez homogènes et les inégalités sociales relativement faibles. Il s'avère donc que là où, de fait ou de droit, les établissements accueillent des publics clairement distincts, cette ségrégation scolaire revêt une dimension sociale et les acquis des élèves sont

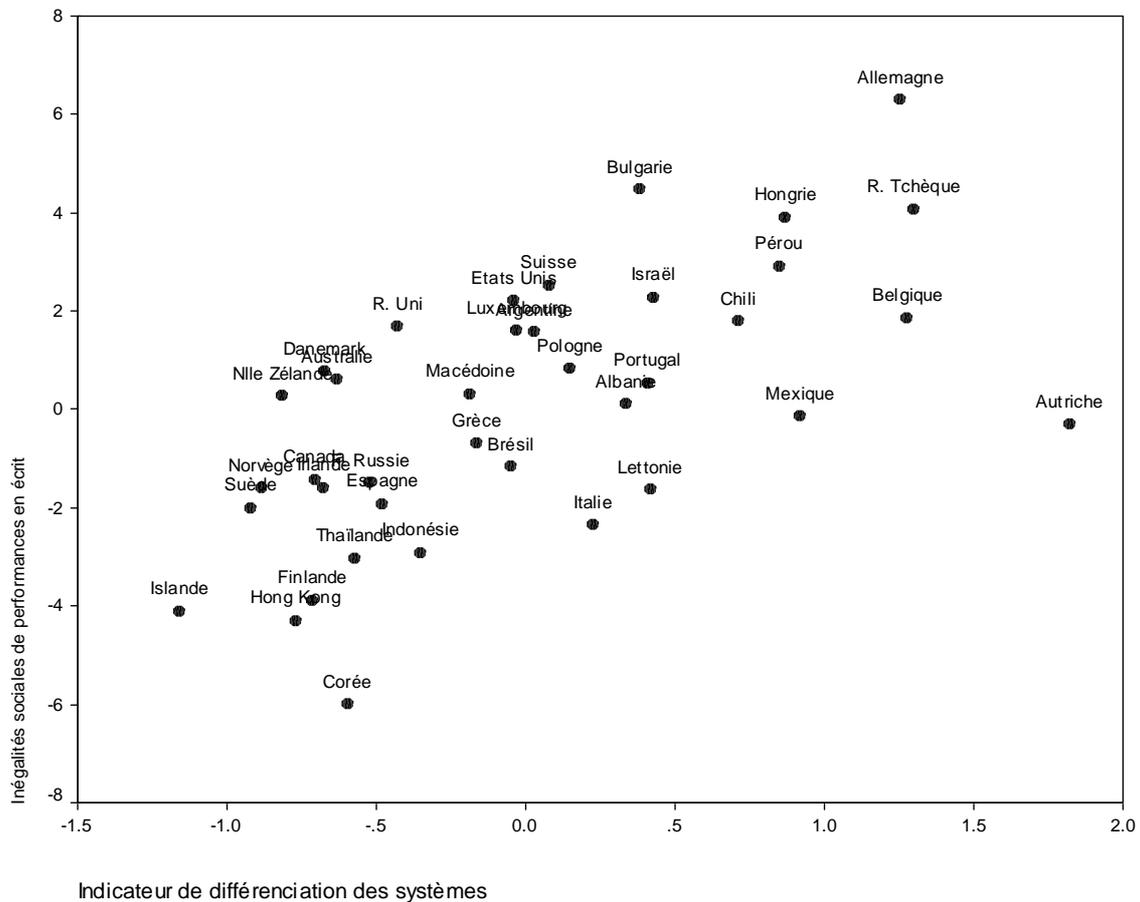
davantage marqués par leur origine sociale que dans les pays où davantage d'hétérogénéité prévaut au sein des écoles.

Qu'il s'agisse de filières ou d'établissements distincts, ou encore de l'existence du redoublement (puisque le redoublement s'efforce d'homogénéiser le niveau des classes, de même que l'existence de filières précoces et d'établissements de niveaux différents), il y a là des manières de regrouper ces élèves qui permettent d'opposer, comme le suggèrent Crahay et Delhaxhe (2003), les systèmes selon une « culture de l'intégration » et une « culture de la différenciation ». Ceci a-t-il un effet sur leurs performances et l'ampleur des inégalités sociales ? Pour instruire cette question, nous avons construit une variable agrégée qui prend en compte à la fois l'importance du redoublement en primaire, la brièveté du tronc commun et l'ampleur de la ségrégation scolaire entre établissements. Plus cette variable revêt des valeurs élevées, plus les élèves sont, très tôt, placés dans un environnement pédagogique homogène; à l'inverse plus cette variable est faible plus les systèmes maintiennent une organisation caractérisée par l'hétérogénéité. On observe alors qu'il existe une corrélation négative, significative bien que d'intensité modérée, entre cette variable exprimant la différenciation des systèmes et le score en compréhension de l'écrit ( $r = - 0.33$ ).

Contrairement à ce qui pouvait être attendu, on n'observe pas de corrélation significative entre l'indicateur de différenciation des systèmes et la dispersion des performances. Ce qui est clair, c'est que la différenciation du système ne permet pas de dégager une « élite » scolaire (appréhendée par le pourcentage des élèves les meilleurs, ceux dont le score est supérieur à 625), au contraire puisque ce sont les systèmes les moins différenciés qui s'avèrent les mieux classés à cet égard ( $r=-0,35$ ). Ces deux relations entre le caractère différenciateur des systèmes et à la fois le score moyen et l'importance de « l'élite » se maintiennent quand on tient compte de la richesse économique des pays et de la couverture scolaire à l'âge de 15 ans).

La relation entre différenciation des systèmes et ampleur des inégalités sociales est encore plus intense (graphique 12) : plus les systèmes scolaires maintiennent un niveau élevé d'hétérogénéité (peu de redoublements, tronc commun long, établissements peu différenciés), moins les inégalités sociales de réussite entre élèves sont importantes ( $r = +0,63$ ). Les pays dont les systèmes cultivent la différenciation ont donc des élèves plutôt moins performants. A nouveau, les différences entre pays se structurent par grandes régions. S'opposent des types de pays bien distincts : d'un côté les pays germaniques, d'Europe de l'Est et du Sud (les plus pauvres) et ceux d'Amérique du Sud, aux systèmes éducatifs à la fois différenciateurs et socialement inégalitaires ; dans une situation opposée, on trouve les pays asiatiques et les pays d'Europe du Nord ainsi que certains pays comme l'Islande.

Le fait que la différenciation des systèmes soit corrélée davantage avec l'ampleur des inégalités sociales qu'avec le niveau moyen ou la dispersion des résultats souligne, combien dans les systèmes éducatifs, l'accès à l'école est un enjeu, pour des stratégies familiales actives, qui s'emparent de toute différence (choix de filière, choix d'établissement...) pour en faire un critère de distinction.



Graphique 12 : Inégalités sociales de performances et différenciation des systèmes éducatifs

Enfin, à titre de synthèse, nous avons estimé un modèle intégrant uniquement les deux variables dispersion des scores et indicateur de différenciation : il explique 67% des différences entre pays en matière d'inégalités sociales de performances. En d'autres termes, si les pays laissent exprimer plus ou moins d'inégalités sociales de performances entre leurs élèves, c'est avant tout en raison de deux facteurs d'importance voisine, qui sont assez peu corrélés, et viennent donc s'ajouter : l'existence d'inégalités inter-individuelles en leur sein d'une part et le caractère plus ou moins différencié de leur système éducatif d'autre part.

Tableau 11 : Modèle expliquant les inégalités sociales par la différenciation des systèmes et la dispersion des résultats des élèves en écrit

Variabes	Coefficient	Significativité
Indicateur de différenciation	+0,72	***
Dispersion des résultats à l'écrit	+0,061	***
Constante	-5,85	***
R <sup>2</sup>	0,67	

\*\*\* : significatif au seuil de 1%, \*\* : significatif au seuil de 5%, \* : significatif au seuil de 10%

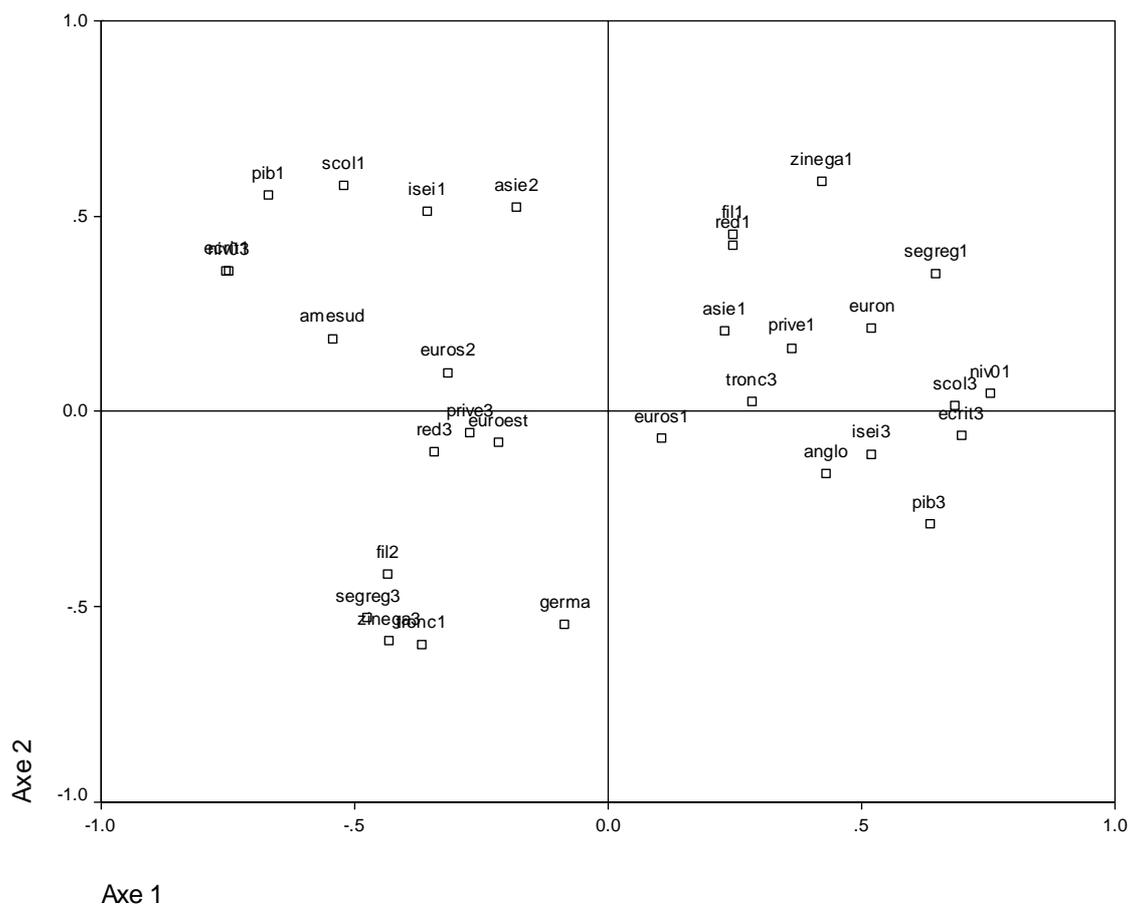
Pour fournir une vision d'ensemble des différences entre pays, nous avons réalisé une analyse en composantes principales qui intègre les caractéristiques institutionnelles majeures des systèmes éducatifs, l'ampleur des inégalités sociales, le niveau de richesse des pays et les performances en compréhension de l'écrit<sup>27</sup>. Le graphique ci-après fait apparaître un premier axe horizontal qui explique 23% de la variance du nuage de points et qui oppose globalement les pays où la scolarisation est élevée, avec de bonnes performances (sur la droite du graphique), et qui sont plutôt des pays riches, aux pays qui sont dans une situation inverse (à gauche du graphique), comme l'expriment les variables *pib1*, *isei1*, *scol1*, *écrit1*, *niv03*. Il se confirme donc que ce qui structure avant tout les différences entre pays, ce sont le niveau de la scolarisation et des performances, et le niveau de richesse. L'axe vertical explique 13% de la variance et représente quant à lui le degré d'homogénéité structurelle du système et l'ampleur des inégalités (*zinega1*, *red1*, *fill*, *segreg1*). En haut du graphique se trouvent les variables représentant les pays les plus égalitaires et aux structures les moins différenciées, et inversement pour la partie basse du graphique.

Nous avons également projeté les groupes de pays par grandes régions du monde<sup>28</sup>. On retrouve ainsi l'opposition entre des pays plutôt performants et peu inégalitaires (dans le cadran haut-droit du graphique), les pays de l'Europe du Nord et les pays riches d'Asie d'une part, et d'autre part, les pays germanique (dans le cadran bas-gauche) dotés de caractéristiques opposées. Dans le cadran haut-gauche figurent des pays moins performants mais aussi

<sup>27</sup> Les variables participant à l'analyse factorielle ont pour la plupart d'entre elles été construites sur la base d'un découpage de la variable initiale en trois tiers, allant de 1 (valeurs moins élevées) à 3 (valeurs les plus élevées). Par exemple « *niv01* » correspond à un faible pourcentage d'élèves en-dessous du niveau 1 de compétences en écrit ; de même « *Zinega1* » représente les pays les plus égalitaires. La variable exprimant l'existence de filières est dichotomique : « *fill* » : pas de filières.

<sup>28</sup> On distingue dans ces regroupements : les pays asiatiques du Nord (Japon, Corée) notés Asie 1 sur le graphique, les pays les plus pauvres d'Asie (notés Asie 2 sur le graphique) c'est-à-dire Thaïlande et Indonésie, les pays anglo-saxons (Royaume-Uni, Nouvelle Zélande, Australie, Etats-Unis...), les pays d'Europe du Nord (Finlande, Suède...), les pays de l'Europe de l'Est (Hongrie, Fédération de Russie, Pologne...), les pays germaniques (Allemagne, Luxembourg, Autriche...), les pays les plus riches de l'Europe du Sud (notés 1 sur le graphique), soit la France, l'Italie... , les pays les moins riches de l'Europe du Sud (Bulgarie, Macédoine..) et enfin les pays d'Amérique latine (Brésil, Argentine, Mexique, Pérou).

relativement peu inégalitaires (pays pauvres d'Asie notamment), que l'on peut opposer aux pays anglo-saxons, plus bas et à droite sur le graphique, c'est à dire plutôt performants mais plus inégalitaires. Si cette analyse factorielle ne donne évidemment pas des résultats foncièrement différents que ce que nous avons établi ultérieurement, ils confortent le sentiment d'une réelle cohérence des relations entre les caractéristiques des systèmes et le niveau de richesse des pays d'une part, les performances des élèves et les inégalités sociales qui marquent ces performances, d'autre part.



Graphique 13 : Analyse en composantes principales des caractéristiques des pays

## Conclusion

Même s'il est sans doute excessif de tirer de ces constats des enseignements politiques, ne serait-ce que parce qu'il ne s'agit que de corrélations, insuffisantes pour établir des relations de cause à effet, deux points méritent néanmoins d'être soulignés. Tout d'abord, le fait que les acquis des élèves et les inégalités qui les marquent (inégalités entre élèves, inégalités sociales) varient sensiblement selon les pays montre qu'ils relèvent de l'action politique ; sachant que le poids important de la richesse des pays pointe aussi l'effet du même coup limité de politiques qui se limiteraient à la sphère éducative. D'autre part, il apparaît assez clairement que le mode d'organisation globale des systèmes est corrélé avec leur degré d'équité : toute limitation de la scolarisation ou mise à part précoces de certains élèves, tout groupement par niveau ou filières distinctes (dans le cadre de la scolarité obligatoire), ou encore des phénomènes de ségrégation entre établissements (organisant une « offre » de qualité inégale), tendent à accroître l'inégalité sociale des performances sans améliorer pour autant le niveau moyen ou même le niveau de l'élite. Ceci alors même que seules les compétences scolaires sont prises en compte. Pour qui s'intéresse aussi à l'intégration civique et sociale des jeunes d'un pays, ce constat est sans doute encore plus valable.

Certes, les données PISA, qui ne concernent pas les pays les plus pauvres de la planète et qui n'incluent pas d'observations du fonctionnement précis des systèmes, ne permettent pas d'explorer toutes les voies d'action concevables. Mais elles montrent en tout cas que l'école peut quelque chose, face aux inégalités entre élèves, et ceci mérite d'être rappelé et exploré plus encore par de nouvelles recherches. Quant à la situation de la France, elle apparaît dans cette dimension sociologique relative aux inégalités sociales de réussite, dans une position encore moyenne dans le contexte international.

### *Chapitre 3*

*Le contexte de la classe : constitution et effets sur les élèves*

En France, les travaux sociologiques spécifiquement consacrés à l'appréhension des facteurs contextuels sont très rares, pour des raisons de nature diverse. L'une est la prégnance, chez les sociologues de l'éducation, des explications structurelles jusqu'au début des années 80 : avec la thèse de la reproduction, les structures sont perçues comme déterminantes quant au devenir scolaire des individus, et du même coup, les processus qui lient ces facteurs structurels aux comportements individuels ne sont guère l'objet d'une attention spécifique. Une autre explication est la relative faiblesse méthodologique des sociologues français, alors même que la simple appréhension de phénomènes contextuels exige la mise en œuvre de méthodologies reposant sur des analyses multivariées, pour démêler effets de composition et effets proprement contextuels. Enfin, l'absence de grosses équipes dotées de moyens importants pour engager des enquêtes nécessairement lourdes est sans doute une des raisons de la faiblesse française. Pourtant, dans le même temps, le thème de la ségrégation sociale connaît une vogue croissante (Van Zanten, 1996, 2001) : dans le cadre d'approches dites qualitatives très répandues dans la sociologie de l'école, les descriptions ethnographiques d'environnements scolaires contrastés s'accumulent et suggèrent, par le terme même de ségrégation, que la concentration d'élèves ayant tel ou tel profil a des effets spécifiques, en particulier que la concentration d'élèves faibles et/ou de milieu social défavorisé a des effets délétères. Mais de fait, ce type d'approche ne permet pas d'établir que ces effets existent, a fortiori qu'ils sont forts ou faibles ou qu'ils concernent plus les apprentissages que les attitudes.

Depuis la fin des années 80, l'Irédu-CNRS a été l'équipe la plus impliquée, en France, dans l'analyse quantitative des effets de contexte, notamment du fait de collaborations possibles, en son sein, entre sociologues et économistes, et aussi de son ouverture aux problématiques anglo-saxonnes. C'est ainsi qu'ont été mises en évidence les notions d' « effet maître » et d' « effet établissement », aujourd'hui triviales (Duru-Bellat et Mingat, 1988 ; Bressoux, 1994), puis l'influence spécifique de l'hétérogénéité/homogénéité scolaire de la classe fréquentée au niveau du collège (Duru-Bellat et Mingat, 1997).

Une question préalable à toute analyse des effets de telle ou telle caractéristique du contexte porte sur le « niveau de contexte » le plus pertinent, la classe ou l'établissement. Cette question du niveau de contexte le plus pertinent (établissement, classe, éventuellement groupe d'élèves) est une question à la fois théorique et empirique, comme le soulignent Bressoux et al. (1997). En effet, la distinction des différents niveaux constitue toujours « une hiérarchie construite, qui ne va donc pas de soi et qui repose toujours sur l'hypothèse de la pertinence du découpage effectué quant au phénomène étudié ». De fait, on ne dispose guère de théories intégrant dans un modèle unique les différents niveaux de contexte, emboîtés, où évolue l'élève, et permettant de comprendre que les apprentissages ou les attitudes d'un élève puissent varier selon telle ou telle caractéristique de sa classe ou de son établissement.

Certains travaux empiriques soulignent par ailleurs que, selon la cible de l'analyse, le « niveau de contexte » le plus pertinent peut varier (Duru-Bellat et Mingat, 1988) : en collège, si les progressions restent liées avant tout au niveau de la classe (Fraïne, Van Damme, Van Landeghem, Opdenakker, Onghena, 2003), l'établissement fréquenté aurait davantage d'influence sur les processus d'évaluation et d'orientation. Les résultats disponibles sur le niveau primaire ne sont pas toujours convergents, ne serait-ce que parce que les écoles sont sans doute moins différentes entre elles que ne le sont les établissements du secondaire. A ce niveau (Bressoux, 1995), pour ce qui concerne les progressions des élèves, le niveau de la classe est plus pertinent que le niveau de l'école : le fait d'appartenir à telle ou telle école n'explique que 3 à 4% de la variance des acquis des élèves, contre 15% environ pour l'appartenance à telle classe (sur la base des travaux existants, estimant pour leur immense majorité des modèles mono-niveau).

Sur les estimations chiffrées de l'importance des différents niveaux, il faut rappeler ce point aussi fondamental qu'évident : pour qu'un effet de tel ou tel niveau de contexte soit visible, il faut qu'il soit caractérisé par une certaine variété ; si tous les établissements sont strictement identiques sous l'angle d'un certain nombre de caractéristiques, on ne pourra déceler d'effet établissement, même si ces caractéristiques sont fondamentales eu égard au phénomène étudié, et l'impact des autres facteurs (ou niveaux de contexte) introduits dans les analyses, les caractéristiques des élèves ou des classes par exemple, sera du même coup sur-estimé. Un second préalable, pour analyser l'influence éventuelle des caractéristiques du contexte sur les acquisitions des élèves, est d'appréhender ces dernières de manière comparable d'un contexte à l'autre. Or il faut rappeler qu'il est nécessaire de disposer, en la matière, d'épreuves standardisées de connaissances, vu l'existence de biais contextuels de notation. Des travaux antérieurs (Duru-Bellat, Mingat, 1993), toujours confirmés depuis (Bressoux, 2003), montrent en effet qu'il existe une relation entre le niveau moyen des notes et certaines caractéristiques du contexte, avec comme tendance dominante une plus grande indulgence dans les contextes populaire (c'est à dire des notes en moyenne plus élevées pour des élèves ayant par ailleurs un niveau de connaissance identique). Cette faible comparabilité des notes scolaires d'un contexte à l'autre conduit à se fonder, autant que faire se peut sur des épreuves standardisées de connaissances scolaires. Ces épreuves, en général construites spécifiquement à cette fin, ne sont toutefois jamais dépourvues d'un certain arbitraire et rarement parfaites d'un point de vue psychométrique. Dans certains cas, c'est la fidélité aux programmes scolaires qui sera mis au premier plan, dans d'autres cas c'est davantage l'allure de la distribution des résultats des élèves qui prime ; la principale vertu de ce type d'épreuves est leur comparabilité d'un site à l'autre rendue notamment possible par une homogénéisation des conditions de passation et de correction. La nature des épreuves est une question capitale, parfois un peu négligée dans les approches sociologiques, alors même que les évaluations des effets de contexte semblent très sensibles à la qualité métrologique des indicateurs retenus (Opdenakker, Van Damme, 2002).

Au niveau strictement méthodologique, la modélisation multivariée constitue en sciences sociales un substitut à l'expérimentation. Dès lors que l'on travaille sur des données nombreuses et diversifiées, les analyses vont permettre de raisonner « toutes choses égales par ailleurs », et d'évaluer l'impact d'une variable net de toutes les relations éventuelles qu'elle entretient avec les autres variables dans le monde réel. Tout comme l'approche expérimentale, la modélisation permet d'effectuer des inférences statistiques sur la base des échantillons mobilisés bien que ceux-ci soient de nature différente selon l'approche retenue.

Dans la recherche en Education, la modélisation multivariée revêt la forme de modèles de progressions ou plus largement de transformations entre un état initial et un état final (acquisitions scolaires, attitudes...), intégrant les principales caractéristiques individuelles dont la littérature montre qu'elles affectent les apprentissages ; on examine ainsi, par exemple, si la nationalité de l'élève est associée significativement à ses progressions, à autres caractéristiques personnelles identiques, notamment le milieu social d'origine. On peut ensuite intégrer dans ces modèles le contexte fréquenté dont l'impact sera lui aussi évalué « toutes choses égales par ailleurs », c'est-à-dire pour des élèves de caractéristiques individuelles données ; on évaluera ainsi, par exemple, si le fait de fréquenter un établissement à tonalité populaire affecte les progressions des élèves, quelles que soient leurs caractéristiques personnelles et notamment leur appartenance sociale. Cela dit, cette clause du « toutes choses égales par ailleurs » se limite aux variables introduites dans l'analyse et il peut toujours exister des variables inobservées susceptibles de biaiser les résultats. Cela vient notamment de la « non randomisation » des élèves dans les différents contextes, problème auquel peuvent échapper les psychologues grâce à l'expérimentation. On notera qu'un courant récent de l'économie dite « économie expérimentale » s'efforce de contourner cette difficulté que les économistes désignent sous le terme de « problèmes d'endogénéité » en mobilisant des « expérimentations naturelles ». Par exemple, on évaluera l'impact de la taille des classes à partir de situations d'ouvertures et de fermetures des classes dans les écoles qui entraînent des variations significatives et brusques des tailles de classe sans lien avec les caractéristiques des élèves et des écoles mais uniquement le fait de l'application de seuils en termes d'effectifs décidés par l'administration (Piketty, 2004).

Il est clair que si l'on pouvait disposer d'un échantillon d'écoles aussi diversifiées que possible et d'y allouer aléatoirement les élèves, la dispersion de leurs apprentissages révélerait l'efficacité différentielle des écoles et les variables de processus qui lui sont attachées. Evidemment, la prise en compte du niveau scolaire initial des élèves limite quelque peu le jeu de ces facteurs inobservés, dans la mesure où ce niveau est vraisemblablement corrélé avec un certain nombre de ces facteurs inobservés (attitudes ou comportement des élèves notamment) ; mais on ne peut jamais être sûr que si impact du contexte il y a (tel effet de la

fréquentation d'une classe homogène faible par exemple), cela s'explique par certaines caractéristiques non observées des élèves regroupés dans ces contextes. C'est le prix à payer quand on travaille en milieu dit « naturel », quelle que soit la sophistication des méthodes.

Le dispositif empirique requis pour conduire une analyse rigoureuse des effets de contexte est nécessairement lourd. Il faut en effet disposer d'une mesure des progressions des élèves (avec des mesures en début et en fin d'année scolaire), pour ensuite estimer, au-delà de l'influence des variables individuelles, d'éventuels effets établissement et effets classe. Ceci exige un échantillon quantitativement important, puisque, nous l'avons évoqué, les contextes observés doivent être assez variés pour que leur influence éventuelle puisse être évaluée au mieux. Ce n'est qu'ensuite que l'on pourra estimer le poids des caractéristiques du contexte que l'on souhaite étudier. Enfin, si des effets contextuels spécifiques sont décelés, il conviendra d'en rechercher les processus générateurs, ce qui suppose des observations qualitatives sur le déroulement de l'enseignement en classe ou encore l'observation des interactions entre élèves. On travaille donc nécessairement par étapes emboîtées, depuis l'appréhension d'un effet du contexte établissement et/ou classe comme une « boîte noire » jusqu'à une exploration des facettes de ces contextes que l'on suppose exercer une influence.

Nous nous proposons dans ce chapitre d'étudier de façon relativement approfondie deux dimensions du contexte d'enseignement qui apportent des éléments empiriques nouveaux dans le paysage de la recherche française sur cette problématique de l'influence de l'environnement scolaire sur les acquisitions des élèves. Ces deux dimensions concernent la composition des groupes d'élèves, à savoir les modes de groupement pédagogiques des élèves et la mixité sociale. Auparavant, il n'est pas inutile de se pencher sur une question très peu abordée dans la littérature française et qui se situe en amont de la mesure de l'efficacité du contexte scolaire, il s'agit en effet d'examiner comment ce contexte peut partiellement se construire de part l'intervention des enseignants.

## **I La fabrication du contexte : procédures de constitution des classes à l'école élémentaire**

Le courant du « school effectiveness », solidement ancré dans la culture de la recherche en éducation dans les pays anglo-saxons, a pour objectif d'identifier ce qu'est une école efficace, en mettant en perspective les caractéristiques des établissements et les apprentissages de leurs élèves. Les résultats de ces travaux sont maintenant bien connus (Duru-Bellat, Van Zanten, 1999) qui indiquent notamment qu'un environnement incitant au travail, des attentes élevées et positives vis-vis des élèves, des évaluations fréquentes, un climat discipliné sont autant d'éléments qui concourent aux progrès des élèves. Parmi ces constats, la notion de climat a donné lieu à de nombreux développements et il semble acquis que les caractéristiques de l'équipe pédagogique (stabilité, mobilisation, motivation, coopération...) et le type de "leadership" exercé par le chef d'établissement en sont des éléments importants (Meuret, 2003). Ces recherches s'appliquent également à l'enseignement primaire dans les pays anglo-saxons puisque les directeurs y disposent d'un statut hiérarchique. Elles montrent alors que la façon dont le directeur remplit sa fonction "managériale" au sein de l'équipe a un rôle crucial sur la vie de l'école et ses résultats (Bossert, Dwyer, Rowan, Lee, 1982), notamment parce que son pouvoir de décision concerne à la fois le recrutement des enseignants et leur assignation dans une classe dans laquelle ils exercent, contrairement au niveau secondaire, à temps complet (Goldhammer, Becker, 1971; Warren, 1975).

En France, les recherches sur cette thématique sont relativement récentes et concernent pour l'essentiel les collèges. Il a en effet fallu attendre que l'établissement soit reconnu comme une réalité organisationnelle et politique (Demailly, 2000) pour que se développe un questionnement sur les effets du contexte de scolarisation. Dans cette perspective, la sociologie des organisations a apporté un éclairage particulier à l'analyse de l'identité des établissements, qui finalement se révèle être "un montage composite entre des registres d'actions qui se réfèrent à des principes de légitimité différents et souvent antagonistes" (Derouet, 1987). Le rôle du chef d'établissement y est évidemment primordial, à qui il revient de faire construire un projet éducatif aux membres de l'équipe et de favoriser l'implication et la mobilisation de tous. Quand bien même ce rôle lui est clairement et institutionnellement dévolu dans l'enseignement secondaire, sa capacité d'action dans la gestion interne de l'établissement n'est, en réalité, que partielle et souvent difficile à mettre en place (Dupriez, 2003).

Que dire alors de ce qui se passe dans les écoles primaires? Les recherches sur ce niveau d'enseignement sont très rares. En effet, dans le contexte français, le directeur d'école primaire n'a aucun pouvoir hiérarchique dans l'établissement et, hormis les décharges horaires dont il peut éventuellement bénéficier, sa charge d'enseignement est identique à celle

de ses collègues. Il n'est finalement qu'un enseignant parmi les autres et il ne peut en principe imposer une quelconque décision. Ce statut, manifestement difficile à assumer, a été l'objet ces dernières années de nombreuses revendications et de « grèves administratives », bloquant la remontée des informations réclamées par l'administration centrale. Si l'augmentation de la prime et l'élargissement des décharges de classe liées à la fonction sont clairement revendiquées par les syndicats, la question de la création d'un véritable corps des directeurs d'école reste en suspend, depuis longtemps d'ailleurs puisque dès 1987, la création d'un corps de « maître directeur » par René Monory avait engendré une levée de boucliers. Les problèmes se posent pourtant de façon similaire aux deux niveaux d'enseignement. Ecoles, collèges et lycées partagent en effet des cadres structurels porteurs de concepts communs : le premier concerne les acteurs et « suppose que les systèmes sont agis de l'intérieur par des agents, et donc susceptibles d'être transformés », le second concerne l'autonomie et postule qu'il existe des marges de manoeuvre laissées à la disposition des établissements (Cousin, 2000). Enfin, et quel que soit le niveau, l'autorité instituée n'est en aucun cas suffisante et c'est la légitimité reconnue au chef d'établissement qui importe dans le style de leadership adopté (Cousin, 1998) et dans la nature et l'efficacité des interactions au sein de l'équipe.

Cette analyse qualitative, issue d'une recherche plus large consacrée dans son ensemble aux modes de groupements des élèves (Leroy-Audouin, Suchaut, 2005), se focalise précisément sur le niveau primaire et, dans l'objectif d'identifier les dynamiques relationnelles internes, elle se base sur une décision essentielle dans le fonctionnement des écoles : celle de l'assignation des enseignants à leur classe, ou plutôt du choix des enseignants de leur classe. La façon dont se prend cette décision est d'autant plus intéressante à analyser que vont se confronter, en l'absence d'arbitrage institutionnel et hiérarchique, des intérêts individuels, souvent concurrents car identiques.

Des entretiens semi-directifs, d'une durée d'une heure trente environ, ont été conduits auprès de 18 directeurs d'écoles, interrogés entre autre, sur la façon dont se déroulaient les procédures relatives au choix des enseignants de leur classe. Leur ancienneté moyenne dans la fonction s'élève à 7 ans et le nombre d'écoles dans lesquelles cette fonction a été exercée est faible en général (moins de 2 écoles). Par conséquent, l'ancienneté sur le poste actuel est en moyenne relativement longue (5 ans). La majorité des ces directeurs/trices a la responsabilité d'une classe (seuls 3 d'entre eux sont totalement « déchargés »), dans plus de 80% des cas à hauteur de un quart de temps, ce qui correspond à environ une journée de « décharge » par semaine. Par ailleurs, des données relatives à la structure des classes dans ces écoles ont été collectées. Il s'agit d'abord du nombre total d'élèves et de leur répartition par niveau, du nombre de classes constituées, de leur effectif et de leur type (simple ou multiple), et enfin, pour chacune des classes dans chaque école, des caractéristiques des 141 enseignants concernés (âge, sexe, ancienneté générale et dans l'école, poste occupé précédemment).

### ***1.1 La priorité aux intérêts individuels***

Le point commun à toutes les écoles est qu'au-delà du relatif isolement de chacun dans sa classe, les directeurs insistent sur la notion d'équipe ; en est pour preuve le très grand nombre d'occurrences des termes « collègues », « équipe » et « consensus » qui apparaissent d'autant plus fréquemment dans les discours que les acteurs sont en présence depuis longtemps dans l'école : « *mais vous voyez, ce sont des gens qui travaillent ensemble depuis longtemps donc ça ne pose aucun problème* » (école 3) ; « *oui, c'est vraiment collégial, c'est-à-dire qu'il y a une espèce de confiance réciproque qui est établie depuis longtemps* » (école 5).

L'importance de cette histoire commune est soulignée par Gather-Thurler (1994) qui montre que les organisations subissent le poids du passé qui s'inscrit dans la mémoire des membres et les unit. Se crée ainsi un fond culturel commun aux enseignants d'une école et celui-ci, d'autant plus riche que la stabilité de l'équipe est grande, véhicule « des solutions construites, héritées et partagées en son sein » (Gather-Thurler, 1994). Cette cohésion apparente est particulièrement visible au niveau des priorités pédagogiques établies dans l'organisation de l'école. En fait, l'évitement des cours multiples reste l'argument fondateur de la constitution des classes (Leroy-Audouin, Suchaut, 2005) et il est partagé par l'ensemble des enseignants, quand bien même il pourrait engendrer des déséquilibres très grands au point de vue des effectifs d'élèves : « *alors en général, on évite un maximum les cours doubles et la politique est dans notre école : il vaut mieux une classe chargée que trop de cours doubles* » (école 3).

Une raison principale justifie aux yeux des directeurs la place première donnée à ce critère dans la constitution des classes. Elle est strictement d'ordre pédagogique et renvoie au

surcroît de travail que les cours multiples occasionnent aux enseignants, en terme de préparation d'abord : « *en tant qu'enseignant, on accepte mal les cours doubles parce que ça fait quand même une double préparation. Ça fait du jonglage, pas continu mais presque entre les deux niveaux* » (école 4) ; « *c'est beaucoup de travail pour l'enseignant, c'est ça le problème* » (école 13). La conduite de la classe au quotidien est également jugée plus difficile dans les cours multiples : « *un enseignant en cours double fait la girouette toute la journée. C'est un one-man-show, quoi !* » (école 16) ; « *physiquement, il fallait que je me mette comme ça face aux CM2 pour leur parler et je parlais aux CM1. C'était affreux, j'étais constamment insatisfait. Il m'avait manqué de temps pour ceux-là, j'avais donné trop de temps aux autres* » (école 14).

Si l'évitement des cours multiples est la règle dans la constitution des classes, il n'en reste pas moins que la France est, selon une étude réalisée dans le cadre de l'OCDE (M.E.N., 1997), l'un des pays dans lesquels, en raison de la petite taille des écoles, les classes à cours multiples sont les plus fréquentes: elles concernent environ 80% des écoles élémentaires. Le consensus très fort observé précédemment sur le rejet des cours multiples peut voler en éclat alors et conduire à l'émergence de situations potentiellement conflictuelles quand les enseignants doivent choisir leur classe, notamment si un cours multiple se présente. C'est la raison pour laquelle, un directeur souligne qu'au moment de la constitution des classes, on songe déjà aux individus : « *il n'y a pas encore trop d'enjeu personnel, quoique ça commence : où je vais me mettre ?... mais attention, là j'avais un CE2, là il va en manquer un...* » (école 5).

Si les enseignants se rassemblent massivement autour de certaines valeurs, par conséquent, une certaine rivalité les oppose aussi et l'esprit d'équipe évoqué précédemment doit respecter, voire servir, les intérêts individuels en présence. Il s'agit donc de trouver un compromis et d'apporter une réponse pratique à ce problème, par la mise en place de négociations entre des acteurs cherchant à préserver leur propre intérêt. Dans cette perspective, Maroy et Dupriez (2000) évoquent la notion de "régulation": celle-ci se comprend comme un processus social de production de règles du jeu, prenant place dans un espace daté et localisé. Cette notion accorde, conformément à d'autres travaux d'ailleurs (Reynaud, 1993; Friedberg, 1993), un poids central à l'histoire de l'équipe, qui favorise (ou freine parfois) la production, l'appropriation, voire le remodelage de ces règles. Les déclarations des directeurs révèlent ainsi qu'il existe à la fois des règles valables dans l'ensemble des écoles qui s'appliquent à tous les enseignants, et des arrangements locaux, liés à l'équipe et à l'histoire de chaque école.

## ***1.2 Le poids des usages***

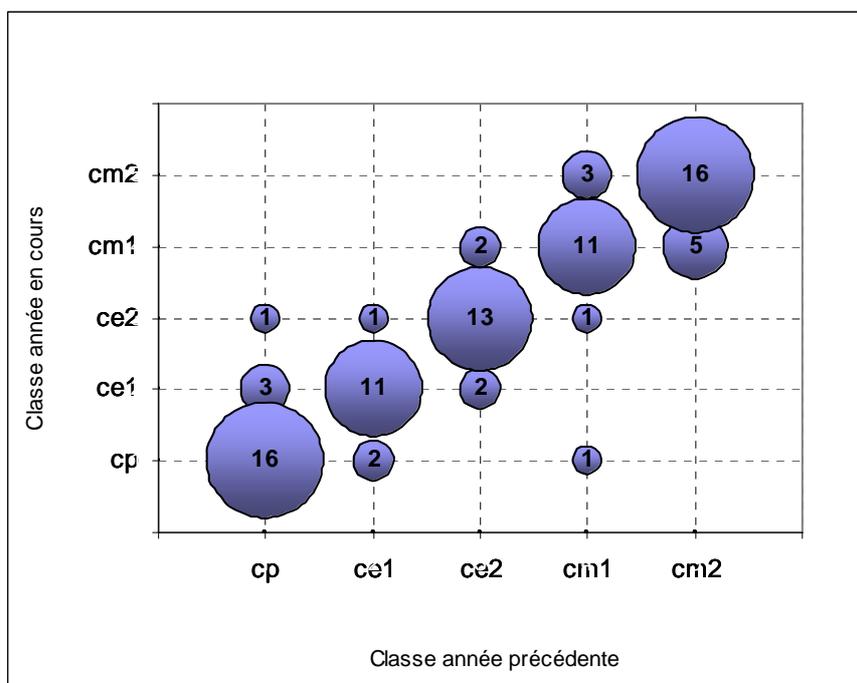
Selon le calendrier rapporté par les directeurs interrogés, après avoir décidé des classes à constituer pour la rentrée suivante, il s'agit ensuite de les attribuer aux enseignants. Ce n'est que dans une dernière étape qu'y seront affectés les élèves. C'est donc en fonction d'une classe considérée dans sa structure (type de cours et niveau) et pas dans sa composition (le groupe d'élèves) que les enseignants doivent se prononcer. Concernant le type de classe, cela a été dit précédemment, la préférence va nettement aux cours simples. Concernant le niveau d'enseignement, s'il apparaît en moyenne une certaine spécialisation des enseignants en école maternelle ou en école élémentaire, en revanche, la question des différents niveaux élémentaires est rarement abordée tant elle semble liée à ce qui semble être une sensibilité personnelle. Selon les directeurs en effet, il n'existe pas *« le sentiment que ce soit plus difficile dans un niveau que dans un autre »* (école 8) et en fait, *« c'est très variable selon les personnes »* (école 1), *« c'est chacun... chacun selon son caractère »* (école 14).

Si tous les niveaux de classe n'exercent pas le même attrait sur les enseignants, on peut imaginer alors que ces derniers tentent, dans la mesure du possible, de conserver d'une année sur l'autre « leur » classe. Les déclarations des directeurs indiquent que c'est effectivement le cas dans la majorité des écoles.

L'attachement des enseignants à un même niveau scolaire est une réalité d'autant plus forte que l'organisation n'est pas remise en cause d'une année sur l'autre, comme c'est le cas dans les écoles suivantes : *« les gens qui avaient ces classes-là cette année reprendront les mêmes »* (école 2) ; *« pour l'instant, on garde nos classes »* (école 12). Par ailleurs, cette tendance se confirme quand les enseignants sont plus âgés ou plus anciens dans le niveau : *« je crois qu'il y a aussi des règles un peu implicites : le collègue qui a des CM2 depuis 15 ans, on chamboule pas tout, on va pas lui donner un CP »* (école 5) ; *« c'est des gens qui sont là depuis un certain temps, qui ont leurs habitudes de cours, de niveau »* (école 18).

Cette « spécialisation », qui correspond effectivement aux préférences des enseignants pour une classe d'âge particulière, renvoie donc également à un certain confort dans l'exercice du métier : *« il y a d'autres collègues qui vont me dire : « attends, tu ne vas pas me donner un cours double alors que je pars en retraite l'année prochaine ? » C'est rarement des arguments pédagogiques mais je me mets à leur place aussi... c'est pas marrant »* (école 2). Cette spécialisation n'est évoquée de façon positive et constructive que par deux directeurs qui y voient, au contraire, l'opportunité d'un approfondissement : *« le collègue qui fait du CP, moi je souhaite qu'il reste sur le cycle 2 puisque c'est quelqu'un qui s'est énormément investi sur l'apprentissage de la lecture, qui fait un excellent travail, d'abord de pédagogue et dont on a énormément besoin au niveau de la formation IUFM »* (école 1).

L'analyse des données concernant les classes et les enseignants de ces 18 écoles confirme cette tendance et le graphique suivant rend compte des mouvements d'une année sur l'autre pour tous les niveaux d'enseignement.



Graphique 14 : Répartition des différentes configurations de classes

Les enseignants qui étaient en cours simple et qui le sont restés sont au nombre de 89. Parmi eux, les trois-quarts (75,3%) ont strictement conservé leur niveau de classe, la « spécialisation » étant particulièrement marquée sur la classe de CP. On observe par ailleurs que les mouvements se réalisent en général au sein d'un même cycle d'enseignement : si le CE2, qui appartient pourtant au cycle 3, a une position très intermédiaire à cet égard, les « glissements » des enseignants sur le niveau précédent ou suivant sont extrêmement fréquents entre le CP et le CE1, et le CM1 et le CM2. En cas de difficulté ou de désaccord dans l'équipe, une règle univoque et commune à toutes les écoles dont le directeur a été interrogé, s'impose avec autorité : c'est celle de l'ancienneté dans l'école, qui détermine l'ordre dans lequel les enseignants vont s'exprimer : « la coutume veut que ce soit les plus anciens dans l'école qui parlent en premier » (école 6) et « le dernier arrivé prend ce qu'on lui donne » (école 1).

Les directeurs évoquent cette pratique en termes de « coutume » ou de « tradition » qui n'est que rarement remise en cause, y compris par les Inspecteurs de l'Education Nationale. Dans

toutes les écoles, cette démarche est la même, qui accorde aux plus anciens dans l'école le droit de choisir en premier alors que les nouveaux arrivants, même ceux dont l'ancienneté générale serait plus élevée, se doivent d'accepter ce qui « reste ». Une exception à cette règle est parfois admise et le premier enseignant par l'ancienneté dans l'école peut être devancé par le directeur : « *en théorie, c'est le directeur qui choisit, ensuite les adjoints se répartissent les classes suivant leur ancienneté dans l'école* » (école 15) ; « *la tradition veut ça... la personne la plus ancienne choisit. En fait, le directeur choisit en premier puisque c'est par ordre de...* » (école 7).

Certains directeurs utilisent cette prérogative tandis que d'autres la refusent, témoignant en cela de la façon dont ils gèrent la question de leur autorité au sein de l'équipe : « *le directeur, par sa fonction, a le choix... mais je n'aime pas imposer* » (école 17) ; « *ça veut dire que je n'utilise pas ma prérogative, normalement si je voulais, je pourrais choisir ma classe en premier et puis bon... comme je voudrais que justement on n'en arrive pas à se répartir les classes par ancienneté...* » (école 8). Par ailleurs, certains contournements de cette règle sont tolérés dans des situations particulières, et notamment en ce qui concerne la classe de CP qui ne doit pas être attribuée à un enseignant débutant ou un enseignant à temps partiel. On peut alors déroger à la règle de l'ancienneté dans l'école, pour éviter que le CP soit la classe qui « reste » à un nouvel arrivant sortant de formation

L'adage selon lequel « *les derniers arrivés sont les derniers servis* » (école 17) se vérifie dans cet échantillon, ainsi que l'illustre le tableau 12 suivant, qui met en relation l'ancienneté des enseignants et le type de classe auquel ils sont affectés. Globalement et conformément aux déclarations des directeurs, il n'existe pas de différence selon l'ancienneté générale des enseignants puisque la règle concerne le dernier arrivé dans l'école. Celui-ci peut avoir un certain nombre d'années d'ancienneté dans l'école si cette dernière connaît peu de mouvements et il sera en général considéré, jusqu'à nouvel ordre, comme le moins ancien et par conséquent, susceptible de conserver une classe qu'il ne souhaitait pas : « *et bien voilà, donc il y en a qui peuvent rester avec des cours doubles pendant des années* » (école 1). Un indicateur a donc été créé, qui rend compte de l'identification dans chaque école de cet enseignant particulier<sup>29</sup>.

---

<sup>29</sup> Dans quelques écoles, deux enseignants arrivés la même année sont considérés comme les derniers arrivés.

Tableau 12 : Type de classe occupé selon que l'enseignant est le dernier arrivé ou pas dans l'école<sup>30</sup>

Enseignants	Cours simple		Cours multiple		Total
les plus anciens	95	86,4%	15	13,6%	110 (100)
les derniers arrivés	19	70,4%	8	29,6%	27 (100)

Beaucoup plus fréquemment que les autres, les derniers arrivés se voient attribuer un cours multiple, ce qui est confirmé par les directeurs : « *alors, c'est très embêtant les cours doubles parce que justement, les instits ne les prennent pas, donc c'est le dernier qui arrive qui, en général, se prend le cours double* » (école 14) ; « *et l'autre personne nouvelle, elle a pris ce qui restait : le cours double ; ce n'était pas un cadeau empoisonné, on le sait maintenant, mais en septembre, on pouvait penser autrement* » (école 17).

Globalement, on observe donc que l'attribution des classes selon ce critère d'ancienneté n'est pas optimale, dans la mesure où on confie à des enseignants débutants des classes plus complexes à gérer sur le plan pédagogique. On peut rapprocher ce constat des modes d'affectations dans les écoles, collèges et lycées, qui voient en général les enseignants les moins expérimentés nommés dans les établissements réputés les plus difficiles, en terme social notamment.

Au-delà de ces deux règles qui valent dans toutes les écoles, certains arrangements sont néanmoins consentis. Ils prennent en compte le contexte particulier de chaque école et l'histoire, voire le passif, de l'équipe, mais s'inscrivent en définitive et à leur tour, sous la férule de l'usage. Deux grandes tendances se dégagent des discours des directeurs. La première a trait aux concessions qui sont faites aux enseignants qui se « dévouent » en acceptant les classes les moins prisées parce que « *si on facilite le travail du maître, on facilite l'ambiance de la classe* » (école 4). C'est notamment le cas des cours multiples : « *c'est une garantie d'avoir un peu moins d'élèves ; ça compense la surcharge de boulot... enfin, ça équilibre* » (école 18) ; « *sur un cours double, c'est certain que de toute façon, c'est plus de boulot qu'un cours simple. Mais enfin, on essaie d'arranger l'emploi du temps de la personne, les effectifs, les services pour que, d'un autre côté, il s'y retrouve un petit peu* » (école 2).

Au-delà de leur nombre, les élèves dont le maître aura la charge peuvent être objet de négociations également : « *mme X., cette maîtresse ayant un cours double, nous concevons*

<sup>30</sup> L'effectif de 137 enseignants diffère ici légèrement de celui annoncé auparavant du fait de données manquantes pour 4 enseignants.

*tous que la prise d'un cours double implique certaines difficultés et certaines tâches supplémentaires. Donc, on va tous dans l'idée de faciliter le travail du maître, on lui accorde le choix de garder ou non ses élèves » (école 4) ; « la maîtresse qui va accueillir un cours double, en général, on lui met que des bons élèves » (école 14).*

La seconde tendance, elle aussi citée par l'ensemble des directeurs, a trait à la mémoire du groupe et à la reconnaissance collective des renoncements individuels. Implicitement, s'instaure un roulement dans l'école, qui fait que chacun devra un jour ou l'autre se dévouer pour prendre la classe dont personne ne veut : *« le collègue qui fait une grosse concession, l'année suivante devient prioritaire. Ca, c'est la mémoire du groupe ; c'est pas écrit mais on le sait [...] ça, ça fonctionne bien, les gens ont cette sorte de reconnaissance » (école 5) ; « en général, ça s'organise, ça va. Il y en a toujours un qui dit : « bon, c'est bon, je le prends... mais l'an prochain, si... est-ce que tu acceptes de le prendre ? » (école 14) ; « par exemple, quelqu'un qui a été en cours double, on essaie l'année d'après de le remettre en cours simple... enfin d'alterner » (école 6). Evidemment, si l'enseignant n'a pas honoré son engagement, alors son « dédommagement » est moins assuré, comme dans l'école 14 où *« c'est celle qui a deux ans d'ancienneté qui se tape le cours double... si ça se produit encore l'année prochaine, comme elle a bénéficié de son congé maternité, c'est encore elle qui l'aurait je pense. Les collègues diraient : « attends, attends, t'as pas été beaucoup là l'an dernier ».**

En fait, c'est toute la dialectique de l'articulation entre l'équipe et l'individu que résumant ces discussions autour de l'attribution des classes aux enseignants. La conjugaison des différentes règles qui coexistent dans les écoles peut s'avérer extrêmement complexe à mettre en œuvre, même si le poids de l'ancienneté est prédominant. Il revient alors au directeur d'initier et piloter les débats et de faire parvenir son équipe à un accord, sans avoir le pouvoir d'imposer une solution. La tâche s'avère donc délicate et mieux connaître le positionnement du directeur au sein de l'équipe permet d'identifier les voies d'accès au consensus.

Les recherches montrent à cet égard que l'influence que peut avoir un directeur sur ses collègues dépend de son pouvoir et de son autorité (Bossert, Dwyer, Rowan, Lee, 1982). Selon Swidler (1979), l'autorité représente le pouvoir investi dans une personne à qui le droit de décider a été reconnu : on peut donc avoir le pouvoir sans l'autorité, et c'est ce que tendent à montrer d'ailleurs les travaux concernant les chefs d'établissement du secondaire. Les directeurs n'ayant pas de pouvoir institutionnel en France, doivent « gagner » la reconnaissance de leur autorité au cours des discussions et négociations avec leurs collègues, en conduisant avec succès un certain nombre d'opérations délicates (conflits, répartition des classes, gestion des ressources...). Finalement, plus leur autorité est reconnue, plus leur pouvoir est grand et plus leur influence au sein de l'école est importante.

### **I.3 Le rôle du directeur d'école**

Différents positionnements ont été relatés au cours des entretiens, qui se caractérisent selon leur plus ou moins grande directivité vis-à-vis des autres enseignants, la position intermédiaire étant celle qui, sans surprise, est le plus souvent adoptée.

A un extrême se situent les directeurs les plus prescriptifs, qui évoquent un fonctionnement très hiérarchisé de l'équipe. Un tiers d'entre eux a cité ce type de relations, qui tend à asseoir l'autorité dans des rapports de force. Cela dit, il est intéressant de constater que 5 de ces 6 directeurs/trices le font pour narrer des pratiques qui ne sont pas les leurs ou alors exceptionnellement : « *je donne un avis... et je peux trancher mais c'est très exceptionnel... Sinon on va au clash* » (école 1) ; « *bon je sais que ma collègue précédente, ses dernières années de carrière, préférait organiser, enfin faire sa répartition de classes avec l'inspectrice, sans se préoccuper du choix des enseignants... ça... je trouve ça gênant quand même* » (école 12). Seuls deux directeurs revendiquent un positionnement intransigeant, qui selon eux, s'avère positif. On notera avec intérêt que ces deux directeurs viennent des deux plus grandes écoles de l'échantillon (respectivement 14 et 15 classes) et n'ont pas de charge d'enseignement : « *Ca m'est arrivé une année, c'est l'arbitraire. Bon, je suis le directeur, on n'arrive pas à se mettre d'accord, j'assume mon autorité et c'est comme ça puisqu'on n'arrive pas à se mettre d'accord. Alors, bon, en général, on évite... les gens ont vu que j'étais capable de le faire donc ils savent* » (école 5) ; « *autant quand je suis arrivé, pendant quelques années, j'ai insisté sur mon rôle de directeur en étant très ferme mais maintenant les choses sont en place, les gens me connaissent. Ils savent que c'est 8H30 et pas 31, que la récré dure 20 minutes et pas 21... mais maintenant j'essaie de laisser les choses tourner et les gens me disent qu'autant au début, je paraissais très directif et je déclenchais de l'animosité, autant maintenant les gens reconnaissent que c'était une bonne méthode parce qu'ils se sentent encadrés par des règles et un système qui est installé... et qui protège* » (école 14).

Au-delà de ce type de positionnement, marginal et sans doute lié au contexte scolaire, à la taille de l'équipe et l'arrivée récente du directeur notamment, ce sont des rapports plus nuancés et conciliateurs que les directeurs évoquent le plus fréquemment. Les trois-quarts d'entre eux ont effectivement insisté sur les échanges, les discussions et les débats dans l'équipe, qui trouvent toujours ou presque une issue dans le consensus. Cela dit, une nuance peut être apportée à ce type de discours selon qu'est utilisé dans l'entretien le « je » (par opposition aux « eux ») qui témoigne d'une certaine distance, voire de hauteur par rapport aux enseignants et le « nous » ou « on » qui, au contraire, inclut le directeur dans l'équipe et rend compte alors à la fois de l'équivalence des statuts et d'une conciliation plus participative. Le premier cas, le moins fréquent, s'observe dans l'école suivante dans laquelle le directeur déclare : « *moi je les laisse d'abord se chamailler, je les laisse d'abord évacuer les choses ;*

*ça fait du bien de se dire ce qu'on a envie de dire ou qu'on garde sur le cœur... et puis en général, ça s'arrange bien comme ça ; sinon, quand je sens qu'il y a trop de torsions, il y a des critères ... des critères administratifs ».*

Dans le second cas, le fait que les enseignants s'«entendent» bien, au sens littéral comme au sens figuré, est toujours évoqué par les directeurs, qui mentionnent en outre souvent des affinités personnelles : *«ça se passe bien, on est une école où on s'entend bien ; dans les écoles où on s'entend moins bien, c'est beaucoup plus difficile »* (école 1) ; *« je crois que, oui heu..., c'est vraiment collégial, c'est-à-dire qu'il y a une espèce de confiance réciproque à la parole de l'autre et donc, il y a un enjeu important, mais ça se passe plutôt bien »* (école 5).

Enfin, à l'autre extrémité du spectre, un dernier type de fonctionnement a été relaté, dans lequel le directeur peut aller jusqu'à s'« effacer » afin de préserver la cohésion de l'équipe. On notera que ce positionnement correspond plus à des situations ponctuelles qu'à un mode relationnel routinier, même s'il est parfois systématique lors de l'attribution des classes. C'est le cas de la directrice suivante : *« bon, moi... pour faciliter les choses et parce que je veux que l'équipe pédagogique soit quand même soudée, qu'il n'y ait pas trop de problèmes dans l'équipe et tout..., généralement je prends la classe qui reste, dont personne ne veut [...] alors c'est vrai que ça me fait beaucoup de travail mais pour moi, c'est moins dur que s'il y a des tensions dans l'équipe, pour moi c'est moins dur à gérer »* (école 8).

Les autres directeurs s'estiment eux « arrangeants » ou conciliants mais leurs concessions ne vont pas aussi loin : *« j'ai choisi en m'arrangeant pour que mon choix ne gêne pas trop les anciens ; j'ai fait en sorte que mon choix respecte un peu tout le monde »* (école 7) ; *« si vous voulez, quelques fois... bon... moi, j'essaie d'être arrangeante, mais pas à mon détriment quand même... je vais pas porter toute la charge de l'école, tous les problèmes de l'école sur mes épaules »* (école 12).

Les déclarations des directeurs au sujet de la classe dont ils ont la charge confirment ces positionnements : tous n'ont manifestement pas choisi leur classe dans la mesure où sont entrées en conflit deux règles citées précédemment, la priorité donnée à l'ancienneté dans l'école et celle accordée à la fonction de direction. Il apparaît ainsi que si le directeur nommé est un nouvel arrivant, alors la règle de l'ancienneté s'applique à lui comme aux autres et il peut arriver qu'il prenne la classe qui « reste ». C'est ce qui s'est explicitement produit pour 6 des 18 directeurs interrogés, et notamment ceux qui se sont vus attribuer les cours doubles. Ces directeurs nouvellement nommés se sont donc trouvés face à une surcharge de travail, administrative et pédagogique. C'est peut-être la raison pour laquelle il est si difficile actuellement de pourvoir les postes de direction dans les petites écoles : *« ... parce que le CMI-CM2, c'était la classe qui restait... dont personne ne voulait »* (école 8) ; *« disons que*

*quand je suis arrivée, le poste qui restait était une classe de CM1-CM2 et donc je me suis adaptée à la classe et par la suite, bon...ben c'est vrai que j'y ai pris plaisir et personne ne souhaitait avoir cette classe » (école 12).*

D'autres niveaux de classe, pourtant généralement moins redoutés, ne sont néanmoins pas ceux qu'auraient choisi personnellement certains directeurs : « *sur la commune, il existe des classes de neige et tous les enfants des classes de CM1 partent 13 jours ; il est très difficile de trouver un enseignant qui accepte [...] donc, contrainte et forcée, je fais la classe de CM1 depuis 8 ans » (école 13) ; « j'ai un CE1, c'est ce qui restait » (école 16).*

Enfin, la mise en perspective des deux types de discours, concernant d'une part les règles mobilisées pour l'attribution des classes et d'autre part, le mode de leadership du directeur, fait apparaître clairement qu'un directeur nouvel arrivant dans une école n'est en définitive qu'un enseignant parmi les autres, notamment s'il a une charge d'enseignement. Dans ce cas, il se soumet aux règles en vigueur pour le choix de sa classe en s'effaçant devant les plus anciens, et son autorité doit se construire dans le temps. En revanche, s'il est déchargé, il peut se permettre d'agir avec plus d'autorité, voire de « passer en force », ainsi qu'en ont témoigné précédemment deux directeurs qui reconnaissent dans le même temps avoir suscité une certaine hostilité à leur égard.

On retrouve ici la notion de consensus, telle qu'elle est développée par Francfort, Sainsaulieu et al. (Draelants, 2002), qui montrent que le consensus renvoie, d'une part à un phénomène de partage de valeurs et d'orientations relatives à l'organisation et, d'autre part, à un système de relations sociales au sein du groupe qui a pour objectif d'éviter les conflits interindividuels. Le système de négociations se caractérise alors par le fait qu'il n'y a pas d'acteur clé exerçant sa suprématie sur les autres. La plupart des acteurs s'engagent sur des problèmes de coopération ponctuels, mènent des négociations et réalisent des ajustements dans un espace assez souple et peu formalisé : ce « compromis négocié » (Dupriez, 2003) ou « working consensus » (Bossert, Dwyer, Rowan, Lee, 1982) s'appuie principalement sur des éléments pragmatiques et concrets. A cet effet, l'application de quelques règles, simples et univoques, comme celles qui ont été évoquées, augmente la prédictibilité des décisions, réduisant ainsi les ambiguïtés et les malentendus (Fuller, Wood, Rapoport, Dornbush, 1982), et par conséquent les conflits.

Les déclarations des directeurs interrogés dans cette recherche valorisent en effet le maintien d'un climat pacifique et convivial dans l'école plus qu'un objectif d'efficacité pédagogique dans l'appariement entre enseignants et classes. On rejoint à cet égard un axe de lecture tracé par Blake et Mouton (cités par Bossert et al., 1982), qui distinguent, dans les discussions au

sein des équipes, les considérations relatives aux collègues (“affective leadership”) de celles orientées vers les résultats (“instrumental leadership”).

De toutes les typologies réalisées sur la question de savoir quels étaient les modes relationnels entre enseignants dans les écoles (travaux cités par Gather-Thurler, 1994), on peut sans doute en retenir une qui semble particulièrement bien témoigner de la réalité rapportée par les directeurs interrogés. Il s’agit de la « grande famille » qui caractérise un mode relationnel à l’intérieur duquel les membres de l’équipe sont parvenus à une certaine forme de paix sociale, grâce à la soumission de chacun à un jeu de règles explicites et implicites. Dans ce type de relations, c’est le bien-être des uns et des autres qui importe et la sociabilité prime dans la recherche de solutions pragmatiques au cours de laquelle le directeur joue le rôle de médiateur (Gather-Thurler, 1994).

Si les auteurs précédents comparent cette « grande famille » à une « fanfare de village », le consensus, tel qu’il est décrit par les directeurs interrogés, pourrait n’être néanmoins qu’un paravent de la rivalité des désirs et intérêts individuels de chacun. On peut déplorer à cet égard de ne pas disposer dans cette recherche de l’avis des enseignants sur le sujet. En effet, il est possible que certains directeurs aient eu tendance à valoriser des pratiques positives plutôt que d’évoquer les conflits inter-personnels au sein de l’école et les difficultés, voire les échecs, à les résoudre.

Les travaux anglo-saxons montrent que les directeurs d’écoles usent en général seuls de leur pouvoir de décision pour affecter les enseignants dans les classes mais que la prise en compte des préférences de ces derniers (voire des parents) constitue une variable d’ajustement permettant de préserver la stabilité de l’équipe et la collégialité (Heck, Marcoulides, Glasman, 1989). C’est ainsi que ce type de décisions et la façon dont elles sont prises au quotidien sont centraux et témoignent des compétences des directeurs (Glasman, Heck, 1987) : l’efficacité des chefs d’établissement pourrait alors transiter par cet élément organisationnel, en influençant non seulement le climat de l’école, mais aussi les progressions des élèves (Monk, 1992). Cette hypothèse est relayée empiriquement par un certain nombre de recherches, qui montrent que le niveau de satisfaction des enseignants est lié, certes indirectement, aux performances de leurs élèves (Lee, Dedrick, Smith, 1991).

Finalement, l’ensemble de ces observations interroge la question de l’attribution des classes aux enseignants dont les principes généraux semblent reposer sur une régulation centrée sur l’habitude et qui donne la priorité aux intérêts individuels et à l’ancienneté. L’intérêt général et celui des élèves ne constituent donc pas une priorité, alors même que les recherches ont montré l’influence des enseignants sur les progressions des élèves était très forte (Bressoux, 1994). On peut se demander alors s’il est possible de réaliser des appariements entre maîtres

et classes (envisagés sous l'angle de la composition des groupes d'élèves) optimaux dans le même contexte institutionnel ou s'il est nécessaire de faire évoluer le statut des directeurs. A l'heure actuelle, la primauté de l'ancienneté et de l'habitude constitue manifestement un frein au changement organisationnel et pédagogique.

Plus largement, ces résultats interrogent sur la gestion pédagogique de l'école élémentaire et la liberté accordée aux acteurs dans l'exercice de leur métier qui, s'exprime, non seulement à travers les pratiques d'enseignement (Altet, Bressoux, Bru, 1994, 1996 ; Suchaut, 1996b), mais également, comme nous venons de l'exposer, au niveau des procédures de constitution des groupes classes. Il est alors important de savoir si ces usages en vigueur à l'école élémentaire sont pertinents. Nous avons souligné que certains critères apparaissaient prioritaires aux yeux des enseignants lors de la composition des classes, parmi ceux-ci, le nombre de sections en présence est capital et les enseignants sont rarement volontaires pour prendre en charge une classes à plusieurs cours. C'est précisément ce contexte particulier dont nous allons à présent tester l'efficacité pédagogique sur des données récentes. L'analyse de la composition des classes sera poursuivie dans la mesure où les choix réalisés pour la répartition des élèves entre classes à cours simple et classe à cours multiple seront interrogés sous l'angle de l'efficacité.

## **II Efficacité pédagogique des classes à cours double**

Certes moins fréquentes qu'au siècle dernier, les classes élémentaires comptant plusieurs niveaux d'enseignement sont encore très présentes en France. La plupart des enquêtes le montrent, le grand public est en général hostile à ce type de classes, et ce dans tous les pays développés (Carleton Board of Education, 1990 ; Walsh, 1989). Les parents craignent que leur enfant n'apprenne pas aussi bien dans un cours multiple que dans un cours simple et leur réticence vis-à-vis de cette configuration de classe est forte. Les enseignants y sont eux aussi opposés mais pour d'autres raisons ; ce ne sont pas tant les effets pédagogiques qui sont évoqués, que les difficultés pour eux à enseigner dans ce type de classe, qui nécessite, outre un travail de préparation plus lourd, un « jonglage » permanent entre plusieurs niveaux d'enseignement. Les directeurs d'école s'avouent également peu favorables à la constitution de cours multiples dans leur école, certes parce qu'ils sont enseignants eux-mêmes, mais aussi parce que la procédure d'assignation des enseignants à leur classe est alors beaucoup plus conflictuelle.

On cherchera ici à soumettre ces opinions à l'épreuve des faits, en apportant un éclairage nouveau à des résultats qui s'inscrivaient jusqu'à présent dans la problématique de l'école rurale en France. La base de données mobilisée à cette fin provient d'une enquête qui s'est

déroulée dans 132 écoles élémentaires de l'Académie de Dijon dont les caractéristiques s'inscrivent très largement dans la structure académique et nationale de l'enseignement élémentaire. La collecte de données a porté plus précisément sur deux niveaux, le CE1 (3053 élèves) et le CM1 (2791 élèves). Ces deux niveaux sont particulièrement intéressants, d'une part parce qu'ils n'ont été que rarement examinés par les recherches relatives aux modes de groupement des élèves, et d'autre part, parce que leur situation dans le cursus primaire les rend propices à être intégrés dans des cours multiples.

Après avoir rappelé dans un premier temps ce que l'on sait de l'efficacité pédagogique des classes à plusieurs cours et les éléments qui ont présidé à une problématisation nouvelle de cette question, nous présenterons les différentes analyses et discuterons les résultats produits.

### ***II.1 Les travaux antérieurs sur l'efficacité des classes à cours multiple***

Les recherches conduites sur la constitution des classes dans les écoles montrent que les stratégies d'évitement des cours multiples sont manifestes (Burns, Mason, 1998) alors même que les progressions scolaires des élèves y sont au moins égales et parfois supérieures à celles d'élèves fréquentant un cours simple. En effet, les travaux réalisés sur l'école rurale en France à partir des années 80 (Vogler, Bouissou, 1987 ; Oeuvarard, 1990 ; Leroy-Audouin, Mingat, 1995) témoignent de l'efficacité des classes à plusieurs cours, et notamment des classes uniques. Cette tendance s'observe globalement, quel que soit le nombre de sections dans la classe mais aussi indépendamment de l'âge des élèves.

Dès l'école maternelle, les élèves réalisent de meilleures progressions lorsqu'ils sont scolarisés en cours double et c'est particulièrement le cas quand les enfants de l'autre section sont plus âgés. Les élèves de grande section obtiennent ainsi des résultats supérieurs s'ils fréquentent dans leur classe des élèves de cours préparatoire plutôt que des élèves de maternelle (Leroy-Audouin, Suchaut, 1994 ; Briquet-Duhazé, 2005). Dans le cursus élémentaire, l'avantage des élèves issus de cours multiples se confirme, que ce soit à l'entrée au CE2 (Oeuvarard, 1990 ; Jarousse, Mingat, 1993) ou à l'entrée en 6ème (Leroy-Audouin, Mingat, 1995), même si la configuration de la classe révèle des nuances sensibles : les cours triples en effet semblent être les moins efficaces au plan pédagogique (Bressoux, 1994). Enfin, les risques de redoublement au collège sont plus faibles pour les élèves ayant fréquenté essentiellement des classes élémentaires à plusieurs cours (Leroy-Audouin, Mingat, 1995). Une étude récente (Davezies, 2005), basée sur les analyses du panel d'élèves entrés au CP en 1997, expose des résultats plus mitigés avec un léger effet positif des cours multiples sur les progressions au niveau du CE1, mais aucun au niveau du CM1, des effets négatifs sont mêmes relevés pour les élèves de CM1 fréquentant un cours double associé à une section inférieure.

Les travaux anglo-saxons aboutissent à des résultats nettement plus mitigés. La synthèse réalisée par Veenman (1995) sur la base de 56 études en provenance de 12 pays conclut globalement à l'inexistence d'effets pédagogiques des classes à plusieurs cours sur les progressions des élèves. Que l'auteur considère les dimensions cognitive (lecture, langage, mathématiques, sciences) ou non cognitive (attitudes et compétences relationnelles, estime de soi, motivation) des apprentissages, aucune différence n'apparaît entre élèves fréquentant cours simples et cours multiples et ce constat rejoint celui réalisé auparavant par Pratt (1986) et Miller (1990). Burns et Mason (1996), après avoir observé que les élèves de bon niveau scolaire et les plus autonomes sont prioritairement affectés en cours multiples, avancent un effet négatif de ces classes sur les progressions des élèves : ainsi, les cours multiples auraient une efficacité pédagogique spécifique négative mais celle-ci serait contrebalancée par des effets positifs de la composition du groupe d'élèves. Selon Wilkinson et Hamilton (2003), plus que le type de classe, ce sont les pratiques pédagogiques mises en œuvre qui importent. Quels que soient les facteurs évoqués en tout cas, jamais, à la lumière des recherches anglo-saxonnes, les élèves ne progressent significativement mieux en cours multiple qu'en cours simple. Les résultats obtenus dans le contexte français peuvent paraître à la fois datés, dans la mesure où ils impliquent généralement simultanément type de classe et milieu géographique (dont on sait par ailleurs qu'il est associé à certaines particularités des écoles et des enseignants) et relativement spécifiques puisqu'ils se démarquent des résultats obtenus dans les autres pays. Plusieurs éléments nouveaux incitent à questionner les résultats des recherches françaises.

La première dimension concerne la spécificité rurale des classes à plusieurs cours qui, à la lumière des statistiques disponibles (MEN, 1982 ; MENRT, 1999) peut être considérée comme obsolète. Aujourd'hui, les classes à cours multiple, qui sont d'ailleurs en grande majorité des classes à cours double, se répartissent presque également entre les zones urbaines et les zones rurales. La seconde dimension a trait à l'affectation des élèves dans les classes, dont Burns et Masons (2002) déclarent qu'elle interagit fortement avec l'efficacité pédagogique. On peut distinguer dans cette perspective les écoles qui ont une certaine liberté dans l'affectation des élèves et celles qui n'en disposent pas. Les premières sont celles qui offrent plusieurs classes proposant un niveau d'enseignement considéré ; les enseignants ont alors le choix d'affecter tel élève dans telle classe plutôt que dans telle autre, sur la base de critères qu'eux seuls décident. Les secondes sont celles dans lesquelles il n'existe qu'une classe, les élèves y étant inscrits « d'office ». La troisième dimension concerne les méthodes d'analyse statistique mobilisées. En France, jusqu'à une période récente, les analyses statistiques utilisées pour mesurer l'impact des variables de politique éducative se limitaient à l'utilisation des modèles de régression dits « mononiveau », nommés dans un langage statistique, modèles M.C.O. (Moindres Carrés Ordinaires). Si ces modèles sont bien adaptés à

la mesure des effets des variables relatives aux élèves, la mesure des effets de contexte est plus problématique et donne lieu à des estimations incorrectes sur le plan statistique (Bressoux, Coustère, Leroy-Audouin, 1997). Depuis quelques années, l'utilisation de techniques plus sophistiquées permet d'évaluer les effets de contexte avec beaucoup plus de précision, il s'agit des modèles multiniveau et ce sont ces modèles que nous utiliserons pour conduire les analyses<sup>31</sup>.

## ***II.2 Cours simple versus cours multiple : quelle efficacité globale?***

Les progressions de 2294 élèves de CM1 ont pu être évaluées, 1381 d'entre eux fréquentant des cours simples, 913 des cours multiples. En CE1, 3053 élèves (2111 en cours simples et 942 en cours multiples) ont été concernés. Deux types d'informations ont été collectés. Le premier concerne les classes (type de cours, nombre de sections, nombre d'élèves total et par section), le second concerne les élèves de ces classes : origine sociale, sexe, retard scolaire et niveau de compétences en français et mathématiques mesuré par des épreuves standardisées administrées en début et en fin d'année scolaire. Ces épreuves sont comparables à celles utilisées dans les évaluations nationales, tout en comportant un nombre moins important d'items. Elles explorent plusieurs domaines des apprentissages : compréhension, outils de la langue, production d'écrits, résolution de problèmes, mesure et numération. Ces épreuves rendent compte dans cette recherche plus d'un niveau moyen d'acquisitions des élèves que de la maîtrise de compétences spécifiques, et sont donc assimilables à celles qui ont été utilisées dans les travaux antérieurs. Par ailleurs, L'échelle de mesure des scores des élèves présente une moyenne de 100 et un écart-type de 15.

Le tableau 13 présente quelques caractéristiques des élèves qui permettent de comparer la composition des classes selon le type de cours fréquenté. Au CE1, quand on oppose les élèves de cours simple à ceux de cours doubles, la seule différence statistiquement significative porte sur le niveau initial de mathématiques : les élèves de cours doubles ont en moyenne un score supérieur de 3,2 points (différence significative à .05). En revanche, quand on compare le type de cours selon les sections qu'il associe, les élèves se distinguent de façon beaucoup plus flagrante. Il apparaît clairement (et de façon significative sur le plan statistique) que les élèves qui fréquentent un CE1-CE2 ont dans l'ensemble des scores plus élevés que ceux qui fréquentent un CP-CE1 (3 points de différence en français et 2,5 en mathématiques). En outre, les élèves en retard scolaire sont plus nombreux en CE1-CE2 qu'en CP-CE1, et de façon liée, les filles sont moins représentées dans le premier type de cours double. La même analyse

---

<sup>31</sup> Plusieurs logiciels permettent de réaliser des analyses multiniveaux ; nous avons pour notre part utilisé la procédure « *proc mixed* » intégrée dans le logiciel SAS depuis une dizaine d'années. On trouvera in Singer (1998) une présentation détaillée de son utilisation dans un exemple directement appliqué au champ de la recherche en éducation.

répliquée en CM1 conduit à des résultats moins tranchés. Seul le niveau de mathématiques distingue significativement les élèves de cours simple et de cours multiple dans un premier temps (différence de 1,4 points significative à .10), puis de façon plus marquée les élèves de CE2-CM1 et CM1-CM2 dans un second temps (différence de 3,7 points significative à .01) : les élèves dont le niveau de mathématiques à l'entrée au CM1 est plus élevé fréquentent plus systématiquement un cours double avec les élèves de la section supérieure.

Tableau 13 : Description des caractéristiques des élèves selon le type de cours fréquenté

Elèves de CE1	Total	Cours simple	Cours multiple	CP-CE1	CE1-CE2
Niveau initial français	98,7	98,4	99,3	98,0	100,9
Niveau initial mathématiques	98,4	97,3	100,5	99,4	101,8
% d'élèves ayant déjà redoublé	15,7	15,1	16,9	13,8	20,6
% de filles	49,1	49,9	47,8	50,8	44,3
% d'enfants de père cadre	7,3	6,5	8,7	8,7	8,8
% d'enfants de mère active	48,6	49,6	46,5	44,3	49,1
Elèves de CM1	Total	Cours simple	Cours multiple	CE2-CM1	CM1-CM2
Niveau initial français	99,2	98,9	99,5	98,8	100,0
Niveau initial mathématiques	99,3	98,7	100,1	98,0	101,7
% d'élèves ayant déjà redoublé	21,2	20,6	22,1	25,1	20,0
% de filles	50,5	49,8	51,5	53,1	50,3
% d'enfants de père cadre	8,5	8,1	9,2	8,4	9,7
% d'enfants de mère active	48,9	50,3	46,5	45,4	47,4

Nous distinguerons plusieurs étapes dans la présentation des estimations relatives à la mesure de l'efficacité pédagogique : la première concerne les modèles estimés sur l'échantillon complet, prenant en compte l'ensemble des classes. On opposera ainsi dans l'analyse et en premier lieu, les classes à cours multiples aux classes à cours simple, puis en second lieu les classes à cours simple aux différentes configurations de cours multiples : élèves regroupés avec des élèves de la section inférieure ou élèves regroupés avec des élèves de la section supérieure.

Le tableau 14 présente les estimations réalisées au niveau du CE1 dans les trois dimensions des acquisitions des élèves. Dans un premier temps, un modèle dit « vide » (*unconditional model*) n'incluant aucune variable explicative a été estimé (non présenté ici). Ce modèle produit une simple décomposition de la variance totale des acquis scolaires en fin d'année en

une part de variance inter-classes et une part de variance intra-classe. Ainsi, pour le CE1, la part de la variance totale expliquée par la variance entre les classes est de 21,17% en français et de 26,42% en mathématiques. Ces chiffres témoignent de l'intérêt de prendre en compte le niveau contextuel dans les analyses puisque environ un quart des différences de scores entre les élèves s'explique par l'appartenance à une classe.

*Tableau 14 : Modèles multiniveau de progression en CE1, score de français, de mathématiques et global (échantillon complet)*

Paramètres	Modèle A		Modèle B		Modèle C	
	Français		Mathématiques		Score global	
Effets fixes						
Constante	41,59	(1,59) ***	37,68	(1,54) ***	31,44	(1,41) ***
Score individuel initial	0,58	(0,01) ***	0,63	(0,01) ***	0,69	(0,01) ***
Sexe : Fille (référence garçon)	1,99	(0,41) ***	0,28	(0,38) n.s.	0,89	(0,32) ***
Profession du père (référence père ouvrier)						
père artisan-commerçant	1,73	(0,73) ***	0,55	(0,68) n.s.	0,98	(0,57) *
profession intermédiaire	1,93	(0,64) ***	1,31	(0,60) **	1,08	(0,50) **
père employé	1,48	(0,59) ***	-0,00	(0,55) n.s.	0,54	(0,46) n.s.
père cadre	3,74	(0,83) ***	1,67	(0,77) **	2,06	(0,65) ***
Mère avec emploi qualifié (référence mère au foyer ou non qualifiée)	0,22	(0,44) n.s.	0,79	(0,41) *	0,25	(0,34) n.s.
Elève redoublant (référence élève à l'heure)	-5,99	(0,62) ***	-5,83	(0,57) ***	-5,69	(0,48) ***
Cours multiple (référence cours simple)	-1,24	(0,79) n.s.	-2,76	(0,88) ***	-2,04	(0,72) ***
Effets aléatoires						
Niveau 2 : variance interclasse	18,82	(2,84)	27,41	(3,73)	17,96	(2,46)
Niveau 1 : variance intraclasse	106,43	(3,02)	91,43	(2,59)	62,86	(1,79)
Pourcentage de variance interclasse expliquée	61,14%		54,4%		63,98%	
Pourcentage de variance intraclasse expliquée	40,96%		45,45%		54,66%	
- 2 log L	20206,2		19983,7		18695,3	

N=2684 (français) N=2682 (Mathématiques) N=2662 (score global)

Seuil de significativité des effets fixes : n.s. : non significatif, \* : significatif à .10, \*\* : significatif à .05, \*\*\* : significatif à .01. Les erreurs type des coefficients figurent entre parenthèses.

Note de lecture : les modèles donnent 1) des effets fixes, c'est-à-dire l'impact spécifique de chaque variable explicative sur la variable dépendante, 2) des effets aléatoires, permettant de décomposer la variance globale du phénomène, et donnant pour chaque niveau un coefficient aléatoire représentant la part de variance résiduelle (c'est-à-dire non expliquée par le modèle) qui lui est attachée, 3) les parts de variance expliquée à chacun des niveaux, qui sont un indicateur de la pertinence globale du modèle testé.

Les trois modèles A, B et C fournissent des informations relatives à l'influence des caractéristiques des élèves sur leurs progressions en cours de CE1. Nous retrouvons là des résultats à présent bien connus, et notamment le poids des différenciations sociales. Celles-ci sont particulièrement fortes en Français, discipline dans laquelle les enfants d'ouvriers sont ceux qui progressent le moins par rapport à tous les autres et les enfants de cadres le plus : près de 4 points séparent ces deux catégories d'élèves. Les écarts sont moins marqués en Mathématiques et seuls les enfants dont le père est cadre ou de profession intermédiaire connaissent des progressions significativement meilleures (respectivement 1,67 et 1,31

points). Les filles progressent légèrement plus que les garçons en cours d'année mais en français seulement. Là encore, ce constat rejoint la plupart de ceux effectués sur cette question (Duru-Bellat, 1990). Enfin, le retard scolaire apparaît très pénalisant pour les élèves : ceux qui ont déjà redoublé de façon précoce (le CP ou le CE1) progressent significativement moins bien que les autres, toutes choses égales par ailleurs (et notamment à niveau initial donné) et cet écart est très fort puisqu'il se situe aux environs de 6 points dans chacune des deux disciplines : c'est dire que cette mesure ne remplit pas ses objectifs pédagogiques, au moins dans le court terme et c'est d'ailleurs ce que montrent les travaux conduits sur le sujet (Paul, 1996).

De façon globale, il apparaît que les élèves de CE1 ont tout à gagner de fréquenter une classe à cours simple plutôt qu'une classe à cours multiple. Hormis le français, discipline pour laquelle le coefficient n'est pas significatif, les deux autres modèles concluent à un effet négatif et significatif des classes à cours multiples. Ainsi, des élèves de caractéristiques par ailleurs comparables connaissent des écarts de progressions de l'ordre de 2,8 points en mathématiques et de 2 points sur le score global selon le type de classe fréquenté. On constate ainsi que le type de cours a autant d'influence sur les progressions des élèves que l'origine sociale puisque les écarts entre enfants de cadres et enfants d'ouvriers sont globalement du même ordre que ceux liés au type de cours. Afin d'explorer davantage ces premiers résultats, les mêmes estimations ont été réalisées en distinguant dans les cours multiples, ceux qui associent la section inférieure (le cours préparatoire) de ceux qui associent une section supérieure (le plus souvent une section de CE2). Le tableau 15 en consigne les résultats.

*Tableau 15 : Effets des différentes configurations de type de cours sur les progressions au CE1 (échantillon complet)*

	Français	Mathématiques	Global
Cours multiple avec section inférieure	-1,29 (1,03) n.s.	-2,29 (1,14) **	-1,73 (0,93) **
Cours multiple avec section supérieure	-1,27 (0,95) n.s.	-3,27 (1,06) ***	-2,42 (0,86) ***

Seuil de significativité des effets fixes : n.s. : non significatif, \* : significatif à .10, \*\* : significatif à .05, \*\*\* : significatif à .01. Les erreurs type des coefficients figurent entre parenthèses.

Une structure similaire à la précédente est relevée, à savoir, des effets négatifs et significatifs en mathématiques et sur le score global et des effets négatifs mais non significatifs en français. Quelle que soit la configuration de classe envisagée, les progressions réalisées sont moins bonnes dans les classes à cours multiples et l'examen des coefficients invite à penser que cet effet négatif est plus marqué quand le cours multiple associe les élèves d'une section supérieure. On notera que ce résultat va à l'encontre des déclarations des enseignants selon lesquels la configuration la plus problématique parmi toutes est celle du CP-CE1.

Le tableau 16 réplique les analyses précédentes au niveau du CM1. L'examen des modèles « vides » nous indique tout d'abord que les différences entre classes sont moins élevées qu'en CE1 : les pourcentages de variance inter-classes sont en effet plus faibles (16,62% pour le français et 23,13% pour les mathématiques), ce qui tend à montrer que le contexte classe exerce moins d'influence.

*Tableau 16 : Modèles multiniveau de progression en CMI, score de français, de mathématiques et global (échantillon complet)*

Paramètres	Modèle D		Modèle E		Modèle F	
	Français		Mathématiques		Score global	
Effets fixes						
Constante	33,01	(1,39) ***	43,37	(1,63) ***	26,68	(1,35) ***
Score individuel initial	0,67	(0,01) ***	0,57	(0,01) ***	0,94	(0,01) ***
Sexe : Fille (référence garçon)	-0,16	(0,34) n.s.	0,33	(0,38) n.s.	0,11	(0,28) n.s.
Profession du père (référence père ouvrier)						
père artisan-commerçant	-0,41	(0,61) n.s.	0,21	(0,68) n.s.	-0,39	(0,50) n.s.
profession intermédiaire	1,20	(0,55) **	2,30	(0,61) ***	1,34	(0,45) ***
père employé	0,04	(0,50) n.s.	0,55	(0,56) n.s.	0,04	(0,41) n.s.
père cadre	4,18	(0,63) ***	3,69	(0,70) ***	3,13	(0,51) ***
Mère avec emploi qualifié (référence mère au foyer ou non qualifiée)	1,08	(0,37) ***	0,95	(0,42) **	0,69	(0,31) **
Elève redoublant (référence élève à l'heure)	-5,93	(0,48) ***	-6,76	(0,53) ***	-5,02	(0,40) ***
Cours multiple (référence cours simple)						
	-1,11	(0,69) n.s.	-0,27	(0,99) n.s.	-0,76	(0,75) n.s.
Effets aléatoires						
Niveau 2 : variance interclasse	15,99	(2,34)	37,38	(4,83)	21,54	(2,76)
Niveau 1 : variance intraclasse	74,03	(2,09)	92,52	(2,62)	48,58	(1,40)
Pourcentage de variance interclasse expliquée	57,61%		29,44%		46,74%	
Pourcentage de variance intraclasse expliquée	60,85%		47,45%		67,51%	
- 2 log L	19508,2		20112,8		17880,5	

N= 2719(français) N=2706 (Mathématiques) N=2634 (score global)

Seuil de significativité des effets fixes : n.s. : non significatif, \* : significatif à .10, \*\* : significatif à .05, \*\*\* : significatif à .01. Les erreurs type des coefficients figurent entre parenthèses.

Les caractéristiques individuelles ont une influence similaire à celle observée en CE1 sur les progressions, mais de façon plus atténuée. Les filles réalisent des progressions identiques à celles des garçons dans les deux disciplines. Les différenciations sociales existent mais elles sont moindres dans la mesure où seuls les enfants de cadre ou de père occupant une profession intermédiaire se distinguent des enfants d'ouvrier par leurs meilleurs progressions (respectivement 3 et 1,3 points au niveau du score global). Cela dit, l'activité de la mère apparaît influente à ce niveau d'enseignement et les enfants dont la mère occupe un emploi qualifié progressent significativement mieux que les autres (environ 1 point). L'effet négatif du retard scolaire se confirme : comme au CE1 et à niveau initial donné, les élèves qui comptent 1 an (ou plus) de retard à ce niveau de la scolarité (et quelle que soit la classe redoublée) réalisent des progressions bien inférieures à celles des autres élèves : près de 6

points de moins en Français et près de 7 en Mathématiques. La conclusion pourrait être la même que précédemment, si ce n'est que ces difficultés se cumulent visiblement au fil des années, fragilisant ces élèves qui finalement ne rattrapent pas le niveau de progression des autres. Enfin, le type de classe, contrairement à ce qui avait été observé en CE1, n'influence pas significativement les progressions des élèves, même si on peut souligner le signe négatif des coefficients dans les différents modèles. Le tableau 17 confirme ce résultat et indique qu'il n'existe pas d'effet significatif spécifique selon la configuration du type de cours.

*Tableau 17 : Effets des différentes configurations de type de cours sur les progressions au CM1 (échantillon complet)*

	Français	Mathématiques	Global
Cours multiple avec section inférieure	-1,11 (0,91) n.s.	-1,05 (1,29) n.s.	-0,89 (0,98) n.s.
Cours multiple avec section supérieure	-1,10 (0,83) n.s.	0,33 (1,19) n.s.	-0,67 (0,90) n.s.

Seuil de significativité des effets fixes : n.s. : non significatif, \* : significatif à .10, \*\* : significatif à .05, \*\*\* : significatif à .01. Les erreurs type des coefficients figurent entre parenthèses.

Si l'on compare les deux niveaux d'enseignement, seuls les élèves de CE1 pâtissent de la fréquentation d'un cours double, quelle que soit sa configuration. A ce stade de l'analyse, il est difficile de trancher en faveur d'une piste d'explication unique. Deux dimensions peuvent être évoquées : d'une part, l'âge des élèves qui fait que les plus jeunes pourraient être plus sensibles à l'influence du contexte scolaire et d'autre part, le contenu des programmes (densité, difficulté) qui, selon le niveau d'enseignement, pourrait se plier plus ou moins bien aux exigences des classes à plusieurs cours. Ce résultat global mérite d'être exploré davantage. En effet, les travaux évoqués précédemment invitent à penser que l'effet du type de cours peut varier selon que les élèves ont été ou non choisis car les écoles adoptent des pratiques assez tranchées en la matière, insistant sur l'autonomie des élèves affectés en cours multiples.

On peut alors faire l'hypothèse, à la suite de Burns et Mason (1998), que les caractéristiques spécifiques des élèves concernés peuvent contrebalancer les effets du type de classe. Si cette hypothèse se vérifie, qui renvoie au fait que les élèves qui fréquentent les cours multiples sont ceux qui peuvent se dispenser le plus facilement de la présence du maître, alors quand les écoles ont l'opportunité de réaliser délibérément cette affectation, les effets des cours multiples peuvent être atténués. En revanche, dans le cas où tous les élèves d'une cohorte, sans distinction, fréquentent un cours multiple parce qu'il n'y a pas d'autre solution possible dans l'école, alors certains d'entre eux ne manifestent sans doute pas les « qualités » que ce contexte particulier requiert : les effets des cours multiples peuvent être alors être plus forts et négatifs qu'il n'y paraît de prime abord. C'est ce que nous allons examiner à présent.

### *II.3 L'affectation des élèves dans les classes*

Les analyses précédentes ont été répliquées en distinguant deux catégories d'écoles : celles dans lesquelles les enseignants ont la possibilité d'intervenir sur l'affectation des élèves et celles dans lesquelles ceci n'est pas possible, les élèves étant assignés directement dans la seule classe existante.

La première situation rend compte des cas dans lesquels les enseignants n'ont pas le choix pour constituer la classe de CE1. Sont donc ici concernées toutes les écoles qui n'ont qu'une classe de CE1, qu'il s'agisse d'un cours simple ou d'un cours multiple dans le cas où les effectifs ont été insuffisants pour constituer une classe complète à cours simple. Pour rappel, ces écoles sont au nombre de 84 et totalisent 51 classes à cours simples (1087 élèves) et 33 cours multiples (399 élèves). Dans ces écoles, les effets pédagogiques des cours multiples apparaissent franchement négatifs (tableau 18). Dans les deux disciplines, les coefficients sont plus élevés et plus significatifs que dans les modèles généraux estimés précédemment : 2,5 points en Français, 3 en mathématiques et 2,5 au niveau du score global séparent les progressions d'élèves comparables, selon qu'ils fréquentent un cours simple ou un cours multiple dans une école qui les y a placés « d'office », sans possibilité de choix individualisés.

Les détails donnés par le tableau 19 montrent qu'il existe des effets symétriques selon la discipline. Si les progressions globales sont également et négativement affectées (un peu plus de 2,5 points quelle que soit la configuration du cours multiple), en revanche, en français, les effets sont plus marqués dans le cas d'un CP-CE1 (plutôt qu'un CE1-CE2) alors qu'en mathématiques, c'est l'inverse : -3,7 points pour des élèves fréquentant un CE1-CE2 et un effet non significatif en CP-CE1.

*Tableau 18 : Modèles multiniveau de progression en CE1, score de français, de mathématiques et global  
(écoles qui ne comptent qu'une classe de CE1)*

Paramètres	Modèle G		Modèle H		Modèle I	
	Français		Mathématiques		Score global	
Effets fixes						
Constante	47,28	(2,29) ***	44,16	(2,38) ***	37,57	(2,14) ***
Score individuel initial	0,53	(0,02) ***	0,58	(0,02) ***	0,64	(0,02) ***
Sexe : Fille (référence garçon)	1,94	(0,58) ***	0,16	(0,56) n.s.	0,73	(0,46) *
Profession du père (référence père ouvrier)						
père artisan-commerçant	1,55	(0,98) n.s.	-0,75	(0,97) n.s.	0,29	(0,79) n.s.
profession intermédiaire	2,01	(0,87) **	1,28	(0,85) n.s.	1,20	(0,70) *
père employé	0,14	(0,84) n.s.	-0,72	(0,81) n.s.	-0,41	(0,67) n.s.
père cadre	3,37	(1,13) ***	0,81	(1,10) n.s.	1,56	(0,91) *
Mère avec emploi qualifié (référence mère au foyer ou non qualifiée)	-0,14	(0,61) n.s.	0,37	(0,59) n.s.	0,01	(0,49) n.s.
Elève redoublant (référence élève à l'heure)	-5,88	(0,88) ***	-5,49	(0,86) ***	-5,39	(0,71) ***
Cours multiple (référence cours simple)	-2,50	(1,21) **	-3,05	(1,36) **	-2,51	(1,14) **
Effets aléatoires						
Niveau 2 : variance interclasse	19,37	(4,23)	26,73	(5,47)	19,03	(3,84)
Niveau 1 : variance intraclasse	95,89	(3,98)	90,58	(3,75)	60,23	(2,52)
Pourcentage de variance interclasse expliquée	43,03%		41,30%		49,38%	
Pourcentage de variance intraclasse expliquée	40,52%		39,84%		51,00%	
- 2 log L	9321,8		9303,1		8681,8	

N=1247 (français) N=1251 (Mathématiques) N=1234 (score global)

Seuil de significativité des effets fixes : n.s. : non significatif, \* : significatif à .10, \*\* : significatif à .05, \*\*\* : significatif à .01. Les erreurs type des coefficients figurent entre parenthèses.

*Tableau 19: Effets des différentes configurations de type de cours sur les progressions au CE1 (écoles qui ne comptent qu'une classe de CE1)*

	Français	Mathématiques	Global
Cours multiple avec section inférieure	-3,63 (1,96) *	-1,95 (2,17) n.s.	-2,58 (1,82) n.s.
Cours multiple avec section supérieure	-2,20 (1,34) *	-3,76 (1,50) ***	-2,76 (1,26) **

Seuil de significativité des effets fixes : n.s. : non significatif, \* : significatif à .10, \*\* : significatif à .05, \*\*\* : significatif à .01. Les erreurs type des coefficients figurent entre parenthèses.

La seconde situation rend compte des écoles qui offrent plusieurs classes avec des élèves de CE1; dans ce cas, les équipes affectent intentionnellement les élèves dans l'une ou l'autre des classes possibles. Les tableaux 20 et 21 offrent une comparaison intéressante sur deux plans : d'une part, le rôle des caractéristiques individuelles des élèves et d'autre part, l'influence de la fréquentation d'un cours multiple. Concernant le premier point, on observe que les effets des caractéristiques individuelles sont plus forts quand les élèves sont délibérément affectés dans une classe.

Tableau 20 : Modèles multiniveau de progression en CE1, score de français, de mathématiques et global (écoles qui offrent le choix, dont un cours double)

Paramètres	Modèle J		Modèle K		Modèle L	
	Français		Mathématiques		Score global	
Effets fixes						
Constante	32,68	(2,84) ***	29,71	(2,60) ***	22,51	(2,42) ***
Score individuel initial	0,64	(0,03) ***	0,68	(0,02) ***	0,75	(0,02) ***
Sexe : Fille (référence garçon)	2,16	(0,73) ***	0,36	(0,62) n.s.	1,15	(0,54) **
Profession du père (référence père ouvrier)						
père artisan-commerçant	1,70	(1,27) n.s.	2,34	(1,08) **	1,68	(0,94) *
profession intermédiaire	2,77	(1,19) **	2,33	(1,02) **	1,98	(0,88) **
père employé	2,50	(1,06) **	0,63	(0,91) n.s.	1,33	(0,79) *
père cadre	4,04	(1,39) ***	2,45	(1,19) **	2,59	(1,03) ***
Mère avec emploi qualifié (référence mère au foyer ou non qualifiée)	0,66	(0,80) n.s.	1,58	(0,68) **	0,73	(0,59) n.s.
Elève redoublant (référence élève à l'heure)	-5,99	(1,13) ***	-6,80	(0,97) ***	-6,05	(0,85) ***
Cours multiple (référence cours simple)	1,39	(1,23) n.s.	-1,00	(0,68) n.s.	-0,04	(1,13) n.s.
Effets aléatoires						
Niveau 2 : variance interclasse	15,89	(4,27)	30,65	(6,49)	16,24	(3,71)
Niveau 1 : variance intraclasse	121,61	(5,77)	87,45	(4,16)	65,54	(3,14)
Pourcentage de variance interclasse expliquée	75,43%		56,19%		72,98%	
Pourcentage de variance intraclasse expliquée	40,61%		53,74%		59,21%	
- 2 log L	7413,7		7186,8		6791,1	

N= 964 (français) N= 969 (Mathématiques) N= 955 (score global)

Seuil de significativité des effets fixes : n.s. : non significatif, \* : significatif à .10, \*\* : significatif à .05, \*\*\* : significatif à .01. Les erreurs type des coefficients figurent entre parenthèses.

C'est le cas du niveau initial notamment qui est lié de façon significativement plus forte au niveau final : le coefficient atteint 0,75 (il était de 0,64 pour les élèves issus d'écoles n'ayant pas réalisé de choix d'affectation). C'est le cas également de la profession du père qui génère des différenciations sociales plus fortes qu'auparavant : les écarts de progression entre élèves s'accroissent ainsi dans ce type de contexte, les enfants de cadres ayant un avantage de 2,6 points sur les enfants d'ouvriers. Les filles progressent également significativement mieux que les garçons et les difficultés des élèves en retard scolaire augmentent : en effet, leurs progressions, à caractéristiques individuelles comparables, sont inférieures de plus de 6 points à celles de élèves « à l'heure ». Il semble donc que l'opportunité qui est donnée aux enseignants de composer les groupes d'élèves ne conduit pas à la constitution de classes moins différenciatrices et plus équitables sur le plan des progressions individuelles. Concernant le second point, un résultat majeur apparaît : la disparition des effets négatifs attachés aux cours multiples. Quelles que soient la discipline évaluée et la configuration de classe (tableau 21), les progressions des élèves ne sont pas affectées: autrement dit, quand les enseignants procèdent délibérément à l'affectation des élèves dans les différents types de classes proposés, alors il n'existe plus d'effet négatif des cours multiples, sans doute et cela demandera à être confirmé, parce que les élèves qui y sont affectés sont ceux qui sont les mieux « armés » pour y réussir.

Tableau 21 : Effets des différentes configurations de type de cours sur les progressions au CE1 (écoles qui offrent le choix, dont un cours double)

	Français	Mathématiques	Global
Cours multiple avec section inférieure	1,59 (1,41) n.s.	-0,61 (1,70) n.s.	0,45 (1,29) n.s.
Cours multiple avec section supérieure	1,14 (1,49) n.s.	-1,46 (1,77) n.s.	-0,62 (1,35) n.s.

Seuil de significativité des effets fixes : n.s. : non significatif, \* : significatif à .10, \*\* : significatif à .05, \*\*\* : significatif à .01. Les erreurs type des coefficients figurent entre parenthèses.

L'analyse a été répliquée au niveau du CM1, et à nouveau, ne sont concernées dans la première phase que les écoles qui n'ont qu'une classe de CM1, qu'il s'agisse d'un cours simple ou d'un cours multiple dans le cas où les effectifs ont été insuffisants pour constituer une classe complète à cours simple. On relève sur ce sous échantillon la même répartition en fonction du type de cours qu'au CE1, à savoir environ 60% dans lesquelles le CM1 est un cours simple. Dans ces écoles où n'ont pas eu lieu de procédures d'affectation des élèves, on observe (tableaux 22 et 23) que les effets pédagogiques des cours multiples sont significativement, mais faiblement, négatifs en français et sur le score global (respectivement -1,57 et -1,87 points).

Tableau 22: Modèles multiniveau de progression en CM1, score de français, de mathématiques et global (écoles qui ne comptent qu'une classe de CM1)

Paramètres	Modèle M		Modèle N		Modèle O	
	Français		Mathématiques		Score global	
Effets fixes						
Constante	32,39	(1,95) ***	41,52	(2,44) ***	25,77	(1,94) ***
Score individuel initial	0,68	(0,02) ***	0,58	(0,02) ***	0,75	(0,02) ***
Sexe : Fille (référence garçon)	0,11	(0,50) n.s.	0,67	(0,57) n.s.	0,40	(0,41) n.s.
Profession du père (référence père ouvrier)						
père artisan-commerçant	-0,76	(0,83) n.s.	0,81	(0,95) n.s.	-0,14	(0,70) n.s.
profession intermédiaire	1,58	(0,78) **	1,67	(0,91) *	1,17	(0,66) *
père employé	-0,26	(0,72) n.s.	1,17	(0,85) n.s.	0,26	(0,62) n.s.
père cadre	3,99	(0,85) ***	3,33	(0,98) ***	2,97	(0,72) ***
Mère avec emploi qualifié (référence mère au foyer ou non qualifiée)	1,59	(0,52) ***	1,18	(0,62) **	1,09	(0,45) ***
Elève redoublant (référence élève à l'heure)	-4,36	(0,70) ***	-6,25	(0,81) ***	-3,83	(0,60) ***
Cours multiple (référence cours simple)	-1,57	(0,99) *	-1,81	(1,49) n.s.	-1,87	(1,10) *
Effets aléatoires						
Niveau 2 : variance interclasse	11,69	(2,75)	30,97	(6,30)	16,80	(3,40)
Niveau 1 : variance intraclasse	66,84	(2,85)	90,15	(3,85)	45,76	(1,99)
Pourcentage de variance interclasse expliquée	54,85%		33,58%		51,14%	
Pourcentage de variance intraclasse expliquée	63,48%		47,75%		68,71%	
- 2 log L	8447,4		8793,0		7740,0	

N=1188 (français) N= 1182 (Mathématiques) N=1143 (score global)

Seuil de significativité des effets fixes : n.s. : non significatif, \* : significatif à .10, \*\* : significatif à .05, \*\*\* : significatif à .01. Les erreurs type des coefficients figurent entre parenthèses.

Enfin, le tableau 23 révèle que ce sont particulièrement les cours doubles associés à la section inférieure qui grèvent les progressions des élèves : en effet, à caractéristiques individuelles comparables, les élèves qui fréquentent un CE2-CM1 ont des progressions inférieures de 2,7 points en français et 3,6 points en mathématiques à celles des élèves scolarisés en cours simples ou en CM1-CM2.

*Tableau 23 : Effets des différentes configurations de type de cours sur les progressions au CM1 (écoles qui ne comptent qu'une classe de CM1)*

	Français	Mathématiques	Global
Cours multiple avec section inférieure	-2,70 (1,36) **	-3,63 (2,04) *	-3,07 (1,50) **
Cours multiple avec section supérieure	-0,75 (1,20) n.s.	-0,46 (1,82) n.s.	-0,99 (1,33) n.s.

Seuil de significativité des effets fixes : n.s. : non significatif, \* : significatif à .10, \*\* : significatif à .05, \*\*\* : significatif à .01. Les erreurs type des coefficients figurent entre parenthèses.

Comme au niveau du CE1, la seconde phase de l'analyse se centre sur les écoles qui offrent plusieurs classes scolarisant des élèves de CM1 ; dans ce cas, les équipes affectent sciemment les élèves dans l'une ou l'autre des classes possibles.

Les tableaux 24 et 25 offrent une comparaison avec les résultats précédents. Concernant les caractéristiques des élèves, et contrairement à ce qui s'observe en CE1, les écarts de progressions sont d'ampleur comparable. Que ce soit la profession du père, l'activité qualifiée de la mère, voire le sexe de l'élève, aucune de ces variables n'intervient différemment selon que les élèves sont affectés volontairement ou non dans les classes. Le retard scolaire est la seule variable dont l'influence varie : paradoxalement, son effet négatif est beaucoup plus fort quand les enseignants procèdent à des choix d'affectation (-3,8 points au niveau du score global quand il n'y a pas de « choix » et -6 points quand il y a « choix »). Tout semble se passer comme si l'opportunité pour ces élèves de bénéficier d'un traitement particulier et d'une affectation « ciblée » avait en fait des conséquences néfastes pour leurs progressions.

*Tableau 24 : Modèles multiniveau de progression en CM1, score de français, de mathématiques et global  
(écoles qui offrent le choix, dont un cours double)*

Paramètres	Modèle P		Modèle Q		Modèle R	
	Français		Mathématiques		Score global	
Effets fixes						
Constante	31,25	(2,51) ***	42,58	(2,80) ***	25,31	(2,40) ***
Score individuel initial	0,68	(0,02) ***	0,56	(0,02) ***	0,74	(0,02) ***
Sexe : Fille (référence garçon)	-0,77	(0,57) n.s.	-0,05	(0,61) n.s. ;	-0,25	(0,46) n.s.
Profession du père (référence père ouvrier)						
père artisan-commerçant	-0,24	(1,11) n.s.	-1,80	(1,18) n.s.	-1,35	(0,89) n.s.
profession intermédiaire	0,54	(0,90) n.s.	3,08	(0,96) ***	1,48	(0,73) **
père employé	-0,36	(0,82) n.s.	-0,24	(0,88) n.s.	-0,39	(0,66) n.s.
père cadre	3,74	(1,07) ***	3,98	(1,15) ***	3,18	(0,86) ***
Mère avec emploi qualifié (référence mère au foyer ou non qualifiée)	1,33	(0,62) **	0,51	(0,67) n.s.	0,66	(0,50) n.s.
Elève redoublant (référence élève à l'heure)	-6,80	(0,79) ***	-7,41	(0,83) ***	-6,04	(0,64) ***
Cours multiple (référence cours simple)	0,93	(1,16) n.s.	1,93	(1,74) n.s.	1,63	(1,30) n.s.
Effets aléatoires						
Niveau 2 : variance interclasse	17,35	(4,06)	44,49	(8,48)	25,08	(4,83)
Niveau 1 : variance intraclasse	79,03	(3,61)	89,02	(4,06)	49,23	(2,28)
Pourcentage de variance interclasse expliquée	64,57%		29,61%		49,32%	
Pourcentage de variance intraclasse expliquée	57,40%		48,68%		66,34%	
- 2 log L	7723,4		7885,5		7082,5	

N= 1059 (français) N= 1057 (Mathématiques) N=1030 (score global)

Seuil de significativité des effets fixes : n.s. : non significatif, \* : significatif à .10, \*\* : significatif à .05, \*\*\* : significatif à .01. Les erreurs type des coefficients figurent entre parenthèses.

Concernant l'impact du type de cours fréquenté, on retrouve une tendance identique à celle observée en CE1. La possibilité d'une affectation choisie des élèves par les enseignants s'accompagne d'une réduction, voire d'une disparition dans le cas du CM1, des effets négatifs des cours multiples sur les progressions des élèves. Le français et le score global étaient les deux champs affectés, certes faiblement mais négativement, par la fréquentation d'un cours multiple, et notamment par la configuration CE2-CM1, quand les équipes ne disposaient pas de ce choix. Le tableau 25 montre en outre que ces coefficients sont certes non significatifs mais toujours positifs, et sensiblement plus forts quand ils sont associés à la configuration CE2-CM1, celle-là même qui est la plus néfaste aux progressions quand les élèves y sont placés d'office.

*Tableau 25 : Effets des différentes configurations de type de cours sur les progressions au CM1 (écoles qui offrent le choix, dont un cours double)*

	Français	Mathématiques	Global
Cours multiple avec section inférieure	1,46 (1,37) n.s.	1,69 (2,04) n.s.	2,02 (1,53) n.s.
Cours multiple avec section supérieure	0,50 (1,31) n.s.	2,12 (1,95) n.s.	1,31 (1,46) n.s.

Seuil de significativité des effets fixes : n.s. : non significatif, \* : significatif à .10, \*\* : significatif à .05, \*\*\* : significatif à .01. Les erreurs type des coefficients figurent entre parenthèses.

En conclusion, il apparaît nettement que la fréquentation d'un cours double, en CE1 comme en CM1, n'est jamais plus efficace que la fréquentation d'un cours simple ; elle est au mieux neutre quand les élèves font l'objet d'une affectation délibérée dans ce type de classe. En outre, la configuration des cours doubles éclaire un résultat intéressant. En effet, en CE1, c'est la fréquentation d'un CE1-CE2 qui s'avère la moins favorable en termes de progressions alors qu'en CM1, c'est la configuration CE2-CM1 qui est la plus désavantageuse. Ce constat n'est pas complètement en phase avec les perceptions des enseignants selon lesquelles il est en général plus positif pour les élèves de faire partie de la section inférieure d'un cours double.

Pour expliquer les différences entre ces deux niveaux d'enseignement, on peut évidemment évoquer l'âge des élèves. Ainsi, pour une question de maturité, les élèves de CE1 peuvent être particulièrement pénalisés quand ils sont scolarisés avec des élèves de CE2, pourtant du niveau scolaire supérieur au leur, alors que les CM1 sont quant à eux désavantagés quand ils fréquentent un cours double avec de plus jeunes. On peut aussi évoquer la nature des apprentissages à réaliser, propres à chaque niveau et plus ou moins adaptés à l'organisation spécifique des cours multiples. Les arbitrages réalisés dans ces classes se font à la fois sur le temps consacré par le maître à chacune des sections et sur les priorités accordées aux différentes disciplines (Suchaut, 1996b). Les différences relevées dans nos analyses entre les disciplines (mathématiques et français) en terme d'efficacité pédagogique montrent bien toute l'importance de ces arbitrages qui peuvent avoir une répercussion sur les opportunités d'apprentissage des élèves qui fréquentent un cours double. Chaque configuration de classe est singulière et sans approfondir cette question, on peut néanmoins supposer que ces arbitrages ont des répercussions différentes selon le niveau d'enseignement : ils sont sans doute plus aisés pour l'enseignant et moins risqués pour les élèves dans un CM1-CM2 que dans un CE1-CE2 par exemple, d'autant plus que cette dernière configuration concerne deux cycles d'enseignement différents.

Un second résultat majeur concerne la possibilité qui existe ou non d'affecter intentionnellement les élèves dans les classes à plusieurs cours. Quand cette possibilité n'existe pas dans les écoles, les effets pédagogiques sur les progressions des élèves sont clairement négatifs ; quand cette possibilité existe, alors les progressions des élèves de cours

multiples ne sont pas significativement affectées. Cette distinction « choix-non choix des élèves » introduit donc une nuance sérieuse dans les résultats et elle permet en outre d'alimenter le débat évoqué précédemment concernant les biais de sélection des élèves. Veenman (1996, 1997) nie leur existence tandis que Mason et Burns (1996, 1997) avancent que c'est l'affectation des meilleurs élèves et des plus autonomes d'entre eux qui permet de contrecarrer la moindre efficacité des classes à plusieurs cours. La recherche présentée ici est basée sur une hypothèse instrumentale plus simple car indépendante des critères de sélection potentiellement mobilisés. Elle distingue deux situations seulement : celle dans laquelle il existe plusieurs classes d'un même niveau dans une école, les enseignants pouvant alors choisir quels élèves fréquentent ces classes et celle dans laquelle il n'y a qu'une seule classe, les élèves y étant inscrits d'office. Cette dichotomie permet de dépasser certaines ambiguïtés apparues précédemment dans la composition des classes. En effet, si les directeurs déclaraient effectivement que les élèves les plus autonomes étaient ceux dont on privilégiait l'affectation en cours multiples, les analyses empiriques réalisées ex-post ont montré qu'il existait manifestement une grande diversité des pratiques selon les écoles, certaines y affectant les élèves les meilleurs, d'autres, les élèves les plus faibles. Ce qui importe donc dans ces résultats, c'est d'observer que les choix opérés, quels qu'ils soient, sont efficaces puisqu'ils permettent de neutraliser les effets négatifs des classes à plusieurs cours ; autrement dit, les enseignants mobilisent les « bons » critères d'affectation et apprécient avec pertinence les qualités requises pour « profiter » ou ne pas pâtir d'une scolarisation dans ce contexte particulier.

Finalement, les résultats dégagés malmènent ce qu'on pensait être une spécificité française. On pourrait donc légitimement s'interroger sur leur robustesse et leur sensibilité au type d'épreuves utilisées pour mesurer les acquisitions des élèves (Leamer, 1983). A ce titre, des sélections aléatoires d'items ont été réalisées pour tester successivement plusieurs représentations du niveau des élèves : quelle que soit la configuration d'items mobilisée, les estimations associées à la variable « cours double » sont stables et apparaissent dans tous les cas, des différences entre le français et les mathématiques.

Par conséquent, on peut évoquer d'autres éléments pour expliquer ces résultats. En premier lieu, le contexte a évolué depuis les conclusions positives des premiers travaux consacrés aux petites écoles rurales et l'évaluation n'a porté ici que sur l'efficacité pédagogique des cours doubles, configuration dorénavant la plus courante. En second lieu, les progrès réalisés en matière de modélisation statistique laissent penser que certains effets jugés positifs hier pourraient ne plus l'être aujourd'hui. L'une des conclusions encourageantes de cette recherche reste que l'affectation intentionnelle des élèves dans les classes à plusieurs cours pourrait permettre d'en neutraliser les effets négatifs. Evidemment, il reste délicat d'en tirer des enseignements en terme d'action dans la mesure où les écoles n'ont pas toujours cette

possibilité. Nul doute que des évaluations complémentaires sur cette question seront les bienvenues pour alimenter le débat et les politiques éducatives en la matière.

En complémentarité avec la diversité des organisations pédagogiques présentes dans les classes dont nous venons en partie de tester l'efficacité, nous allons à présent nous intéresser à la seconde dimension du contexte scolaire retenue pour ce chapitre, à savoir la composition sociale du groupe classe.

### **III. L'influence de la composition sociale**

La question de l'influence du contexte sur les comportements des individus est, à des niveaux d'analyse différents, au cœur des problématiques de la sociologie et de la psychologie sociale. Dans le domaine scolaire, cette dernière va prendre en compte à la fois « l'état interne et l'environnement du sujet » en admettant que « les activités intellectuelles et leurs produits peuvent dépendre, et parfois de façon majeure, du rapport que la personne entretient... avec la situation dans laquelle elle se trouve et de la nature de la tâche qu'elle accomplit » (Monteil, Huguet, 2002). Une attention particulière sera alors portée à des paramètres tels que la composition du groupe classe au sein duquel des processus de comparaison sociale vont prendre place, ou encore les modalités des interactions pédagogiques (valorisation, par le maître, de la compétition entre élèves, par exemple). De manière parallèle, et sans que des échanges systématiques ne se soient développés avec les psychologues sociaux, les sociologues travaillent très normalement ce type de question. En effet, la question de l'influence des caractéristiques du contexte scolaire sur les élèves (leurs acquis, leurs attitudes) est une déclinaison de la question sociologique très générale de l'influence de la structure sociale sur les conduites des individus.

On trouve ce type d'interrogation chez des fondateurs de la discipline comme Durkheim, qui, dans son analyse du Suicide (1897), faisait des hypothèses sur les processus liant certaines propriétés de la structure familiale et les conduites des individus. Dans les années 40-50, plusieurs sociologues américains, notamment Merton (1938), reprennent cette problématique pour expliquer, grâce à certaines caractéristiques de la structure sociale, l'émergence de la déviance. On souligne alors que les individus évoluent dans des ensembles organisés de relations sociales leur offrant des structures d'opportunité, affectant les probabilités de développer telle ou telle conduite et imposant telle définition prégnante de la situation. Aujourd'hui, cette perspective est très développée dans la recherche anglo-saxonne, même si l'on s'en tient aux thématiques éducatives (Bressoux, 1995 ; Duru-Bellat, 2003b). En particulier, tout un courant extrêmement vivace étudie l'impact de cette dimension du contexte que constitue l'établissement fréquenté (courant dit du « school effectiveness »),

tandis qu'une masse de travaux qui peuvent relever de la psychologie sociale ou de la sociologie se centre sur l'effet, sur les acquisitions et les attitudes, de la composition sociale ou scolaire du contexte (ce que les anglo-saxons désignent sous le vocable de *school mix*), avec par exemple, dans les pays anglo-saxons où ce mode de groupement des élèves est répandu, des analyses des effets des classes de niveau.

Dans une recherche récente (Duru-Bellat, Landrier-Le Bastard, Piquée, Suchaut, 2004), sur laquelle s'appuie en partie cette section, c'est l'influence de l'hétérogénéité/homogénéité sociale du contexte qui a constitué la variable cible de l'analyse. Les données empiriques, sont issues de la même base de données mobilisée dans le premier point de ce même chapitre ; elles portent sur 91 classes de CE1 et 86 classes de CM1 de l'Académie de Dijon (avec environ 1700 élèves à chacun de ces niveaux) ; elles permettent d'engager une analyse quantitative fiable des effets du "school mix" sur les progressions, dans la mesure où des informations sur le niveau scolaire initial et sur le niveau final ont été engrangées.

L'ensemble des travaux empiriques existants à ce jour dans la recherche en Education à orientation sociologique produit sans conteste des connaissances mais expose aussi à de nombreux problèmes méthodologiques que l'on ne peut ignorer. Notre démarche exposera les différentes étapes de l'analyse, allant du constat de différences entre contextes aux modélisations multiniveaux, puis elle abordera rapidement l'apport d'approches plus qualitatives, avant de conclure sur l'enjeu inhérent à ces débats méthodologiques, vu l'importance de la portée politique des résultats en de domaine.

### *III.1 Des différences globales entre contextes*

La variable centrale aux analyses qui vont suivre (la « tonalité sociale » du contexte étudié), pose des problèmes d'opérationnalisation, notamment sur quelles bases est-on fondé à parler de tonalité sociale populaire ou favorisée ? Cette question est très documentée dans la littérature sur la ségrégation sociale (cf. Gorard et Taylor, 2002), en ce qu'elle n'est évidemment pas seulement « technique » mais implique une spécification de ce qu'on entend par ségrégation. On s'accorde pour considérer comme centrale dans cette notion ce qu'on appelle l'équi-répartition, à savoir l'homogénéité de la représentation d'un groupe dans les différentes unités étudiées. Il est alors possible d'estimer la distance entre la situation observée et une situation d'équi-répartition, avec des indicateurs comme le « pourcentage d'élèves à déplacer » pour atteindre cette situation. Mais si cet indicateur très classique chiffre le degré global de ségrégation –telle sous-population présente des caractéristiques sociales très typées par rapport à l'ensemble-, il ne dit rien du ou des groupes qui sont sur ou sous représentés, et deux établissements dont l'indicateur revêt la même valeur peuvent avoir des profils sociologiques très différents. Dans la recherche présentée ici, après avoir testé

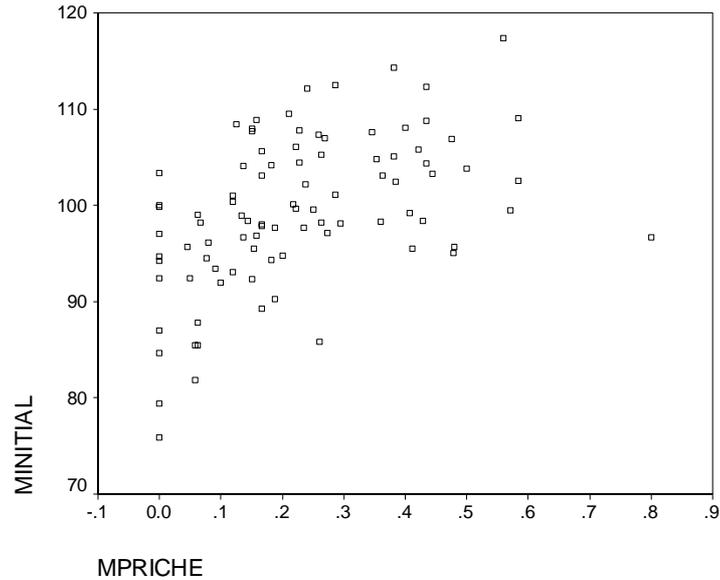
plusieurs de ces indicateurs classiques de ségrégation, nous nous sommes ralliés à des mesures plus élémentaires telles que le pourcentage d'élèves (dans la classe ou l'établissement) appartenant à telle ou telle catégorie sociale. Néanmoins, dans la mesure où les travaux existants invitaient à nous méfier d'indicateurs continus<sup>32</sup>, nous avons construit plusieurs typologies alternatives pour caractériser la tonalité sociale de la classe, à partir de l'examen des distributions. C'est ainsi que nous avons utilisé dans certaines analyses des nomenclatures spécifiques opposant des classes dites favorisées (où le pourcentage d'enfants dont le père est cadre ou profession intermédiaire est supérieur à un certain seuil<sup>33</sup>), et des classes dites défavorisées, avec des résultats qui peuvent différer selon le mode d'opérationnalisation retenu.

Dans cette section, nous nous référerons principalement à la variable nommée « mpriche », soit la somme du pourcentage d'enfants de cadres et du pourcentage d'enfants des professions intermédiaires (par exemple, un chiffre de 70% signifie que l'on a dans la classe 70% d'enfants dont le père est cadre ou de profession intermédiaire). Cette façon d'appréhender la tonalité sociale est plus intéressante que le pourcentage d'enfants n'appartenant pas à des catégories défavorisées (ouvriers, autres et inconnues, soit la variable « mpauvre », en ce sens qu'elle fait apparaître une plus grande discrimination entre les classes. C'est sur cette base que nous pouvons produire une première description des différences entre classes, au niveau du primaire, sur les données collectées. Un premier constat, que fait apparaître le graphique ci – après est l'existence d'une relation forte entre le niveau initial moyen des classes et leur tonalité sociale appréhendée par le pourcentage d'élèves de milieu favorisé (mpriche).

---

<sup>32</sup> Dans sa récente étude sur la ségrégation ethnique dans les collèges, Felouzis (2003) montre ainsi que la variable « pourcentage d'élèves allochtones dans l'établissement » ne joue qu'au delà d'un certain seuil. De manière générale, les chercheurs anglo-saxons (cf. par exemple Thrupp, 1999) insistent, en matière de *school mix*, sur l'importance de la « *critical mass* », c'est à dire du seuil à partir duquel la culture du groupe classe bascule du côté de la culture des classes moyennes, *a priori* plus conforme à la culture scolaire, ou au contraire du côté d'une culture populaire qui en est plus éloignée, si ce n'est en opposition.

<sup>33</sup> Aux deux niveaux, le codage des professions des parents s'est fait suivant les nomenclatures de la DEP-Ministère de l'Education, distinguant un groupe « favorisé » (professions intermédiaires, cadres supérieurs et professions libérales), un groupe intermédiaire regroupant les agriculteurs, les commerçants, les employés, et un groupe dit défavorisé (ouvriers et inactifs). Bien qu'il soit évidemment préférable de tenir compte des professions des deux parents, l'existence d'une proportion importante de mères « inactives » amène souvent à privilégier la seule profession paternelle.



*Graphique 15 : Niveau initial des classes et proportion d'élèves de milieu favorisé, en CE1.*

On observe tout d'abord qu'il existe des disparités très importantes de niveau initial moyen entre les classes, dès ce niveau, pourtant précoce, de la scolarité. De plus, ces fortes différences d'une classe à l'autre apparaissent, dans ce type d'approche transversale, comme corrélées avec la tonalité sociale de la classe. Les situations sont extrêmement contrastées entre les classes les plus typées socialement. Ainsi, en CE1, les 10 classes dans lesquelles le pourcentage d'enfants de cadres ou de professions intermédiaires (soit mpriche) est le plus fort (soit plus de 44%) ont un score moyen de 103 alors que les 10 classes dans lesquelles le pourcentage d'enfants d'employés, d'ouvriers ou d'inactifs est le plus fort (supérieur à 71%) ont un score moyen de 92, soit un écart de 11 points (dans une distribution normalisée où la moyenne est 100 et l'écart-type est de 15). L'écart est du même ordre de grandeur au CM1, où les 10 classes les plus privilégiées ont un score moyen de 105 alors que les 10 classes les plus défavorisées ont un score moyen de 95.

Une manière de chiffrer cette relation consiste à introduire dans une analyse de régression expliquant le score moyen des classes, leur tonalité sociale comme variable explicative (cf. tableau 26). Au niveau du CE1 (91 classes), le pourcentage d'élèves de milieu favorisé (mpriche) explique 24% de la variance du niveau initial moyen, et une proportion légèrement supérieure au niveau CM1 (86 classes).

Tableau 26 : Relation entre score initial et tonalité sociale de la classe au niveau agrégé

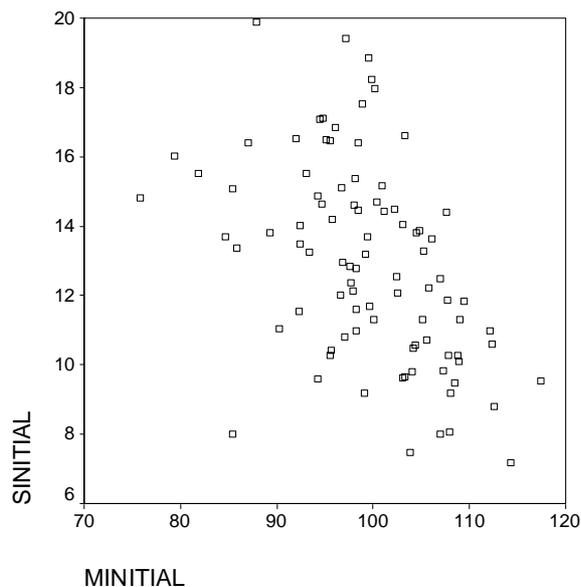
Variables	CE1		CM1	
	Coef.	t	Coef.	t
% d'enfants de cadres ou de profession intermédiaires	0,22	***	0,23	***
Constante	94,30	***.	94,25	***
R <sup>2</sup>	0,24		0,26	

\*\*\* : significatif au seuil de 1%, \*\* : significatif au seuil de 5%, \* : significatif au seuil de 10%,  
n.s. : non significatif

A ce propos, la question se pose de l'interprétation du pourcentage de variance expliquée dans un modèle, qui varie dans l'absolu de 0 à 100%. Si les écoles ou les classes fréquentées expliquaient 100% de la variance des scores, cela voudrait dire qu'elles sont à l'origine de l'intégralité de la dispersion des performances des élèves mais cela pourrait aussi vouloir dire qu'elles n'ont aucun effet en tant que telles. Ce serait le cas si les écoles ou les classes étaient complètement homogènes en terme de composition de leur public, les élèves étant semblables au sein des écoles (ou classes) et différents d'une école à l'autre. Dans ce cas, même si les écoles n'ont en elles-mêmes aucun effet sur la réussite des élèves, les caractéristiques des écoles prédiraient complètement les inégalités de performances des élèves, ces effets étant attribués de façon inexacte aux écoles alors qu'ils ne refléteraient que les caractéristiques sociales et scolaires de leur public (Sorensen, Morgan, 2000).

Il importe de pouvoir distinguer ces deux interprétations opposées. La solution idéale serait de recourir à une expérimentation où les élèves seraient affectés de façon aléatoire dans les classes ou les écoles. Ceci est évidemment impossible à mettre en œuvre mais certaines méthodes statistiques peuvent être considérées comme un substitut à cette approche expérimentale, sachant qu'une théorie serait à cet égard bien commode mais le champ de la « school effectiveness » est relativement démuné à cet égard.

Notons que la tonalité sociale s'avère par ailleurs également liée à d'autres caractéristiques du groupe classe. Ainsi, il existe une corrélation négative entre le pourcentage d'élèves d'origine favorisée et le pourcentage d'élèves ayant redoublé (la corrélation est de -0,40 au CE1 et de -0,48 au CM1, ces corrélations étant significatives à .01). Il existe aussi des relations entre le niveau initial des classes et leur degré d'hétérogénéité (cf. graphique 16 ci-après).



*Graphique 16 : Niveau initial moyen des classes et degré d'hétérogénéité, au niveau CE1.*

Au niveau du CE1, on observe une relation négative entre le degré d'hétérogénéité de la classe (eu égard au niveau initial) et son niveau initial moyen ( $r = -0,47$ , significatif à  $.01$ ). C'est dire qu'en moyenne, les classes sont d'autant moins hétérogènes qu'elles sont de bon niveau moyen ; on peut noter à ce propos que les discours qui déplorent l'hétérogénéité des classes seraient donc « contaminés » par le fait qu'ils concernent le plus souvent des classes faibles. Par conséquent, il y a aussi une tendance (mais la corrélation n'est pas significative) à ce que l'hétérogénéité du niveau initial moyen soit plus faible dans les classes où il a plus d'élèves de milieu favorisé, puisque ces classes sont de niveau moyen plus élevé. Cette relation est également manifeste au niveau du CM1, mais elle est plus faible ( $r = -0,27$ , significatif à  $.01$ , donc plus faible qu'en CE1). Par conséquent, il n'y a pas de relation entre l'hétérogénéité du niveau initial moyen et le pourcentage de « riches » dans la classe au niveau du CM1.

Ces différences moyennes observées de manière transversale invitent à aller plus loin. Elles sont intéressantes et irremplaçables dans la mesure où elles traduisent une réalité forte, celle de contextes d'apprentissage des élèves extrêmement différenciés. Ainsi, pour les cas extrêmes, les différences de scores moyens entre classes sont supérieures à deux écart-types. Les élèves partagent donc bien un environnement très typé au sein d'une même classe. Mais, ces analyses conduites au niveau agrégé ne permettent pas de distinguer ce qui tient aux caractéristiques des élèves, agrégées dans un contexte et produisant ce qu'on appelle des effets de composition, de ce qui relève de phénomènes qui seraient produits spécifiquement au niveau du groupe classe. En d'autres termes, ce niveau d'analyse ne permet pas d'identifier à proprement parler des effets de contexte qui résulteraient spécifiquement du fait d'être

scolarisés dans un contexte ségrégué et expliquant, au moins pour une part, les inégalités de performances constatées entre les classes. De plus, dans la mesure où la prise d'information est instantanée (en début d'année scolaire), on ne sait rien de la genèse temporelle du phénomène étudié, en l'occurrence, du rôle respectif du milieu familial et des premières années d'école.

De fait, les contraintes de temps et de ressources, et plus encore la difficulté à recueillir sur le terrain des informations sur un sujet aussi sensible que les acquisitions « objectives » des élèves, amènent souvent à monter des dispositifs imparfaits (ils le sont toujours comparativement aux dispositifs expérimentaux que construisent de toutes pièces les psychologues). En ce qui nous concerne, et pour articuler une approche quantitative et une exploration qualitative ciblée, nous avons retenu une solution consistant, au niveau de l'école primaire, à coupler une évaluation (quantifiée) de l'effet établissement et de l'effet du school mix sur des données existantes (recueillies dans le cadre d'une précédente recherche) et à collecter spécifiquement des données sur les processus sur un nombre plus réduit d'établissements. Il eut bien sûr été préférable de monter un dispositif spécifique unique, et les données secondaires ne sont jamais parfaitement adaptées à la question précise que l'on se pose. Ces données permettent néanmoins d'évaluer d'éventuels effets du school mix sur les progressions et certaines attitudes des élèves sur une base fiable et sur des échantillons raisonnables, notamment au niveau de l'école primaire. Elles permettent également de croiser la composition sociale des publics des établissements avec une variété de processus, organisationnels, pédagogiques ou psycho-sociaux (de manière toutefois assez distale), susceptibles d'en comprendre le mode d'action sur les résultats et les attitudes des élèves.

### ***III.2 Les modèles de progressions des élèves au niveau agrégé***

Pour évaluer des effets de contexte, une exigence est, nous l'avons vu, d'adopter une approche longitudinale, avec au moins deux phases d'observation. Le point de départ qui peut apparaître comme le plus immédiat consiste à estimer des modèles de progressions au niveau agrégé, en prenant comme unité d'analyse la classe ou l'école et en expliquant les niveaux d'acquisitions des élèves en fin d'année scolaire par leur niveau initial moyen tel qu'agrégé au niveau des classes. Les modèles agrégés de ce type estimés sur la base de nos données font apparaître une relation forte entre acquis initiaux et acquis finaux, ce qui atteste du caractère cumulatif des acquis scolaires ; mais on remarque qu'en CE1, la liaison entre initial et final est plus forte qu'en CM1 (tableau 27).

*Tableau 27 : Progressions des élèves  
(explication du score final par le score initial) au niveau agrégé*

Variables	CE1		CM1	
	Coef.	t	Coef.	t
Niveau initial moyen de la classe	0,74	***	0,69	***
Constante	26,13	***.	31,33	***
R <sup>2</sup>	0,54		0,43	

\*\*\* : significatif au seuil de 1%, \*\* : significatif au seuil de 5%, \* : significatif au seuil de 10%,  
n.s. : non significatif

Dans les deux cas, la relation entre score initial moyen et score final moyen des classes est linéaire (la linéarité de la relation a été vérifiée) : le score final d'une classe est d'autant plus élevé que l'est le niveau initial moyen de ses élèves. Ceci signifie que les progressions au niveau des classes ne revêtent pas une allure différenciatrice, c'est-à-dire que les classes faibles et fortes progressent de la même façon

Les modèles ci-après (tableau 28) montrent également que la prise en compte de la tonalité sociale de la classe apporte un gain de variance de 2% pour les deux niveaux scolaires considérés. Les progressions sont donc d'autant plus fortes que le pourcentage d'enfants de milieu favorisé est élevé, dans la classe, sachant que l'influence du pourcentage d'élèves favorisés (variable mpriche) est moins forte en CE1 qu'en CM1 (le pouvoir explicatif global du modèle étant plus fort au CE1). Un des apports précieux des modèles est qu'ils permettent d'effectuer des simulations. Ainsi, en CE1, pour 10% de plus d'élèves de milieu favorisé dans la classe, le score moyen final de la classe est de 0,7 points de plus à niveau initial moyen donné. Par conséquent, quand on passe de 10% à 70% de « mpriche », l'augmentation de score moyen est de 4,2 points. En CM1, pour 10% de « mpriche » supplémentaire, le gain est de 0,9 points à niveau initial moyen donné. A ce niveau, quand on passe de 10% à 70% de « mpriche », l'augmentation de score moyen de 5,4 points.

*Tableau 28 : Modèles agrégés de progressions intégrant la tonalité sociale de la classe*

Variables	CE1		CM1	
	Coef.	t	Coef.	t
Niveau initial moyen	0,67	***	0,59	***
% d'enfants de cadres ou de profession intermédiaires	0,07	*	0,09	**
Constante	31,43	***.	38,81	***
R <sup>2</sup>	0,56		0,45	

\*\*\* : significatif au seuil de 1%, \*\* : significatif au seuil de 5%, \* : significatif au seuil de 10%,  
n.s. : non significatif

Il s'avère par ailleurs que, au CE1, la forme quadratique est plus adaptée à la relation : au-delà d'un certain seuil de « mpriche » l'effet sur les progressions moyennes des classes devient moins important<sup>34</sup>. On peut également intégrer dans l'analyse le degré d'hétérogénéité du niveau initial moyen de la classe (tableau 29). On observe alors un impact négatif de cette variable : plus la classe est hétérogène en début d'année, moins les progressions sont bonnes ; cet impact est plus fort en CM1 qu'en CE1, et les coefficients de « mpriche » ne changent pas par rapport aux modèles précédents.

*Tableau 29 : Modèles agrégés de progressions intégrant la tonalité sociale de la classe, le niveau moyen et l'hétérogénéité*

Variables	CE1		CM1	
	Coef.	t	Coef.	t
Niveau initial moyen	0,60	***	0,53	***
% d'enfants de cadres ou de profession intermédiaires	0,07	**	0,09	**
Ecart-type du niveau initial moyen	-0,36	*	-0,48	**
Constante	43,00	***	51,21	***
R <sup>2</sup>	0,57		0,48	

\*\*\* : significatif au seuil de 1%, \*\* : significatif au seuil de 5%, \* : significatif au seuil de 10%, n.s. : non significatif

Mais cette relation reste difficile à interpréter ; d'une part parce que dans la réalité les classes hétérogènes sont plus fréquemment de moins bonnes classes (toutes les choses ne sont pas égales par ailleurs comme le voudrait la démarche abstraite de modélisation) ; d'autre part parce qu'il est grossier, comme le montrent certains travaux, d'opposer des classes hétérogènes à des classes plus homogènes qui peuvent être homogènes-fortes ou homogènes-faibles, soit deux situations aux effets très différents, respectivement positifs et négatifs sur les apprentissages, du moins au niveau du collège (Duru-Bellat, Mingat, 1997).

Pour approfondir cette question nous avons construit des variables qui permettent de mesurer l'effet conjugué du niveau moyen initial de la classe et de son degré d'hétérogénéité<sup>35</sup>. Les estimations font apparaître qu'effectivement l'effet de l'hétérogénéité n'est pas le même selon le niveau initial de la classe. Aux deux niveaux scolaires étudiés, globalement l'hétérogénéité s'avère moins favorable aux progressions que l'homogénéité, ceci est vrai particulièrement pour les classes faibles alors que les classes fortes, qu'elles soit homogènes ou hétérogènes

<sup>34</sup> Notons que si, plutôt que de travailler au niveau agrégé de la classe, on se situe au niveau écoles (62 écoles dans lesquelles on a au moins 1 CE1 et 1 CM1), les résultats sont comparables ; le modèle de progression (score final=score initial, mpriche), explique 53% de la variance du score moyen final par école, le coefficient de « mpriche » vaut 0,07)

<sup>35</sup> Deux spécificités alternatives ont été testées. La première permet de comparer quatre situations entre elles (classes hétérogènes faibles, classes hétérogènes fortes, classes homogènes faibles et classes homogènes fortes) ; la seconde spécification prend en compte deux variables continues rendant compte du degré d'hétérogénéité pour les classes fortes et pour les classes faibles.

(mais elles le sont plus rarement) progressent davantage. Mais il faut souligner que ces résultats sont établis au niveau agrégés : ils traduisent donc les situations globales des classes mais ne disent rien sur les progressions individuelles des élèves de ces classes. Les situations constatées peuvent en fait renvoyer à des cas de figure très différents : un « noyau » fort d'élèves qui tire la moyenne vers le haut, quelques élèves avec des progressions atypiques, une forte progression moyenne de toute la classe... On ne saurait donc tirer de ces premiers résultats sur des modèles agrégés des conclusions en terme d'effets de l'hétérogénéité des classes (sur les progressions des élèves), au niveau individuel donc, pas plus que sur l'impact de leur niveau moyen, sur la progression d'élèves donnés.

En résumé, au niveau agrégé, on est donc bien en présence de classes différentes, à la fois en terme de composition sociale, d'hétérogénéité et de niveau moyen. Et de façon globale, les progressions moyennes des classes sont d'autant plus fortes que les classes sont homogènes (en terme de niveau initial), de composition sociale favorisée et de niveau initial élevé. Mais avec les modèles agrégés, on ne peut séparer ce qui tient à la dynamique propre au groupe d'élèves et ce qui serait donc un phénomène contextuel, de ce qui tient aux élèves eux-mêmes, et découle de leur agrégation. Car il peut exister ce que l'on désigne dans la littérature sous le terme de « biais d'agrégation », à savoir que des « analyses portant sur des données agrégées peuvent conduire à des résultats notablement différents de ceux portant sur des données individuelles, voir inverses » (Bressoux, Coustère, Leroy-Audouin, 1997). Autrement dit, on ne peut déduire d'une relation établie au niveau agrégé une relation correspondante au niveau des individus.

Une autre limite de l'agrégation des données est qu'il est impossible d'étudier la dimension de l'équité (soit la réduction des écarts d'acquisitions entre les élèves en cours de l'année scolaire), en effet dans les modèles agrégés, chaque classe est caractérisée par un score final moyen et la seule estimation possible est la fonction de transformation moyenne entre le niveau initial moyen et le score final. Or la recherche d'interaction entre un « effet » moyen et les caractéristiques des personnes concernées est aujourd'hui une démarche de routine dans la recherche, et particulièrement dans le domaine éducatif.

### ***III.3 Les modèles multiniveau intégrant variables individuelles et contextuelles***

Pour surmonter ces deux difficultés et dégager véritablement un éventuel effet du contexte, il faut introduire simultanément dans les modèles variables individuelles et variables contextuelles avec des modèles multiniveaux. Rappelons que le principal intérêt de ces modèles est de permettre d'analyser des données hiérarchisées ce qui est particulièrement intéressant dans l'étude des phénomènes ayant trait à l'éducation. Ici, deux niveaux d'analyse sont présents : les élèves (niveau individuel) et les classes (niveau contextuel ou

environnemental). L'objectif est alors de pouvoir distinguer dans l'explication du phénomène étudié, en l'occurrence ici les progressions des élèves, et cela dans une même phase d'analyse, ce qui tient d'une part aux caractéristiques individuelles des élèves (origine sociale, sexe, niveau initial..) et d'autre part aux caractéristiques de la classe (niveau initial moyen, tonalité sociale...). Nous avons estimé ce type de modèle aux deux niveaux scolaires étudiés (tableaux 30 et 31).

Tableau 30 : Modèles multiniveaux de progression en CE1

Paramètres	Modèle A (vide)		Modèle B		Modèle C	
Effets fixes						
Constante	99,59	(0,82) ***	29,21	(1,76) ***	29,49	(1,76) ***
Score individuel initial			0,70	(0,02) ***	0,70	(0,02) ***
Sexe : Fille (référence garçon)			0,96	(0,45) **	0,97	(0,45) **
Profession du père : père cadre ou profession intermédiaire (référence autres profession du père)			1,27	(0,56) **	1,13	(0,57) **
% d'élèves dans la classe dont le père ou la mère sont cadres ou professions intermédiaires					0,99	(0,55) *
Effets aléatoires						
Niveau 2 : variance interclasse	51,76	(9,13)	22,78	(4,11)	22,14	(4,03)
Niveau 1 : variance intraclasse	174,95	(5,98)	87,01	(2,98)	87,01	(2,98)
Pourcentage de variance interclasse expliquée	22,83 %		56,0%		57,2%	
Pourcentage de variance intraclasse expliquée	77,17 %		50,0 %		52,3%	
- 2 log L	14591,5		13326,8		13323,0	

N = 1802

Seuil de significativité des effets fixes : n.s. : non significatif, \* : significatif au seuil de 10% , \*\* : significatif au seuil de 5%, \*\*\* : significatif au seuil de 1%. Les erreurs type des coefficients figurent entre parenthèses

Tableau 31 : Modèles multiniveaux de progression en CMI

Paramètres	Modèle D (vide)		Modèle E		Modèle F	
Effets fixes						
Constante	99,91	(0,75) ***	21,88	(1,54) ***	21,95	(1,55) ***
Score individuel initial			0,77	(0,01) ***	0,77	(0,01) ***
Sexe : Fille (référence garçon)			0,67	(0,40) *	0,68	(0,40) *
Profession du père : père cadre ou profession intermédiaire (référence autres profession du père)			3,04	(0,49) ***	3,01	(0,49) ***
% d'élèves dans la classe dont le père ou la mère sont cadres ou professions intermédiaires					0,37	(0,57) n.s.
Effets aléatoires						
Niveau 2 : variance interclasse	38,33	(7,34)	23,89	(4,18)	24,06	(4,23)
Niveau 1 : variance intraclasse	186,83	(6,55)	64,76	(2,27)	64,77	(2,27)
Pourcentage de variance interclasse expliquée	17,02%		37,67		37,23	
Pourcentage de variance intraclasse expliquée	82,98%		65,34		65,33	
- 2 log L	13965,3		12195,6		12194,5	

N = 1714

Seuil de significativité des effets fixes : n.s. : non significatif, \* : significatif au seuil de 10% , \*\* : significatif au seuil de 5%, \*\*\* : significatif au seuil de 1%. Les erreurs type des coefficients figurent entre parenthèses.

Dans un premier temps, un modèle dit vide (ou inconditionnel), n'incluant aucune variable explicative, produit une décomposition de la variance totale des acquis scolaires en fin d'année en une part de variance inter-classes et une part de variance intra-classe (modèle A). Au niveau du CE1, la part de la variance totale expliquée par la variance entre les classes, qui est ici de 22,8%, donne une évaluation de la ressemblance entre les élèves d'une même classe. La prise en compte de ce « niveau de contexte » s'avère donc justifiée et les modèles suggèrent l'existence d'un effet contextuel. Dans le modèle C, l'ajout de la variable tonalité sociale (% de pères et mère favorisés) procure un accroissement de 1,2% de la variance interclasse. Le coefficient est de +0,99 (significatif au seuil de 7%), soit 3 points de progressions supplémentaires quand on passe de 20 à 80% d'élèves favorisés dans une classe<sup>36</sup>.

Une forme non linéaire de la variable « riches » a été testée, la forme quadratique s'avère mieux adaptée que la spécification linéaire, les coefficients sont les suivants pour les deux termes de la variable (tableau 32).

*Tableau 32 : Effets de la non linéarité de la tonalité sociale sur les progressions au CE1*

	Coefficient	Erreur-type	t	Sign.
% de « riches »	1,74	0,61	2,83	0,00***
% de « riches » au carré	-1,12	0,44	-2,55	0,01***

L'effet de la variable « riches » est positif jusqu'à un certain seuil, environ 60% d'élèves de milieu favorisé dans la classe, puis on relève au-delà un impact marginal décroissant dont atteste le coefficient négatif de la variable au carré.

Au niveau du CM1, le modèle vide (modèle D), qui donne une évaluation de la ressemblance entre les élèves d'une même classe, explique 17% de la variance. Si à nouveau la prise en compte de ce « niveau de contexte » s'avère justifiée, puisqu'est ainsi suggérée l'existence d'un effet contextuel, le pourcentage de variance interclasse s'avère à nouveau plus faible qu'en CE1 : il se confirme qu'il y a moins d'influence globale du groupe classe sur les progressions des élèves. On constate une influence plus forte de l'origine sociale en CM1 qu'en CE1, mais pas d'effet d'une tonalité sociale favorisée (% d'élèves dans la classe dont le père ou la mère sont cadres ou professions intermédiaires) sur les progressions.

Nous avons ensuite intégré la variable qui rend compte du niveau moyen de la classe au test initial, celui-ci ne joue pas de façon significative sur les progressions des élèves. Au niveau CM1 on détecte même une tendance négative de cette variable. Rappelons que dans les

<sup>36</sup> La variable qui rend compte de la tonalité sociale a été centrée réduite (moyenne 0 et écart-type de 1).

modèles agrégés celle-ci était corrélée positivement aux progressions dans les modèles individuels qui tiennent compte des caractéristiques des élèves (origine sociale et niveau initial). Notons que cette tendance à un effet négatif du niveau initial de la classe est contraire à ce qui est observé jusqu'alors dans la littérature sociologique au niveau du secondaire (Fraine, Van Damme, Van Landeghem, Opdenakker, Onghena, 2003). À titre d'hypothèse, on peut suggérer que quel que soit le niveau de la classe, les objectifs à atteindre en terme de connaissances scolaires sont davantage bornés en primaire qu'en secondaire ; les classes déjà fortes ne peuvent alors progresser au-delà des acquis élémentaires requis à ce niveau, alors que dans le secondaire, les enseignants ont des possibilités d'approfondissements moins bornées, qu'ils exploiteraient à fond avec les « bonnes » classes. Mais cette question reste aujourd'hui des plus ouvertes : la psychologie sociale, qui insiste sur les processus de comparaisons sociale au sein du groupe classe, prévoit à l'inverse que les progressions seront amoindries dans les contextes forts, du fait d'une moindre estime de soi des élèves ; Marsh et Hau (2003) démontrent ainsi que l'estime de soi (scolaire) est d'autant plus faible que la qualité académique de l'établissement est élevée.

On peut également observer en ce qui concerne le CMI que la tonalité sociale joue significativement après l'introduction de la variable rendant compte du niveau initial moyen de la classe : les progressions sont d'autant meilleures que le pourcentage d'élèves de milieu social favorisé est élevé. Il y a donc bien un phénomène spécifique d'articulation entre les caractéristiques du public d'élèves : on progresse plus dans les classes faibles et dans les classes à public favorisé sachant que dans la réalité ces deux caractéristiques varient en sens inverse. Ce cas de figure rappelle que l'estimation de l'impact d'une variable est toujours contingente à l'ensemble des variables introduites dans les modèles. Ainsi, l'influence d'une variable peut être sur ou sous estimée selon que l'on a omis de tenir compte d'une autre variable dont les effets viendraient contrarier ou au contraire confortés ceux de la première.

#### ***III.4 Les effets d'interaction***

Une question est d'examiner l'articulation qui existe entre les effets de l'origine sociale des élèves et la composition sociale du groupe qu'ils fréquentent. La littérature montre en effet que, de manière générale, les effets de contexte sont plus marqués chez les élèves les plus faibles et/ou de milieu social défavorisé (Duru-Bellat, 2003). Pour cela on peut mesurer l'effet de la tonalité sociale de la classe sur des élèves d'origine sociale différente.

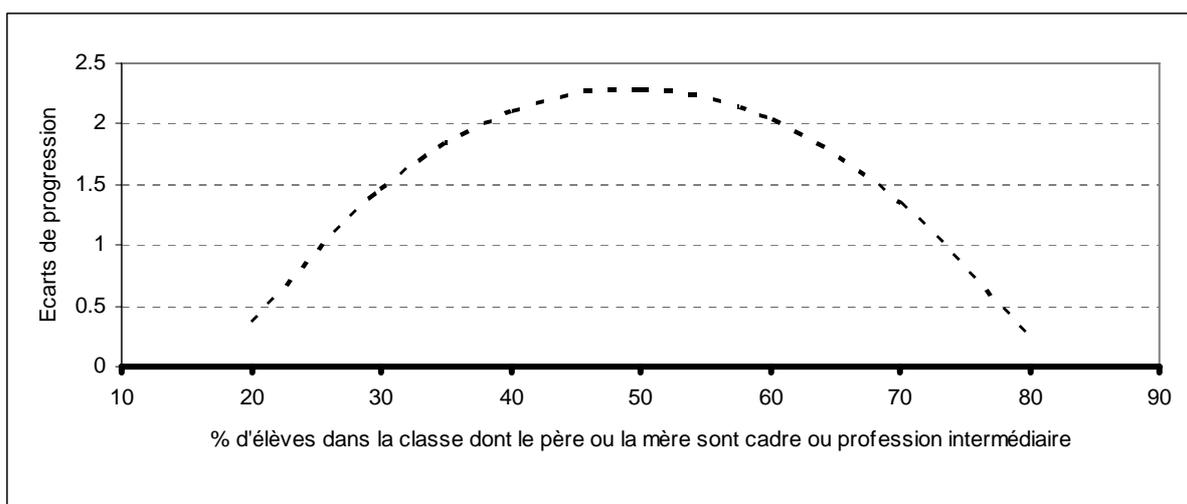
Cela exige de construire et d'intégrer dans les modèles deux variables d'interaction : « riche » (% d'élèves dans la classe dont le père ou la mère sont cadres ou professions intermédiaires pour les élèves dont le père est cadre ou prof intermédiaire) et « richep » (% d'élèves dans la classe dont le père ou la mère sont cadres ou professions intermédiaires pour

les élèves dont le père n'est pas cadre ni de prof intermédiaire). Les résultats, toujours établis sur la base de modèles multiniveaux montrent au niveau du CE1 un effet différencié de la tonalité sociale : seuls les élèves d'origine sociale modeste (ni cadre ni profession intermédiaire) profitent d'une composition sociale de la classe globalement favorisée. Ceci permet de comprendre les résultats détectés précédemment, à savoir l'atténuation des écarts de progressions en fonction de l'origine sociale quand on prend en compte la tonalité sociale de la classe : les élèves de milieu défavorisé comblent d'autant plus leur retard qu'ils fréquentent une classe à tonalité favorisée. Au CM1 on ne note par contre pas d'effet différencié de la composition sociale de la classe selon l'origine sociale de l'élève.

*Tableau 33 : Effets des variables d'interaction entre variables d'origine sociale et de tonalité sociale de la classe dans les modèles de progression au CE1*

	Coefficient	Erreur-type	t	Sign.
Richer	+1,07	3,37	0,32	0,75 n.s.
Richep	+5,94	2,71	2,19	0,03 **

Il est intéressant, de façon complémentaire, de voir si l'effet contextuel de la composition sociale de la classe détecté pour les élèves d'origine modeste, est susceptible de combler le handicap tenant à leur propre origine sociale sur les progressions, donc s'il y a en quelque sorte compensation entre le poids des facteurs individuels et des facteurs contextuels au cours d'une année scolaire donnée. Des simulations de nature comparative effectuées entre des enfants de cadres ou de profession intermédiaires et des enfants de milieu social plus modeste scolarisés dans des classes à composition sociale variées mettent en évidence un effet compensatoire de la tonalité sociale du groupe classe. Ces simulations sont visualisées sur le graphique 17.



*Graphique 17 : Progressions différentielles des élèves selon la tonalité sociale de la classe au CE1*

Note : La ligne horizontale représente la référence, c'est-à-dire les enfants de cadre supérieur ou de prof intermédiaire. Donc, par rapport à cette référence (qui vaut 0), les écarts de progression des enfants de milieu modeste sont matérialisés par la ligne en pointillés. Ainsi, dans une classe où l'on a 30% d'enfants de cadre ou de prof intermédiaire, les enfants de milieu modeste ont (toutes choses égales par ailleurs) des progressions supérieures d'environ 1,5 points à celles des enfants de milieu plus favorisé.

Le graphique montre que dès lors que le groupe classe est composé d'au moins 20% d'enfants de milieu favorisé, les enfants de milieu modeste réalisent (toutes choses égales par ailleurs, et notamment à niveau initial donné) des progressions légèrement meilleures que les enfants de cadres. Cet écart est le solde de deux effets de sens inverse : l'influence négative sur les progressions de l'origine sociale individuelle de l'élève d'une part, de l'effet positif associé chez les élèves de milieu populaire à la fréquentation d'une classe de tonalité sociale plutôt favorisée d'autre part. Par exemple, dans une classe dans laquelle on compte 50% d'enfants de milieu favorisé, un enfant de milieu modeste a des progressions supérieures de 2,29 points à celles d'un enfant de cadre. Ces 2,29 points sont le résultat, d'une part de l'influence négative de l'origine sociale et d'autre part de l'influence positive de la composition sociale de la classe (+3,62 points avec la spécification quadratique). Dans ce graphique, les progressions des enfants de milieu favorisé constituent une référence unique, quel que soit le contexte de la classe, puisque rappelons le, leurs progressions ne sont pas affectées positivement par la tonalité sociale de la classe.

Il faut souligner que c'est surtout dans l'intervalle 40 à 60% d'élèves de milieu favorisé que les écarts deviennent réellement importants (supérieurs à 2 points), ceci renvoie sans doute à des phénomènes de composition de groupes (critical mass d'enfants de cadres, bénéfique aux enfants de milieu populaire, notamment autour d'une situation de relatif équilibre, mais le devenant moins à partir d'un certain seuil), phénomènes dont l'analyse reste à faire.

Evidemment, ces résultats sont des simulations qui restent à interpréter en fonction de la réalité des situations rencontrées dans l'échantillon. Ainsi, dans l'échantillon, il y a moins de 25% des élèves de milieu populaire qui sont effectivement dans des classes où il y a au moins 40% d'élèves favorisés. Il n'y a également que 35 % d'élèves de milieu populaire qui fréquentent des classes où le pourcentage d'élèves favorisés est compris entre 30 et 70%. Autrement dit, la probabilité pour un élève de milieu populaire de se trouver dans une situation dans laquelle la composition sociale de la classe compense le handicap de progression lié à son origine sociale n'est pas très forte. Il faut souligner enfin que cette compensation n'intervient que dans une période de temps donnée. Or, il faut rappeler qu'à l'entrée au CE1, l'écart absolu entre enfants de cadres et enfants d'ouvriers est de l'ordre de 10 points (soit deux-tiers d'écart-type dans la distribution des acquis initiaux). On ne saurait pour autant en déduire qu'il suffirait de 5 années de scolarisation dans ce type de contexte pour que les enfants de milieu défavorisé rattrapent leur retard en terme d'acquisitions. En effet, on ne peut strictement comparer les chiffres qui rendent compte des écarts d'acquisitions d'une année sur l'autre car la nature des acquisitions étant différente chaque année, les échelles de mesure ne sont pas strictement comparables. Cela dit, ces simulations soulignent la portée compensatoire de certaines situations de contexte.

### *III.5 L'exploration des effets de contexte*

Même s'ils sont « quantitativement » modérés, les effets de contexte sont donc importants et on ne saurait s'arrêter à leur seule prise en compte comme des « boîtes noires » : il faut explorer les processus qu'ils recouvrent. Ceci exige des réflexions théoriques sur la nature même du contexte et empiriques sur l'observation de ce contexte.

Dans la recherche en Education, le contexte est tout d'abord perçu comme l'ensemble des « conditions de travail » des élèves, telles que les déterminent des paramètres aussi disparates que les dotations budgétaires, les programmes théoriques et effectifs, les interactions avec le maître et avec les pairs, etc. Tout ceci construit un « destin commun » partagé par les élèves d'un même contexte : c'est l'« offre », comme le disent les économistes, ou le « traitement » auquel ils sont exposés. Evidemment, on ne saurait borner a priori le contexte : sa dimension géographique (un contexte rural, un quartier, comme l'explorent par exemple Goux et Maurin, 2003) peut dans certains cas peser sur la variable cible de l'analyse, et c'est précisément cet empan très large de la notion de contexte (de la zone géographique plus ou moins étendue au « contexte de soi » des psychologues, qui fait la difficulté (mais aussi le caractère oecuménique) de cette notion. Ces derniers se polariseront volontiers sur le contexte interactif : le contexte est alors construit par les interactions maître/élèves et entre élèves, et aussi par les normes et attitudes qui se mettent en place à la fois chez les enseignants et chez les élèves, dans un contexte donné. Dans la recherche d'orientation économique ou

sociologique, on s'attachera davantage à des paramètres du contexte tels que le temps alloué au travail scolaire, la couverture des programmes, la qualification des enseignants, la taille des classes, etc.

Dans ces multiples facettes du contexte, une recherche donnée fait des choix, qui souvent, dans la recherche en « milieu naturel » du moins, dépendent de fait des contraintes de la collecte d'information plus que de réelles considérations théoriques. Dans la recherche sur laquelle nous nous appuyons ici (Duru-Bellat, Landrier-Le Bastard, Suchaut, 2004), nous avons exploré, essentiellement par questionnaires (donc sur des données déclaratives), certains aspects des pratiques enseignantes et des comportements et attitudes des élèves, dans des contextes contrastés. Il ne s'agit là que d'un travail exploratoire, dans la mesure où ces différentes facettes devraient faire l'objet d'une opérationnalisation beaucoup plus stricte, pour qu'elles puissent être introduites ensuite dans un modèle explicatif des progressions permettant d'en vérifier la pertinence.

Un premier volet concerne les pratiques enseignantes. La littérature laisse entendre que les enseignants s'adaptent aux élèves auxquels ils font face, et que ceci est une des voies importantes par lesquelles l'existence de contextes scolaires contrastés tend à accroître les écarts entre les élèves, entre les plus faibles, moins stimulés, et les plus forts, qui le sont davantage. Ont donc été observés, les attentes des enseignants, les objectifs poursuivis ou les contenus déclarés comme étant effectivement mis en œuvre. Ceci repose néanmoins sur les déclarations des enseignants et vu la prégnance des normes nationales, il serait politiquement et pédagogiquement incorrect de reconnaître certaines variations. Effectivement, on n'observe pas de différences sensibles dans les objectifs déclarés par les enseignants. Mais cela n'exclut pas l'existence d'attentes différenciées concernant la réussite des élèves. Elles s'avèrent effectivement très contrastées selon les contextes. Ainsi, les maîtres des écoles favorisées estiment que 75% de leurs élèves obtiendront le bac, contre 43% dans les écoles défavorisées. Les enseignants semblent donc convaincus qu'ils font face à des élèves de niveau inégal, selon les contextes d'enseignement, et ce dès l'école primaire. Ainsi, les programmes sont jugés très inégalement irréalistes selon les classes : d'un tiers dans les classes favorisées à plus des trois quarts dans les classes populaires, soit un écart très important, à ce stade pourtant très précoce de la scolarité. Réciproquement, 40% des maîtres des écoles favorisées disent ne pas renoncer à l'étude de certains points du programme, parce qu'ils leur apparaissent trop difficiles pour leurs élèves, contre 5% dans les écoles défavorisées. Et très massivement, les professeurs des écoles disent modifier le contenu des évaluations en fonction du niveau de la classe, et ce encore plus dans les écoles au public socialement défavorisé (95%) que dans les écoles au public socialement favorisé.

Une première conclusion est donc que les enseignants adaptent leur niveau d'exigence selon les contextes d'enseignement, sans que l'on sache si les variations dans les pratiques pédagogiques effectives sont encore plus marquées que pour ce qui est des pratiques déclarées. Ceci pourrait contribuer à expliquer la tendance à une moindre progression dans les établissements et les classes au public populaire, dans la mesure où il s'agit là de facteurs (en l'occurrence la couverture des programmes et plus largement ce que les anglo-saxons désignent sous le terme d'*opportunity to learn*) dont toute la littérature montre l'influence sur les acquisitions des élèves.

Explorons à présent les normes et les comportements des élèves selon le contexte de la classe, en pointant celles et ceux susceptibles d'expliquer les effets, si modérés soient-ils, du contexte scolaire sur la réussite, sans que l'insertion dans une chaîne causale soit bien sûr immédiate. Nous mettrons ici l'accent à la fois sur les comportements en classe et les normes de groupes, dont on connaît l'importance sur la réussite, et sur la face plus subjective de la scolarité, moins explorée quant à sa sensibilité au *school mix*.

Une famille d'interrogations concernait les normes du groupe classe. Un indicateur de normalité/déviance scolaire a été construit à partir de réponses à des questions telles que « il est normal de... se bagarrer pendant la récréation, mentir pour éviter une punition... ». Globalement, les élèves répondent de manière très conformiste (avec sans doute un biais de désirabilité sociale), et on n'observe pas de différences selon le public de la classe : il existe bien un ensemble de normes partagées par tous, quelle que soit la tonalité sociale de la classe. Quelques items font cependant des différences significatives : « se bagarrer en classe » ou « taper un élève qui a pris vos affaires » est plus souvent considéré comme normal dans les classes défavorisées, alors que « dénoncer un copain qui a volé » ou « critiquer la maîtresse » est moins mal jugé dans les classes socialement favorisées. Il reste qu'au total, les différences de climat des classes et de normes comportementales semblent bien ténues, si on se fonde sur des déclarations d'élèves. On ne saurait donc en faire une piste explicative fiable pour rendre compte des différentiels de progression.

Parmi les dimensions subjectives des scolarités, il est apparu intéressant de prendre en compte un certain nombre d'attitudes dont on sait (Monteil, Huguet, 2002) qu'à la fois elles affectent la performance scolaire et sont sensibles au contexte. On sait par exemple que dans l'espace de la classe, où les élèves font l'objet d'évaluations constantes de la part des maîtres, des processus de comparaison sociale sont sans cesse à l'œuvre, et influent sur les performances, sachant qu'à cet égard, le niveau des autres élèves (et donc la composition du groupe, en termes scolaire ou social) est essentiel. On a donc estimé des modèles logistiques explicatifs de la probabilité de s'estimer d'un niveau faible ou bon. L'évaluation de son propre niveau de réussite n'apparaît pas liée à la composition scolaire de sa classe. Par contre, il existe un effet

du social mix : à caractéristiques individuelles identiques (notamment à valeur scolaire équivalente, mesurée par des épreuves standardisées), les élèves des classes les plus défavorisées socialement s'attribuent des niveaux plus élevés que leurs homologues des classes moins défavorisées. Ceci pourrait s'expliquer par des comparaisons sociales moins défavorables (moins « menaçantes » pour l'estime de soi), mais aussi peut-être par la notation, dont nombre de travaux montrent qu'elle est en moyenne plus indulgente, dans les milieux scolaires défavorisés.

Il faut souligner que cette évaluation plus positive est en décalage avec le niveau objectif, plus faible, de ces classes défavorisées. Au total, les auto-évaluations des élèves sont de niveau comparable dans toutes les classes, quelle que soit leur composition scolaire ou sociale ; de même, les élèves « notent » leur classe de façon parfaitement uniforme. Ces constats indiquent que l'auto-évaluation est contextualisée, ordinale au sein d'un contexte et bornée par lui, tout comme les évaluations en général<sup>37</sup>. En d'autres termes, et comme le dit Felouzis (2003), au vu de résultats convergents au niveau du collège, « les normes d'évaluation mais aussi les normes d'exigences sont étroitement dépendantes du contexte pédagogique et social ».

Une facette distincte de l'auto-évaluation des élèves est la manière dont ils expliquent ce qui leur arrive ; ces processus d'attribution ne seraient pas sans rapport avec la réussite, dans la mesure où prévaut, à l'école comme dans la société, une « norme d'internalité » (chacun est censé être responsable de ce qui lui arrive), avec pour conséquence que les élèves qui sont perçus comme « internes » font l'objet d'évaluations plus favorables (Bressoux, Pansu, 2003). Au niveau individuel et en primaire, les élèves de milieu défavorisé invoquent plus souvent des causes externes (ne dépendant pas d'eux-mêmes telles que « c'est mal expliqué », « les autres élèves me dérangent »), ce qui concorde avec les acquis de la psychologie sociale, montrant que globalement les « dominés » sont plus externes que les « dominants » ; mais cet écart entre des élèves de milieu favorisé et tendanciellement « internes », et des élèves de milieu défavorisé tendanciellement « externes », disparaît dans les classes favorisées, sans doute parce que la « norme d'internalité » s'impose davantage au groupe. Enfin, les projets scolaires et professionnels des élèves ont été examinés, en ce qu'ils sont à la fois un « produit » de la socialisation scolaire et un « ingrédient » de l'investissement dans les études. L'ambition des enfants en fin de primaire apparaît (déjà) variable selon leur milieu social d'origine (les enfants de cadres projettent plus souvent de devenir cadres eux-mêmes). Mais, comme la littérature existante le laissait prévoir, la classe fréquentée exerce une influence sur

---

<sup>37</sup> On sait en effet que les notes mises par les enseignants sont sensibles au contexte et proches d'une classe à l'autre indépendamment du niveau objectif des élèves quand on peut mesurer celui-ci avec des épreuves communes. La conséquence, vu la prégnance d'une norme de notation à la moyenne (autour de 10), est, par rapport à ce niveau objectif, une notation de fait plus sévère dans les classes faibles et plus indulgente dans les classes fortes (Bressoux, Pansu, 2003).

les aspirations des élèves, notamment sur les plus défavorisés : ils sont encore moins nombreux à envisager une profession d'ouvrier quand ils fréquentent une classe au public favorisé que quand ils sont scolarisés dans une classe au public moins favorisé (0% contre plus de 9%), et réciproquement plus nombreux à envisager alors une profession de cadres (51% contre 33%). Les aspirations professionnelles des élèves de milieu favorisé sont moins sensibles au contexte et globalement plus ambitieuses.

En conclusion, il apparaît donc que la composition scolaire et sociale de l'établissement fréquenté affecte les attitudes et les visées scolaires qui vont vraisemblablement influencer sur l'investissement et la réussite scolaire, et aussi sous-tendre les orientations. Le rôle de la tonalité sociale du contexte est peut-être d'ailleurs plus fort concernant les choix scolaires que la réussite stricto sensu. Dans une recherche au niveau collège, Felouzis (2003) constate ainsi que la ségrégation ethnique, si elle jouait peu sur les résultats, semblait davantage influente sur les orientations des élèves.

## Conclusion

La question de l'appréhension des effets de contexte reste difficile et délicate à la fois sur le plan méthodologique et d'un point de vue politique, cela pour plusieurs raisons. Tout d'abord, plus encore que dans d'autres domaines, il est capital de bien spécifier les résultats obtenus. Nous l'avons évoqué, les effets de telle ou telle modalité du contexte peuvent être de sens différents selon ce qui est étudié (acquis scolaires ou attitudes), mais aussi selon le niveau scolaire ou la discipline évaluée. Les résultats doivent également être spécifiés eu égard au champ exploré. Dans la recherche en Education, en milieu « naturel », on ne travaille que sur ce qui existe, avec en particulier la diversité des situations telle qu'elle est ; il est clair qu'on ne peut pas faire de manipulation sur cette variable (de contexte), contrairement à l'approche expérimentale. Les résultats sont donc contingents à la plage de variation des données empiriques intégrées dans l'analyse. De plus, dans la réalité, les diverses dimensions du contexte sont fréquemment liées ; la clause « toutes choses égales par ailleurs » est évidemment une abstraction, qui a sa place dans un raisonnement de type expérimental (Passeron, 1991) mais ne dispense pas d'un retour aux conditions réelles, historiques, sauf à faire de la sociologie fiction. La question de savoir si cela a toujours un sens d'estimer l'impact spécifique de facteurs qui dans la réalité sont très fréquemment confondus, qui se pose dans toutes les sciences humaines, revêt ici une portée politique importante. Ainsi, quand Opdenakker et van Damme (2002) montrent qu'un certain nombre de caractéristiques des établissements associés à une bonne efficacité (un climat d'ordre, une bonne couverture des programmes, etc.) ne s'avèrent plus influents quand on tient compte de la tonalité sociale du public d'élèves, ils soulignent à la fois que le social mix rend plus facile ou au contraire nuit à la mise en œuvre de certaines pratiques pédagogiques et par ailleurs que les politiques qui sanctionnent les établissements peu efficaces sans tenir compte de leur public les culpabilisent sans doute excessivement.

A travers ces travaux et ces débats, c'est évidemment la question de l'influence spécifique de l'école, des marges d'action des enseignants, qui est posée. De façon claire, les deux dimensions explorées dans ce chapitre montrent que l'environnement de la classe, que ce soit par sa structure pédagogique ou sa composition scolaire et sociale, influence les chances de progression des élèves à l'école élémentaire.

La responsabilité « technique » des chercheurs, par leur capacité à produire des analyses fiables et leur souci de croiser les perspectives et les méthodes et de répliquer tant qu'il le faut leurs travaux, est alors d'autant plus lourde que les enjeux sont importants.

## *Chapitre 4*

### *Les apprentissages des élèves : structure et évolution*

L'ensemble des résultats généraux sur l'efficacité et l'efficience de l'école primaire française exposés dans le premier chapitre incite à se pencher plus précisément sur les apprentissages des élèves, ceux-ci étant à la base de la mesure de la qualité de notre système. Si l'amélioration des résultats passe par des mesures globales visant à agir sur le contexte scolaire dans son ensemble, comme on a pu le voir dans les deux chapitres précédents, l'analyse fine des acquisitions des élèves peut également, de façon indirecte, livrer des résultats précieux en termes de politique éducative. Notre questionnement général comporte deux aspects complémentaires abordés par le biais d'une approche micro. Le premier concerne la transversalité des apprentissages avec l'idée que ceux-ci ne peuvent se concevoir de façon isolée mais qu'ils s'inscrivent au contraire dans une logique d'interdépendance. La maîtrise d'une compétence peut être ainsi fortement liée à la maîtrise d'autres compétences appartenant, soit à un même domaine d'acquisition, soit à des domaines différents ; la notion de bloc de compétences est alors mobilisée pour rendre compte de ces liens entre les nombreuses dimensions des apprentissages. Le second aspect porte sur l'évolution des compétences dans le temps. Les acquisitions des élèves se construisent au fil des années et il est essentiel de connaître les principes généraux qui gouvernent cette évolution en faisant apparaître des relations de dépendance entre bloc des compétences et en identifiant quels sont ceux qui semblent être les plus prédictifs de la réussite ultérieure. Cette problématique revêt une importance réelle en terme de politique éducative. On s'interrogera notamment sur la notion de déterminisme dans les acquisitions des élèves et sur les rythmes variés de progression des élèves au cours de la scolarité élémentaire. La question de la nature des contenus d'enseignement est indirectement abordée au travers du rôle central tenu par certaines compétences dans la réussite scolaire. La question des rythmes d'apprentissage et de leur prise en compte concrète par l'institution dans les programmes scolaires et leurs modalités de transmission apparaît également en filigrane.

Deux sources de données sont mobilisées pour traiter cette problématique. En premier lieu, des analyses originales menées sur le panel 1997 permettent de dresser un cadre général de l'évolution des acquisitions des élèves entre la sortie de l'école maternelle et la fin de l'école élémentaire. En second lieu, des données longitudinales récentes (Morlaix, Suchaut, 2007a, 2007b) permettent d'identifier les compétences centrales dans les apprentissages au cours du cycle III. Les deux bases de données présentent l'intérêt d'utiliser les mêmes épreuves d'évaluation des élèves. Ce chapitre se structure en trois parties. La première examine l'évolution des acquisitions des élèves au cours de l'école élémentaire en mettant l'accent sur la prédictivité de certaines compétences. La seconde partie rend compte précisément de la structure des acquisitions des élèves à l'entrée au cycle III et au début du collège ainsi que des relations qui les lient. La troisième partie s'interroge sur le déterminisme des acquisitions en examinant les relations entre les compétences scolaires, les capacités cognitives et l'origine sociale des élèves.

## I. L'évolution globale des acquisitions à l'école élémentaire

Les données de panel offrent des avantages notables par rapport aux enquêtes spécifiques utilisées dans les recherches, ces avantages concernent la représentativité des échantillons et la structure longitudinales des données. Le panel 1997 regroupe des informations sur 9858 élèves scolarisés à la rentrée 1997 dans 1401 écoles élémentaires de la France métropolitaine, 9639 élèves ayant fait l'objet d'un suivi de scolarité jusqu'au collège. Trois prises d'informations sur le niveau d'acquisition des élèves ont été réalisées : début CP, début CE2 et à l'entrée en 6<sup>ème</sup>. Notre analyse repose sur ces différentes mesures effectuées à trois étapes de la scolarité. Le schéma ci-après visualise la structure temporelle de l'évaluation des élèves qui couvre deux des trois cycles de l'école primaire.

Cycle II		Cycle III		
CP	CE1	CE2	CM1	CM2
<b>Evaluation 1</b>		<b>Evaluation 2</b>		<b>Evaluation 3</b>

Graphique 18 : Structure temporelle de l'évaluation des élèves (panel 1997)

A l'entrée au CP, des évaluations spécifiques administrées en septembre 1997 ont concerné douze dimensions qui peuvent être regroupées en cinq grands domaines : connaissances générales, compétences verbales et familiarité avec l'écrit, compétences logiques et familiarité avec le nombre, concepts liés au temps et à l'espace, comportements et attention. Au niveau du CE2, les évaluations nationales (administrées en septembre 1999) évaluent sept dimensions des acquisitions : compréhension, outils de la langue, production d'écrits, travaux géométriques, mesure, travaux numériques et résolution de problèmes. Trois années plus tard à l'entrée en sixième (septembre 2002), les évaluations nationales explorent également sept domaines : traitements opératoires, travaux géométriques, numération, problèmes numériques, traitement de l'information, compréhension, outils de la langue et production de texte. On constate, à la lecture de ces différents domaines que les évaluations réalisées à l'entrée en CE2 et en 6<sup>ème</sup> peuvent se regrouper nettement selon les deux disciplines clés de l'école élémentaire (français et mathématiques), alors qu'à l'entrée au CP, si des domaines distinguent la maîtrise de la langue et celle du nombre, d'autres domaines sont plus généraux et n'apparaissent pas, a priori, liés à une discipline particulière, c'est le cas des concepts liés à l'espace et au temps.

Le niveau d'acquisition des élèves peut dans un premier temps s'apprécier de manière globale ; il est en effet d'usage de calculer des scores moyens pour chacun des niveaux considérés, prenant en compte l'ensemble des items présents dans les épreuves d'évaluation. Le tableau 34 présente les scores moyens (exprimés en pourcentage de réussite) ainsi que leur indice de dispersion sur l'échantillon d'élèves du panel. Les effectifs mentionnés dans le tableau 1 font apparaître une déperdition assez forte d'élèves au cours du suivi longitudinal. Celle-ci s'explique par différents facteurs, dont le redoublement qui a concerné une partie des élèves lors de l'une ou l'autre des cinq années de la scolarité primaire. Nous disposons en fait des résultats aux trois évaluations de CP (1997), CE2 (1999) et 6<sup>ème</sup> (2002) pour une cohorte numériquement plus modeste que l'échantillon initial.

*Tableau 34 : Scores moyens aux épreuves de CP, CE2 et 6<sup>ème</sup> (panel 1997)*

	Moyenne	Ecart-type	Coefficient de variation	N
Score évaluation CP	69,0	12,9	0,19	9531
Score évaluation CE2	67,5	14,4	0,21	7204
Score évaluation 6 <sup>ème</sup>	68,5	14,5	0,21	6222

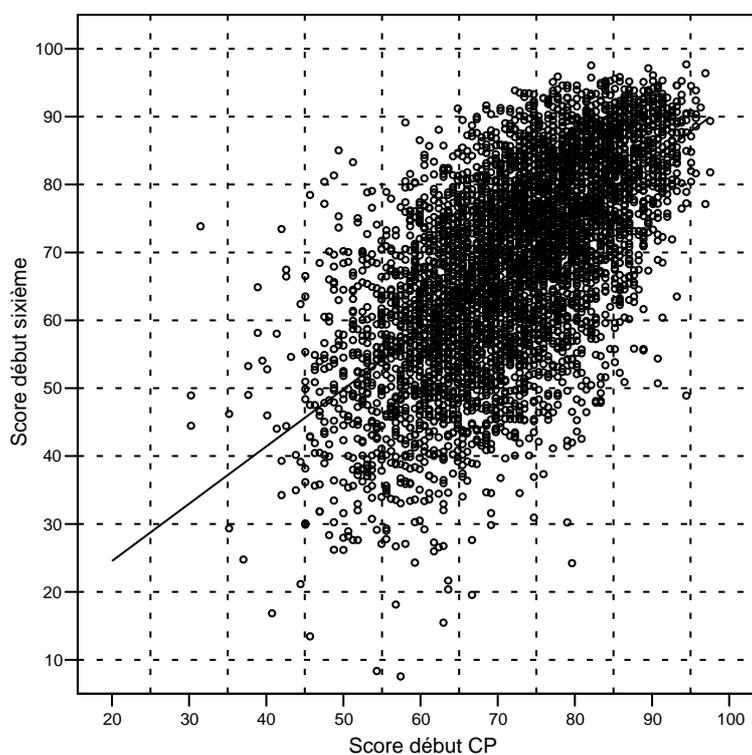
Aux trois niveaux scolaires considérés, les scores apparaissent très semblables puisque les élèves réussissent en moyenne un peu plus des deux-tiers des items. La dispersion des performances est elle-même comparable d'un niveau scolaire à l'autre (le coefficient de variation est toutefois légèrement plus faible pour l'évaluation CP que pour les autres niveaux). Une première analyse très générale consiste à examiner les corrélations entre ces trois scores. Le tableau 35 présente trois modèles de régression univariée permettant d'apprécier l'évolution des acquis des élèves. Le modèle A analyse la variance du score de début de CE2 en fonction du score obtenu à l'entrée au CP alors que les modèles B et C analysent la variance du score de 6<sup>ème</sup> en fonction du score de CE2 (modèle B) et de CP (modèle C). Le niveau de performances des élèves à l'entrée au CP explique 45% des différences de performances à l'entrée en CE2 (modèle M1). Ce chiffre traduit, même à deux années d'intervalle, une liaison assez forte entre les deux évaluations car les études qui ne portent que sur une année scolaire (le CP notamment) affichent généralement un R<sup>2</sup> proche de 0,50 (soit 50% de la variance du score de fin d'année expliquée par le score de début d'année scolaire).

Tableau 35 : Modèles de progression globale des élèves (panel 1997)

Variables explicatives	A		B		C	
	Score CE2		Score 6 <sup>ème</sup>		Score 6 <sup>ème</sup>	
	coefficient	sign.	coefficient	sign.	coefficient	sign.
Score début CP	+0,88	***			+0,85	***
Score début CE2			+0,82	***		
Constante	4,53	***	12,27	***	7,00	***
Pourcentage de variance expliquée	45,0%		58,3%		39,2%	

\*\*\* : significatif au seuil de 1%

A l'entrée au collège, les acquis des élèves sont expliqués à la hauteur de 58% par le niveau scolaire au début du CE2 (modèle B) ; ce chiffre ne peut être comparé au précédent car il ne s'agit pas de la même échelle de mesure (celle-ci étant spécifique à chacun des scores), mais il traduit une forte relation entre les performances des élèves mesurée à trois années d'intervalle. Sur l'ensemble de la scolarité élémentaire, le lien entre les acquisitions des élèves devient beaucoup plus faible puisque le pourcentage de variance du score de 6<sup>ème</sup> expliqué par le score à l'entrée au CP n'est que de 39% (modèle C). Autrement dit, le niveau des élèves à l'entrée au CP ne prédit qu'imparfaitement la réussite en fin d'école primaire. Le graphique 19 permet de visualiser la relation entre les deux scores à cinq années d'intervalle. On remarque que le nuage de points présente une allure différente selon le niveau des élèves à l'entrée au CP.



Graphique 19 : Relation entre les scores de CP et de 6<sup>ème</sup> (panel 1997)

Pour les élèves initialement faibles (sur la gauche du graphique), le nuage de points est bien moins resserré que pour les élèves présentant des scores élevés au début du CP (sur la droite du graphique). D'ailleurs une spécification quadratique de la variable rendant compte du score de CP est davantage adaptée aux données et confirme le fait que la prédiction du niveau à l'entrée en 6<sup>ème</sup> sur la base du niveau de CP est d'autant plus juste que les élèves ont un bon niveau au début de leur scolarité élémentaire. Pour illustrer plus précisément ce constat, nous avons estimé la relation entre le score de CP et celui de 6<sup>ème</sup> pour la population d'élèves la plus en difficulté à l'entrée au CP. Le critère retenu étant l'écart-type du score de début CP, les élèves les plus faibles sont ceux qui obtiennent un score inférieur ou égal à 56 (un écart-type au-dessous de la moyenne) à l'entrée au CP. Ces élèves qui sont au nombre de 1288 ne sont plus que 346 à parvenir à l'entrée en 6<sup>ème</sup> dans notre échantillon (une forte proportion d'entre eux ayant redoublé au moins l'une des 5 années de la scolarité élémentaire). Il est frappant de constater la très faible corrélation entre leurs performances à cinq années d'intervalle ( $r = +0,10$ , significatif au seuil de 10%). Ces élèves, sans doute atypiques, du fait qu'ils parviennent à l'entrée en 6<sup>ème</sup> sans redoublement tout en ayant un niveau très faible à l'entrée au CP, réalisent, pour certains d'entre eux, des progressions très fortes sur cette période de cinq années en obtenant des scores supérieurs à la moyenne à l'entrée en 6<sup>ème</sup>.

A l'opposé, si l'on s'intéresse aux progressions des élèves ayant obtenu les scores les plus élevés à l'entrée au CP (soit un score supérieur ou égal à 82 qui correspond à un écart-type au-dessus de la moyenne), les résultats sont bien différents. La corrélation entre score de CP et de 6<sup>ème</sup> est de  $+0,27$  (significatif à  $.01$ ). Il n'y a qu'une faible proportion d'élèves qui affichent un score inférieur à la moyenne à l'entrée en 6<sup>ème</sup> et la relation entre les deux scores serait beaucoup plus forte si l'effet de plafond n'était pas présent au niveau des scores de début 6<sup>ème</sup>. En effet, les épreuves d'évaluation apparaissent peu discriminantes pour les élèves aux aptitudes initialement élevées : nombre d'élèves affichent des scores supérieurs à 90% de réussite.

En résumé, si, en moyenne, les acquisitions des élèves au cours des cinq années de l'école élémentaires s'inscrivent dans une logique d'évolution classique, il n'en reste pas moins que le score de début CP n'arrive pas à rendre compte de l'ensemble des différences d'acquisitions observées à l'entrée en 6<sup>ème</sup>. Même en prenant en compte les caractéristiques socio-démographiques des élèves (mois de naissance et profession des parents principalement), les modèles sont loin d'expliquer la totalité de la variance du score des évaluations à l'entrée au collège. Les chiffres annoncés peuvent bien sûr être lus différemment selon le point de vue que l'on adopte et il est possible d'avoir une vision plus déterministe du phénomène observé en insistant sur le fait que le niveau à l'entrée à l'école élémentaire est le plus fort prédicteur de la réussite ultérieure (Caille, Rosenwald, 2006). Deux principales conclusions sont dégagées par ces auteurs. La première porte sur le niveau

des élèves mesuré par la première évaluation à l'entrée au CP. D'après les auteurs, les élèves entrent au cours préparatoire avec des niveaux de compétences socialement différenciés, ceux-ci étant liés à la fois à des aptitudes personnelles des élèves et à l'influence du milieu familial. En effet, d'après Caille et Rosenwald, à caractéristiques démographiques (sexe, date de naissance) et familiales (diplôme des parents, taille de la famille, activité de la mère...) comparables, un enfant d'enseignant réussit plus d'items (7% de plus en moyenne) à l'entrée au CP, qu'un enfant d'ouvrier ou qu'un enfant dont le père est inactif (9% de plus). Ces écarts sont deux fois plus élevés que ceux associés à taille de la famille ou à la nationalité et leurs progressions diffèrent selon leur milieu social d'origine (même à niveau initial comparable).

La seconde conclusion porte sur les progressions des élèves au cours de la scolarité élémentaire. Les progressions durant les cinq années du CP au CM2 diffèrent, à niveau initial comparable, selon le milieu social de l'élève, et les écarts se creusent entre les élèves au fur et à mesure de la scolarité élémentaire<sup>38</sup>. Les mêmes auteurs rappellent toutefois que le niveau de compétences en français et mathématiques à l'entrée en 6<sup>ème</sup> est fortement lié au niveau des élèves à leur entrée à l'école primaire. Ainsi, un écolier qui faisait partie des 10% d'élèves les plus faibles à l'entrée au CP réussit à caractéristiques démographiques et familiales comparables, entre 30% (en français) à 40% (en mathématiques) d'items en moins qu'un élève faisant partie des 10% les plus forts à l'entrée au CP. Les disparités de réussite selon le niveau initial, sont en français comme en mathématiques, quatre à cinq fois plus fortes que celles associées à l'origine sociale. Les auteurs estiment d'ailleurs qu'à la fin de la scolarité élémentaire, la moitié des inégalités sociales de réussite est due aux différences de compétences que présentaient les élèves à l'entrée au cours préparatoire. Cependant, au-delà du rôle prépondérant et persistant des compétences initiales, le niveau d'acquis en français et en mathématiques des élèves en 6<sup>ème</sup> apparaît une nouvelle fois liés aux diplômes des parents et à l'origine sociale. Le diplôme de la mère notamment reste l'une des variables les plus intéressantes dans l'explication des différences de progressions.

D'autres recherches (Cosnefroy, Gibert, 2006) basées sur le même panel portent sur l'évolution des compétences dans le temps. Leurs conclusions sont de nature différente, mettant l'accent sur le fait que les élèves dont le niveau initial est le plus élevé à l'entrée au CP ont tendance à développer plus rapidement leurs compétences que les autres, mais aussi à présenter un maximum plus tôt. A l'inverse, les élèves les plus faibles au départ ont tendance à progresser plus lentement que les autres et à observer un maximum plus tardivement. Pour notre part, nous mettrons avant tout l'accent sur le fait que tout n'est pas joué à l'entrée au CP et, qu'au-delà de la tendance moyenne, des profils d'évolution des acquisitions variés existent.

---

<sup>38</sup> Le mois de naissance de l'élève exerce lui aussi une influence très marquée sur les progressions des élèves mais celle-ci tend naturellement à s'estomper au fil des années de l'école primaire (Florin, Cosnefroy, Guimard, 2004).

Pour terminer avec ces constats généraux de l'évolution des acquis des élèves et pour faire le lien avec la suite de notre démarche, il n'est pas inutile de repérer quelles dimensions de l'évaluation de début CP sont les plus prédictives de la réussite à l'entrée au collège. Pour cela, un modèle de régression analysant la variance du score de 6<sup>ème</sup> intègre comme variables explicatives les douze dimensions de l'évaluation de début CP (tableau 36).

*Tableau 36 : Modèle analysant la variance du score de 6<sup>ème</sup> en fonction des différentes dimensions de l'évaluation de début CP*

	coefficients non standardisés	coefficients standardisés	sign.
Score connaissances générales	+0,07	+0,07	***
Score connaissance de l'écrit	+0,04	+0,06	***
Score lecture tâches phonologiques	+0,02	+0,03	**
Scores lecture morphologie et syntaxe	+0,04	+0,06	***
Score compétences épreuves numériques	+0,15	+0,20	***
Scores concepts liés au temps	+0,20	+0,21	***
Score compréhension orale	+0,07	+0,08	***
Scores compétences d'écriture	+0,09	+0,12	***
Scores concepts liés à l'espace	+0,02	+0,02	*
Scores compétences de prélecture	+0,07	+0,09	***
Scores nombre et figures géométriques	+0,05	+0,06	***
Score culture technique	-0,01	-0,01	ns
Constante	7,91		
Pourcentage de variance expliquée	41,9%		

n.s. : non significatif, \* : significatif au seuil de 10%, \*\* : significatif au seuil de 5%,

\*\*\* : significatif au seuil de 1%

Les coefficients standardisés permettent de comparer directement les impacts des différents scores sur le score global (français et mathématiques) de 6<sup>ème</sup>. Deux dimensions semblent particulièrement prédictives de la réussite à l'entrée en 6<sup>ème</sup>, il s'agit des concepts liés au temps (coefficient de +0,21) et des compétences aux épreuves numériques. Les autres dimensions affichent des coefficients plus modestes (de 0,06 à 0,12) alors que deux dimensions apparaissent comme presque (concepts liés à l'espace) ou totalement (culture technique) indépendantes du niveau global de compétences à l'entrée en 6<sup>ème</sup>. Les acquis des élèves à l'entrée à l'école élémentaire n'ont donc pas tous le même poids et certains apparaissent, plus que d'autres, jouer un rôle déterminant dans la réussite ultérieure. C'est donc le cas pour les concepts liés au temps et les compétences numériques, ces deux dimensions expliquant à elles seules plus de 35% de la variance du score global à l'entrée au collège. Ce dernier point permet de s'interroger sur le fait que les acquis des élèves relèvent de dimensions qui n'ont pas toutes la même importance dans l'explication des différences de réussite interindividuelles, il est donc essentiel d'essayer de mieux cerner la nature des acquis scolaires et d'examiner précisément les liens qu'ils entretiennent entre eux.

## II. Structure et évolution des compétences au cycle III

Depuis 1989, le système éducatif français s'est doté d'outils d'évaluation des élèves à différents niveaux de la scolarité primaire et secondaire<sup>39</sup>. Ces évaluations de masse, distribuées et utilisées à l'échelon national ont, d'une part contribué de manière progressive à la diffusion de la culture de l'évaluation à l'ensemble de la communauté éducative et, d'autre part, permis aux enseignants d'approfondir leur réflexion pédagogique (Joutard, Thélot, 1999 ; M.E.N., 2000). L'objectif initial de ces évaluations nationales était de « *permettre l'observation des compétences et apprécier les réussites et les difficultés éventuelles de chaque élève considéré individuellement, à un moment précis de la scolarité. Elles fournissent aux enseignants des repères exploitables quant aux types d'erreurs fréquemment produites par les élèves au cours de leur apprentissage...* »<sup>40</sup>. Même si la dimension diagnostique reste aujourd'hui prioritaire, ces évaluations peuvent être utilisées comme des instruments de pilotage car elles rendent objectivement compte du niveau d'acquisition des élèves à plusieurs niveaux : national, régional, local.

Pour le chercheur en éducation, les évaluations nationales présentent également un grand intérêt dans le sens où celles-ci balayaient un vaste ensemble de compétences et de connaissances alors que les tests couramment mobilisés dans les recherches (françaises notamment) sont beaucoup plus restrictifs sur ce plan. Nous essaierons de montrer en quoi les outils d'évaluation peuvent livrer des informations capitales sur les apprentissages réalisés par les élèves, à condition de ne pas se limiter à la seule approche institutionnelle et à la classification des compétences établies par les concepteurs des épreuves.

### II.1 Les compétences à l'entrée au CE2

Notre démonstration se base sur une exploitation des résultats aux épreuves nationales. Des données ont été collectées sur un échantillon d'environ 700 élèves à différents moments de la scolarité, ces données concernent les résultats détaillés d'une même cohorte d'élèves aux évaluations nationales de 1999 pour le niveau du CE2<sup>41</sup> et de 2002 pour la classe de sixième ; les élèves ont également été testés en fin de cinquième (deuxième année de collège) en 2003.

---

<sup>39</sup> Les évaluations nationales mises en place par la D.E.P. (Direction de l'Evaluation et de la Prospective) ont concerné à l'origine trois niveaux de la scolarité : CE2 (3<sup>ème</sup> année de l'enseignement élémentaire), sixième (1<sup>ère</sup> année de l'enseignement secondaire) et seconde (5<sup>ème</sup> année de l'enseignement secondaire).

<sup>40</sup> Note de service 2002-105 du 30.04.2002, B.O. N°19 du 9 mai 2002.

<sup>41</sup> Les données analysées ici concernent la classe de CE2. L'échantillon comporte 671 élèves scolarisés dans 31 écoles d'une circonscription de la Côte d'Or. Nous disposons pour tous ces élèves des résultats détaillés aux évaluations nationales de septembre 1999 (résultats item par item) et d'informations socio-démographiques : mois et année de naissance, sexe, catégorie socio-professionnelle du père de l'élève.

Nous disposons aussi d'autres informations socio-démographiques et scolaires pour tous les élèves.

En 1999, dans les évaluations du Ministère de l'Education nationale de CE2, sept champs de compétences sont préalablement définis par les concepteurs des épreuves : compréhension (41 items), outils de la langue (40 items), production d'écrits (10 items) pour le français ; travaux géométriques (17 items), mesure (32 items), travaux numériques (32 items), et résolution de problèmes (9 items) pour les mathématiques. Les 91 items de français sont répartis dans 18 exercices, tandis que les 80 items de mathématiques sont issus de 27 exercices. Plusieurs séquences (4 en mathématiques/4 en français) sont prévues pour faire passer aux élèves l'ensemble de ces épreuves. Les cahiers d'évaluations (un en français/un en mathématiques) remplis par les élèves sont alors corrigés par les enseignants, grâce aux consignes de correction nationales.

Au total, les évaluations nationales de CE2 comportent 171 items regroupés en 15 compétences de français et 27 de mathématiques, chacune de ces compétences correspond dans la grande majorité des cas à un exercice du cahier d'évaluation. Ces exercices reflètent les programmes scolaires officiels en vigueur et constituent à ce titre une référence commune pour l'ensemble des enseignants et des élèves. Sur la base des simples taux de réussite aux items, il est possible d'avoir une vision du niveau d'acquisition des élèves à un moment donné de leur scolarité et d'identifier les lacunes de chacun en vue d'y remédier au cours de l'année scolaire. Outre cette perspective diagnostique et pédagogique, les résultats des évaluations fournissent des indicateurs objectifs de la réussite des élèves. Le Ministère insiste toutefois sur l'utilisation inadaptée qui pourrait être faite des résultats des évaluations, ceux-ci ne pouvant, en aucun cas, être considérés comme des normes à atteindre, mais seulement comme un diagnostic, en début de cycle, des réussites et des difficultés des élèves. Il est également clairement énoncé dans le discours officiel que les épreuves d'évaluation ne rendent que partiellement compte des compétences et des connaissances des élèves puisque la passation est collective et limitée à des exercices écrits. La prudence est aussi conseillée quant à l'interprétation de l'évolution des taux de réussite au cours du temps, les épreuves n'étant pas la plupart du temps identiques d'une année sur l'autre comme nous avons pu le mentionner dans le premier chapitre de ce texte. Au-delà de ces limites, l'exploitation des évaluations nationales pose de réels problèmes méthodologiques quand on utilise la catégorisation des compétences proposée par les concepteurs de ces outils. Deux points essentiels méritent d'être soulignés, le premier concerne l'échelle de mesure utilisée, le second la définition des différentes compétences.

L'échelle utilisée pour apprécier la réussite des élèves est très variable selon le nombre d'items retenus pour mesurer chacune des compétences. Par exemple, la compétence «*comprendre un texte...*» est évaluée par 12 items alors que la compétence «*construire une figure simple sur un quadrillage...*» ne concerne qu'un seul item. Selon le cas, la graduation de la réussite est donc très fluctuante d'une compétence à l'autre, allant de la simple dichotomie « échec / réussite » à une échelle en 12 scores. Une autre difficulté est relative au seuil de réussite : à partir de quel score peut-on considérer que l'élève maîtrise la compétence ? On peut bien sûr fixer un seuil de réussite commun à toutes les compétences en adoptant les principes en usage (75% est généralement le seuil retenu), mais les dissymétries entre les différentes échelles rendent cette solution trop imparfaite, notamment pour les compétences qui comportent très peu d'items.

Indépendamment des problèmes liés à l'échelle de mesure, les évaluations nationales peuvent donner lieu à des interrogations encore plus profondes concernant la définition même des compétences évaluées, ces compétences étant définies par les commissions chargées de l'élaboration des épreuves. Quand on teste la pertinence des regroupements d'items qui composent les compétences, deux observations majeures ressortent. En premier lieu, il apparaît que des items provenant de compétences différentes, voir de disciplines différentes, entretiennent des liens statistiques très forts. On peut ainsi imaginer que les items figurant dans les épreuves puissent s'associer différemment pour rendre compte des compétences des élèves. En second lieu, les items censés rendre compte d'une même compétence peuvent n'être que très faiblement corrélés. C'est par exemple le cas de la compétence « *comprendre et savoir appliquer les consignes courantes du travail scolaire* » pour laquelle la plupart des corrélations entre les 5 items la composant ne sont pas significatives.

De nombreux autres exemples pourraient être évoqués pour montrer qu'une définition a priori des compétences, n'est pas adéquate. L'approche institutionnelle relève d'un certain arbitraire dans la définition des compétences et il n'est pas certain que chaque item mesure réellement la compétence visée. En outre, cette approche peut être considérée comme statique dans la mesure où elle se limite à une catégorisation par domaine d'acquisition et par discipline. En celà, elle ne permet pas de prendre en compte la logique d'interdépendance des acquis des élèves.

Toutes ces remarques soulignent la difficulté à saisir concrètement la notion de compétence dans le domaine de l'éducation (Crahay, 2006). Même si, d'un point de vue théorique, les auteurs s'accordent sur une définition générale de la compétence qui peut être vue comme un ensemble intégré de connaissances susceptibles d'être mobilisées pour accomplir des tâches et donc orientée vers une finalité (Gillet, 1991), la question de la mesure demeure. Pour progresser dans ce domaine, nous nous proposons d'identifier, non plus à priori, mais à

posteriori et de manière empirique, les diverses compétences mises en œuvre par les élèves à partir des situations d'évaluation figurant dans les épreuves nationales. Notre travail s'inspire globalement des procédures utilisées pour les validations empiriques a posteriori des épreuves d'évaluation (De Ketele, Gérard, 2005). La démarche méthodologique que nous avons adoptée permet d'appréhender les apprentissages des élèves de façon plus précise et plus dynamique en centrant l'analyse sur les items et non sur les compétences définies préalablement dans les épreuves. Même si cette approche comporte elle aussi des limites, notamment quant au degré de difficulté des items (Demeuse, Henry, 2004), elle présente des avantages certains.

L'objectif principal de notre travail est donc de mieux comprendre comment les acquisitions des élèves s'articulent en identifiant des structures capables de dresser une cartographie des compétences développées par les élèves à un moment donné de la scolarité primaire, en l'occurrence, à l'entrée au cycle III. Le terme de compétence intègre ici dans sa définition les différentes capacités, aptitudes, connaissances et attitudes qui mobilisent chez les individus des ressources cognitives pour réaliser une tâche.

Dans un premier temps, nous soulignerons l'intérêt de l'approche centrée sur les items, puis nous présenterons la démarche statistique suivie pour identifier empiriquement les compétences. Les principaux résultats des analyses, à savoir la structure relationnelle et hiérarchique des compétences, seront ensuite exposés et enfin, les conséquences possibles de notre travail en matière de politique éducative seront évoquées.

### *II.1.1 Une approche basée sur les relations statistiques entre items*

L'approche centrée sur les items a un double objectif : tester statistiquement la pertinence de la nomenclature des compétences établie par les concepteurs des évaluations nationales et identifier de nouvelles compétences à partir d'associations d'items. D'un point de vue concret, cette approche permet d'utiliser une échelle de mesure commune puisque tous les items présentent le même barème de cotation : 0 pour une réponse erronée, 1 pour la réponse attendue. La question du seuil de réussite mentionnée auparavant ne se pose plus, puisque ce seuil est défini objectivement par la réussite ou l'échec, sans possibilité de situations intermédiaires<sup>42</sup>.

Le point de départ de notre démarche est la production d'une matrice de corrélations qui intègre tous les items de l'épreuve de CE2 (171 items de français et de mathématiques).

---

<sup>42</sup> On mentionnera que cette dichotomie (échec/réussite) n'est pas adaptée à une interprétation diagnostique des évaluations nationales et que les barèmes de cotation de certains items comportent à l'origine d'autres paliers qui permettent une analyse des erreurs des élèves.

Compte tenu du nombre important de corrélations (14535 coefficients de corrélation), nous avons décidé de ne retenir que celles supérieures à +0,20, soit 317<sup>43</sup> ; on notera déjà que cette sélection de corrélations intègre la totalité des items de l'évaluation. Sans grande surprise, on remarque qu'une bonne partie des corrélations (57% d'entre elles) se rapporte à des items appartenant à un même exercice et 86% à la même discipline. Par complémentarité, on observe aussi qu'un nombre non négligeable de corrélations concerne des items provenant d'exercices différents et cela confirme l'observation précédente selon laquelle certaines compétences seraient mesurées de façon imparfaite.

Afin d'organiser l'analyse des relations entre items, une phase préparatoire consiste à étudier individuellement chaque corrélation et à dresser ainsi une cartographie de l'ensemble des situations présentes. Le principe de cette étape préalable est d'identifier des blocs de relations au sein desquelles on retrouve le plus souvent les mêmes items. Cette procédure revêt, de fait, un caractère systématique puisque, pour chaque item, on identifie tous les autres items qui lui sont associés dans les corrélations. Au terme de cette phase, on aboutit à des groupements d'items fortement corrélés entre eux ; certains de ces groupements faisant intervenir un nombre important de corrélations alors qu'à l'opposé, d'autres ne contiennent que deux ou trois corrélations, celles-ci n'ayant pu être rattachées à d'autres regroupements. Au total, 29 blocs d'items ont été identifiés. A titre d'illustration, le tableau 37 présente l'un de ces blocs d'items (il s'agit de 8 items reliés les uns aux autres par 9 corrélations bivariées).

---

<sup>43</sup> Même si ce seuil revêt un caractère arbitraire, l'analyse serait peu pertinente si tous les liens entre items étaient examinés en détail. En effet, dans un grand nombre de cas ces liens sont, soit non significatifs, soit très faibles, et il est alors difficile de commenter et de donner un sens à ces constatations.

Tableau 37 : Exemple d'un regroupement d'items sur la base des corrélations<sup>44</sup> (évaluation CE2, 1999)

Corrélations	Description succincte des items
F67 F62	Dictée préparée d'une phrase : bonne orthographe du mot « jouent » Dictée préparée d'une phrase : bonne orthographe du mot « pendant »
F67 F60	Dictée préparée d'une phrase : bonne orthographe du mot « jouent » Respect de la ponctuation dans une production écrite (majuscule et point présents)
F67 F64	Dictée préparée d'une phrase : bonne orthographe du mot « jouent » Dictée préparée d'une phrase : bonne orthographe du mot « garçons »
F67 F66	Dictée préparée d'une phrase : bonne orthographe du mot « jouent » Dictée préparée d'une phrase : bonne orthographe du mot « filles »
F67 F63	Dictée préparée d'une phrase : bonne orthographe du mot « jouent » Dictée préparée d'une phrase : bonne orthographe du mot « récréation »
F67 F12	Dictée préparée d'une phrase : bonne orthographe du mot « jouent » Repérer des types d'écrits à partir d'extraits de livres différents
F67 F16	Dictée préparée d'une phrase : bonne orthographe du mot « jouent » Répondre à des questions concernant un texte
F64 F62	Dictée préparée d'une phrase : bonne orthographe du mot « garçons » Dictée préparée d'une phrase : bonne orthographe du mot « pendant »
F63 F62	Dictée préparée d'une phrase : bonne orthographe du mot « récréation » Dictée préparée d'une phrase : bonne orthographe du mot « pendant »

Parmi ces 8 items de français, 4 items appartiennent au même exercice (60, 62, 63, 64, 66, 67) et sont censés mesurer la même compétence «*écrire sous la dictée des mots courants, de petites phrases ou de petits textes*». Dans cet exercice, il est demandé aux élèves d'écrire la phrase suivante, dictée par l'enseignant : «*pendant la récréation, les garçons et les filles jouent aux billes.* ». Cette phrase a été auparavant écrite au tableau, les élèves étant invités à mémoriser l'orthographe des mots. La phrase a été ensuite effacée et un exercice différent a été proposé aux élèves ; ce n'est qu'après cet exercice que la phrase a été l'objet d'une dictée. La réussite à la dictée dépend de plusieurs éléments : de compétences orthographiques, de la capacité à se remémorer la graphie des mots préalablement observés et, dans une certaine mesure, d'une capacité d'attention. L'item 60 est supposé rendre compte de la compétence «*copier un mot, une phrase, un texte...* ». Cet item fait partie d'un exercice dans lequel il est demandé aux élèves de recopier des phrases faisant partie d'un court texte présenté sous la forme d'une constellation (les phrases sont écrites dans des bulles et reliées par une flèche à un dessin qui représente le thème du texte) et de les réorganiser sous une forme plus conventionnelle (un texte organisé en paragraphes). L'item 60 n'évalue qu'un aspect de la production des élèves : le respect de la ponctuation (majuscules et points).

<sup>44</sup> Ce sont les relations entre items qui figurent dans la colonne « corrélations » du tableau ; par exemple « F67 F62 » traduit la corrélation entre les items 67 et 62 de français.

L'item 12 fait partie d'un exercice dans lequel on demande aux élèves de repérer des types d'écrits à partir d'un extrait de livres différents (poésie, dictionnaire, livre de recettes de cuisine...). La compétence visée est « *distinguer des textes différents en utilisant des indices extérieurs au texte* ». Enfin, l'item 16 vise la compétence : « *comprendre un texte et montrer qu'on l'a compris* ». Un texte est proposé aux élèves et ceux-ci doivent répondre à des questions concernant ce texte. Dans l'item 16, les élèves doivent entourer 3 bonnes réponses parmi 6 propositions.

Selon notre logique d'analyse, ce regroupement d'items fortement corrélés serait censé rendre compte d'une ou plusieurs compétences. On rappellera que dans les évaluations nationales ces 8 items correspondent à 4 compétences distinctes. L'étape suivante consiste à analyser les liaisons statistiques à l'intérieur de chaque regroupement d'items de façon à mettre à jour des variables latentes pouvant être interprétées comme des compétences, des aptitudes ou des capacités mobilisées par les élèves dans les évaluations nationales. Pour cela, des modèles de mesure ont été estimés à l'aide du logiciel LISREL (LInear Structurel RELationship).

### *II.1.2 L'analyse en variables latentes*

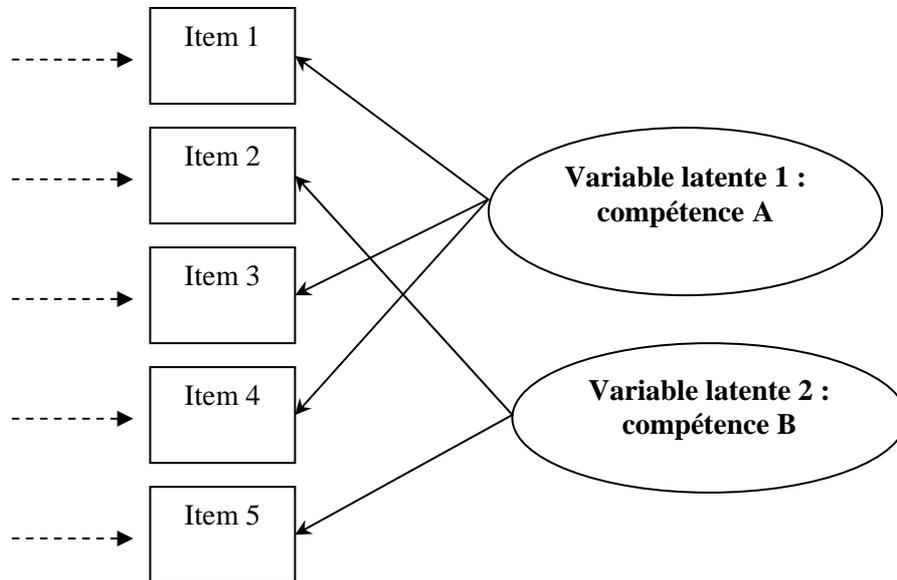
Les modèles de mesure permettent de rendre explicite une dimension latente des phénomènes observés en postulant l'existence de variables inobservées (ou inobservables) qui rendent compte des relations entre variables observées, en l'occurrence ici les résultats aux items. La démarche avec LISREL implique plusieurs étapes (Aish-Van Vaerenbergh, 1997), dans la mesure où le logiciel permet l'estimation de deux types de modèles. Le premier se rapporte aux modèles de mesure qui établissent les relations entre les variables latentes (compétences supposées) et leurs indicateurs (les items), le second concerne le modèle structurel qui détermine les liaisons entre les différentes variables latentes. L'analyse réalisée sur les items de CE2 ne concerne que les modèles de mesure.

Il est d'usage de représenter le modèle de mesure par un diagramme dans lequel les variables latentes sont symbolisées par des ellipses et les variables observées (ou indicateurs de ces variables latentes) sont matérialisées par des rectangles (graphique 20). Les flèches en traits pleins matérialisent l'intensité des relations qui lient chacun des indicateurs (items) à la variable latente (compétence), un coefficient de régression (et son degré de significativité) pour chaque indicateur fournit une indication sur la validité du modèle<sup>45</sup> (Morlaix, 2002).

---

<sup>45</sup> LISREL fournit également un certain nombre de statistiques générales rendant compte de la qualité du modèle de mesure. Le  $\chi^2$  (khi carré) mesure la distance séparant les deux matrices de corrélations, celle théorique et celle portant sur les corrélations estimées. Cet indicateur du  $\chi^2$  doit être le plus petit possible et il doit être aussi proche que possible du nombre de degrés de liberté du modèle. La  $\pi$  value (P value) mesure la probabilité d'obtenir la valeur du  $\chi^2$  correcte. Le modèle est considéré comme recevable si cette valeur est supérieure à 0,5 et on cherchera en outre à ce qu'elle soit la plus élevée possible. Enfin, l'indice RMSEA (Root Mean Square Error

LISREL donne également une mesure de fiabilité du modèle grâce à l'indication de l'erreur de mesure sur chacun des indicateurs (flèches en pointillés sur la figure). Il est donc possible de savoir si les indicateurs retenus sont « robustes » pour mesurer la variable latente.

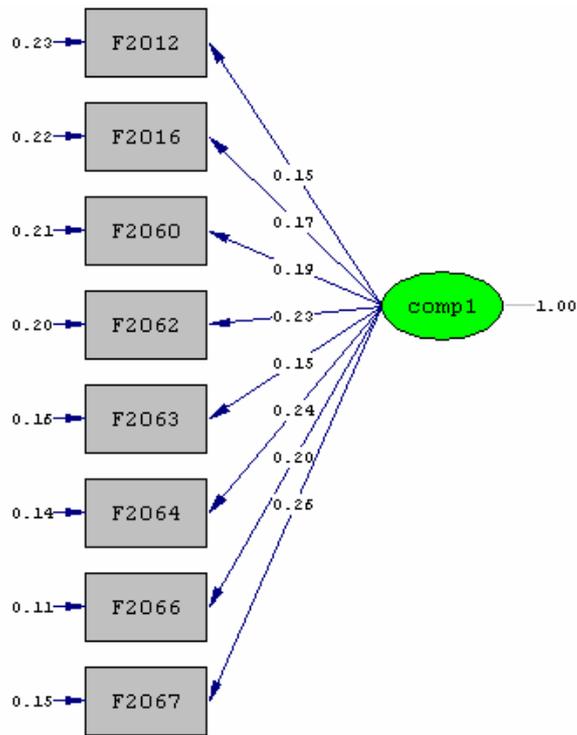


*Graphique 20 : Modèle de mesure entre deux variables latentes et leurs indicateurs*

Des modèles de mesure ont été estimés pour chacun des regroupements d'items ; à titre d'exemple, nous allons présenter l'analyse qui concerne le regroupement d'items évoqué précédemment. Une première étape consiste à définir une variable latente qui rend compte des liens entre les items (graphique 21).

---

of Approximation) se rapporte à la moyenne des résidus du modèle et à leur significativité. Le modèle est acceptable si la valeur de cet indicateur est inférieure ou égale à 0,05.

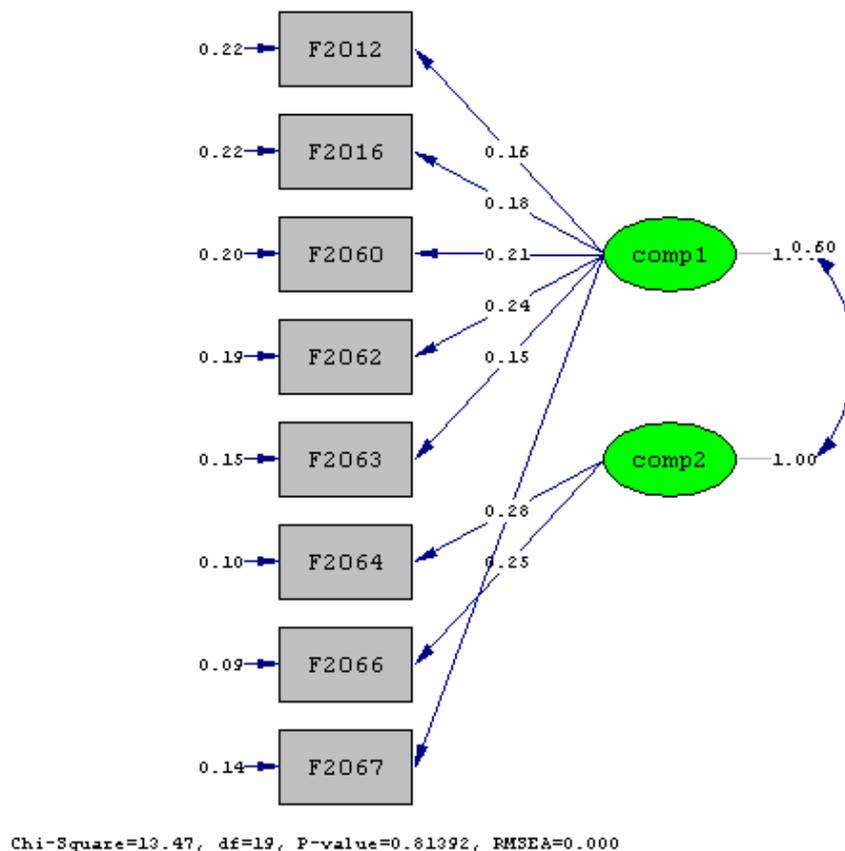


Chi-Square=78.16, df=20, P-value=0.00000, RMSEA=0.066

Graphique 21: Modèle de mesure initial

On peut lire sur les flèches les coefficients des régressions entre chaque item et la variable latente (comp1). Les valeurs des coefficients nous renseignent sur la capacité que peut avoir l’item à rendre compte de la variable latente, celle-ci étant, rappelons le, censée traduire une compétence. C’est pour l’item 67 que la variable latente est la plus prédictive (coefficient de +0,26). D’autres statistiques fournies par LISREL indiquent également que l’item 67 explique 31% de la variance de la variable latente. Les statistiques fournies par LISREL (Khi<sup>2</sup>, P.Value et les résidus) indiquent que ce premier modèle de mesure peut être amélioré et le logiciel suggère qu’une nouvelle variable latente pourrait être introduite dans l’analyse en isolant les items 64 et 66.

Sur la base de cette nouvelle information, un modèle faisant intervenir deux variables latentes est estimé. La première variable est déterminée par les items 12, 16, 60, 62, 63 et 67 (« comp1 » sur le graphique), la seconde par les items 64 et 66 (« comp2 » sur le graphique). Les indicateurs statistiques du nouveau modèle deviennent beaucoup plus fiables (baisse sensible du Khi<sup>2</sup>, augmentation de la P.Value, baisse du R.M.S.E.A.) et le logiciel ne propose alors aucune autre variante pour améliorer la qualité statistique de la représentation. Le graphique ci-après traduit les nouvelles relations entre les items et les deux variables latentes (graphique 22).



Graphique 22 : Modèle de mesure amélioré

On voit aussi apparaître sur le schéma la corrélation entre les deux variables latentes qui affiche une valeur de +0,60, cela témoigne de la proximité des deux compétences. Toutes les relations matérialisées sur le schéma sont en outre significatives. Cette nouvelle configuration demande à être interprétée. Il semble donc que les items 64 et 66 mesurent une compétence spécifique qui rend compte des aptitudes des élèves à orthographier correctement des noms écrits au pluriel (« *garçons* » et « *filles* »). Les 6 autres items (12, 16, 60, 62, 63 et 67) mesureraient pour leur part une compétence différente. Il est difficile d'interpréter rigoureusement ces relations sans avoir recours à des éléments théoriques. En effet, les résultats des élèves aux items de ce premier regroupement semblent dépendre de plusieurs aptitudes : orthographiques, mémorielles, attentionnelles....

Les modèles de mesure estimés sur la totalité des regroupements d'items ont donné lieu à l'identification de 63 variables latentes (Morlaix, Suchaut, 2007a). L'analyse en variables latentes a donc conduit à une recomposition des compétences des élèves à l'entrée au CE2. Parmi les 63 variables mises à jour, 27 d'entre elles (soit 43%) correspondent, souvent de façon partielle, à des regroupements d'items déjà présents dans les évaluations nationales. La correspondance entre les compétences des épreuves et les variables latentes est néanmoins

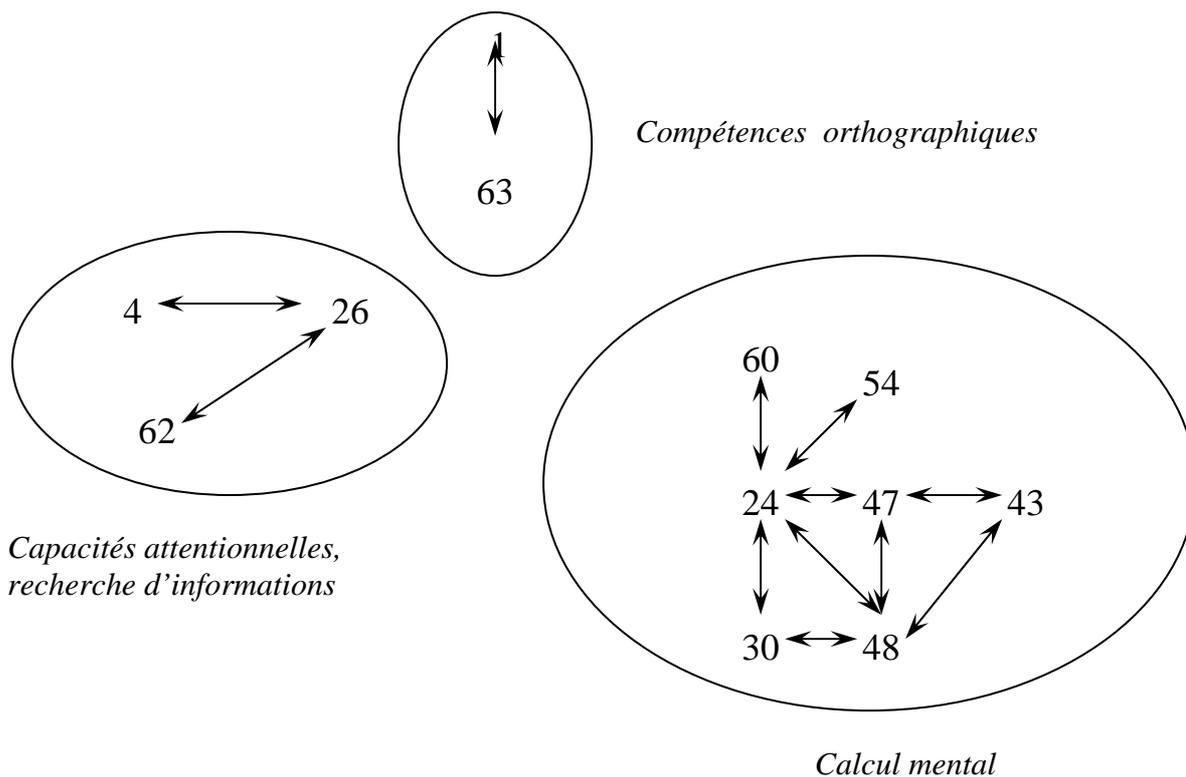
très imparfaite puisque seules 5 variables correspondent exactement à des compétences figurant dans les évaluations nationales (il s'agit uniquement d'exercices de mathématiques).

### *II.1.3 Structure relationnelle et hiérarchique des compétences*

Afin de mettre à jour la structure des acquisitions des élèves, il est à présent indispensable d'étudier les relations statistiques entre les variables latentes. Cela permettra notamment de répondre à une question centrale : quelles compétences rendent le mieux compte des acquisitions globales des élèves ? Pour répondre à cette question, nous avons dans une première étape analysé la matrice de corrélations entre les différentes variables latentes. Les coefficients de corrélation présentent des valeurs allant de 0 à +0,80 (la moyenne sur l'ensemble des corrélations étant de +0,24). Nous avons sélectionné les corrélations les plus importantes (supérieures à +0,70). Cela a permis de dégager trois grands groupes de compétences qui sont représentées sur le graphique 23 (les chiffres correspondent aux numéros associés aux variables latentes). Un premier ensemble est composé des variables « *comp1* » et « *comp63* ». Il s'agit ici clairement de compétences orthographiques qui sont mesurées avec ces deux variables. En effet, la presque totalité des items rendant compte de ces variables latentes concernent deux exercices de dictée. Le deuxième ensemble regroupe les variables « *comp4* », « *comp26* » et « *comp62* ». Cet ensemble d'apparence disparate prend sens quand on l'examine sous l'angle de la psychologie cognitive<sup>46</sup>. Autant les items de français que ceux de mathématiques présents dans ce regroupement évaluent la capacité que peuvent avoir les élèves à rechercher de l'information plus ou moins complexe à partir de supports divers (textes, mots, calendriers, emplois du temps, plans, énoncés de problème). Ce sont donc les capacités attentionnelles des élèves qui sont mises à contribution pour la maîtrise de cette compétence globale. Le troisième ensemble regroupe sept variables latentes. Le lien commun entre ces variables est également clair puisque les items de calcul mental interviennent systématiquement pour chacune d'entre elles.

---

<sup>46</sup> Des collègues psychologues du L.E.A.D. (Laboratoire d'Etude de l'Apprentissage et du Développement - Université de Bourgogne), Pierre Barrouillet et Valérie Camos, ont été associés à cette partie de la recherche, ils ont notamment contribué à donner une signification à certaines variables latentes identifiées par les modèles LISREL



Graphique 23 : Blocs de compétences des évaluations CE2

Les acquisitions des élèves à l'entrée au CE2 s'organisent principalement autour de ces trois compétences qui ne sont pas de même nature. Si l'acquisition des compétences orthographiques dépend principalement d'un enseignement systématique, les deux autres compétences sont davantage associées à des processus plus complexes qui interviennent de façon transversale dans de nombreuses situations d'apprentissage. On devrait s'attendre à ce que ces compétences majeures contribuent fortement à l'explication des différences de réussite entre élèves à l'entrée au CE2. Pour vérifier cela, une régression « pas à pas » a été estimée avec comme variable dépendante le score global moyen de CE2 et comme variables explicatives l'ensemble des compétences mises à jour précédemment. Le tableau suivant présente les résultats en ne conservant que les 12 compétences les plus prédictives. Celles-ci expliquent à elles seules plus de 92% de la variance du score global, soit presque sa totalité.

On constate que les trois compétences les plus prédictives (« comp48 », « comp4 », « comp63 ») appartiennent chacune à un des groupes identifiés auparavant. Ces trois compétences expliquent à elles seules 82% de la variance du score global. Cette analyse de la prédictivité montre que certaines compétences sont au cœur des acquisitions des élèves à l'entrée au CE2. Les habiletés en calcul mental, la capacité à retrouver rapidement des

informations dans des supports variés, la maîtrise de l'orthographe structurent ainsi fortement les résultats des élèves au début du cycle III.

Tableau 38 : Modèle analysant la prédictivité des variables latentes sur le score global de CE2 (régression pas à pas)

Ordre d'introduction	Variables	Coefficients standardisés	t de Student
1	Comp48	+0,23	12,3 ***
2	Comp4	+0,24	14,7 ***
3	Comp63	+0,17	12,4 ***
4	Comp58	+0,13	10,2 ***
5	Comp13	+0,12	10,0 ***
6	Comp17	+0,11	8,9 ***
7	Comp60	+0,13	7,2 ***
8	Comp41	+0,08	6,6 ***
9	Comp16	+0,08	6,8 ***
10	Comp34	+0,08	7,2 ***
11	Comp15	+0,08	6,8 ***
12	Comp3	+0,08	6,5 ***
	Constante	68,76	510,02
	R <sup>2</sup>	0,92	

\*\*\* : significatif au seuil de 1%

Une autre question est d'examiner comment ces compétences les plus prédictives des évaluations nationales se hiérarchisent entre elles<sup>47</sup>. Le tableau 39 expose les relations entre la réussite et l'échec aux 12 compétences les plus prédictives du score global de CE2. Les compétences sont présentées de façon hiérarchisée dans le tableau (de « *comp15* » à « *comp3* »). La réussite à la première compétence « *comp15* » est donc censée être une condition à la réussite de la seconde (« *comp58* »), qui elle-même détermine la réussite à la suivante (« *comp41* »)..., jusqu'à la dernière compétence (« *comp3* ») qui ne peut être acquise que si toutes les autres compétences sont maîtrisées. Dans chaque case, le nombre indiqué représente le pourcentage d'élèves (sur la totalité de l'échantillon) qui réussit une compétence tout en ayant échoué à une autre (il s'agit de résultats issus de tableaux croisés). Ainsi 4% des élèves réussissent « *comp58* » tout en échouant à « *comp15* » ; de même, 3,3% des élèves réussissent « *comp41* » et échouent à « *comp15* ». Si la structure hiérarchique des

<sup>47</sup> Une étape préalable à l'analyse est de rendre compte systématiquement de la maîtrise ou de l'échec à une compétence, ce qui suppose une transformation de l'échelle de mesure, on passe alors d'une échelle d'intervalle (qui varie selon la compétence considérée) à une échelle nominale, représentée de façon dichotomique : échec ou réussite. Il se pose alors la question du seuil à partir duquel on va considérer que l'élève a acquis la compétence visée; il n'existe pas de réponse parfaite mais nous avons considéré qu'un score supérieur à 75% de réussite soit associé à la maîtrise de la compétence.

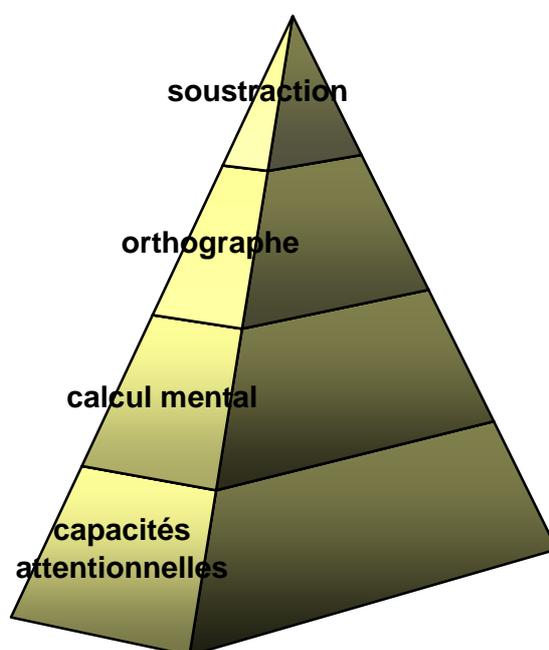
compétences se vérifie, on devrait donc s’attendre à ce que les chiffres du tableau présentent des valeurs faibles. Cela est effectivement le cas, même si certaines valeurs s’écartent de la règle générale. Ainsi, si la maîtrise de « *comp15* » semble bien être une condition indispensable à la réussite à toutes les autres compétences (la première ligne du tableau présente des valeurs très faibles), la relation entre « *comp13* » et « *comp16* » ne semble pas vérifiée dans la mesure où plus de 20% des élèves réussissent « *comp13* » tout en échouant à « *comp16* ». Globalement la structure de dépendance entre compétences se vérifie dans la majorité des cas ; les cases grisées dans le tableau représentent ces situations (pourcentages d’élèves inférieurs ou voisins de 10%).

Tableau 39 : Relations entre l’échec et la réussite aux compétences de CE2  
(% d’élèves)

EHEC	REUSSITE											
	comp15	comp58	comp41	comp17	comp16	comp13	comp34	comp4	comp48	comp60	Comp63	comp3
comp15		<b>4,0</b>	3,3	2,1	3,0	2,4	2,8	0,0	0,4	0,1	0,1	0,6
comp58			<b>8,5</b>	6,7	7,9	7,6	8,0	1,2	1,3	0,7	0,7	1,2
comp41				<b>14,6</b>	14,6	14,0	13,7	2,8	3,1	1,6	1,3	1,6
comp17					<b>19,5</b>	17,7	16,1	4,5	4,9	3,6	2,5	3,1
comp16						<b>20,6</b>	16,7	6,7	6,7	4,3	4,8	3,1
comp13							<b>16,5</b>	4,5	6,1	4,2	2,2	3,3
comp34								<b>6,9</b>	7,2	4,6	5,0	4,9
comp4									<b>11,2</b>	7,9	6,7	8,9
comp48										<b>5,8</b>	7,5	6,1
comp60											<b>8,6</b>	7,5
comp63												<b>9,8</b>

Les compétences semblent donc assez bien hiérarchisées dans l’ensemble ; la compétence qui se trouve au sommet de cette hiérarchie correspond à la maîtrise de la soustraction (variable « *comp3* »). Ceci signifie que l’acquisition de la technique opératoire de la soustraction est un processus qui nécessite de la part des élèves diverses capacités et habiletés préalables. En effet, très peu d’élèves réussissent dans cette dimension des acquisitions sans maîtriser toutes les autres compétences qui figurent dans le tableau. La seconde compétence dont la maîtrise dépend de nombreuses autres est de nature orthographique (« *comp63* ») ; viennent ensuite deux compétences dans lesquelles le calcul mental intervient majoritairement, soit directement (« *comp48* »), soit indirectement par le biais d’autres compétences (« *comp60* »). La compétence qui se situe au plus bas niveau de cette structure hiérarchique correspond à des items évaluant la connaissance des unités de mesure. En fait, c’est davantage la compréhension de consignes simples qui semble être la compétence véritablement visée dans cet exercice. Le graphique 24 symbolise la structure pyramidale des dernières compétences du même tableau.

Cette structure pyramidale des compétences peut fournir des indications didactiques et pédagogiques pour l'enseignement au cycle II, en amont du CE2. Ces indications concernent principalement les contenus d'enseignement et leur programmation dans le temps, ils s'appuient sur le double constat suivant établi à l'entrée du CE2 : certaines compétences sont difficilement accessibles à l'ensemble des élèves, certaines compétences sont essentielles à l'acquisition d'autres compétences



*Graphique 24: Structure pyramidale des compétences des élèves au CE2*

Suite à ces deux constats, on peut déjà suggérer que la définition des programmes scolaires mette l'accent de façon explicite sur les compétences définies comme essentielles précédemment, cela nécessite aussi que des situations d'enseignement concrètes puissent être envisagées pour l'acquisition de ces compétences, ce qui n'est pas immédiat quand elles sont transversales. Il est ensuite utile de s'interroger sur la planification de l'enseignement de ces compétences en prenant en compte la structure hiérarchique d'apprentissage telle qu'elle apparaît sur le graphique précédent, c'est-à-dire en respectant le fait que certains apprentissages ne peuvent se réaliser que si certains autres sont déjà maîtrisés.

#### II.1.4 Conclusion sur les acquisitions à l'entrée au cycle III

Un premier constat qui se dégage des résultats est l'aspect transversal des acquisitions scolaires. La vision institutionnelle des évaluations nationales masque cet aspect par le regroupement, parfois artificiel, de certains items au sein d'un même exercice alors que l'approche statistique permet d'identifier des compétences de nature variée. De façon complémentaire, des items provenant d'exercices variés, voire de disciplines différentes ont été identifiés comme mesurant une seule et même compétence. Ce premier constat peut avoir des implications évidentes quant à l'utilisation des évaluations nationales sur le plan pédagogique dans une perspective diagnostique. Cela permet en effet de mieux cerner les processus d'apprentissage des élèves qui sont mobilisés lors de la réalisation d'un exercice et donc de fournir des éléments pour organiser des situations pédagogiques ciblées sur l'acquisition de certaines compétences.

Parmi la soixantaine de compétences identifiées, certaines ont un statut particulier dans le sens où elles rendent compte, à elles seules de la presque totalité des différences d'acquisition des élèves à l'entrée au cycle III. Ces compétences concernent trois domaines : l'orthographe, l'attention et la recherche d'information, le calcul mental<sup>48</sup>. On rejoint ici un constat réalisé par la D.E.P. sur la base du panel 1997<sup>49</sup> selon lequel l'épreuve d'attention partagée administrée à l'entrée au CP s'est révélée fortement corrélée avec la réussite générale à l'évaluation des CE2, deux années plus tard ; « *la capacité d'attention pourrait donc bien constituer un facteur majeur de la réussite* » (M.E.N., 2000).

Notre travail a aussi permis d'approfondir les relations entre les compétences en mettant en évidence une structure hiérarchique, certaines compétences étant indispensables à l'acquisition d'autres compétences. Sans entrer ici dans le détail de cette structure, signalons que la compétence de CE2 qui se trouve au sommet de la hiérarchie correspond à la maîtrise de la soustraction qui apparaît être un processus nécessitant de la part des élèves diverses capacités et habiletés préalables. La compétence se situant au plus bas niveau de la structure hiérarchique concerne des exercices qui visent la compréhension de consignes simples (alors que les évaluations nationales ont répertorié ces exercices dans le domaine de la mesure). En prenant en compte la structure hiérarchique d'apprentissage effective des élèves, c'est-à-dire en respectant le fait, trivial à présent, que certains apprentissages ne peuvent se réaliser que si d'autres déjà maîtrisés (Bloom, 1979), les résultats évoqués précédemment débouchent naturellement sur un questionnement lié à la planification des activités d'enseignement, et plus largement au contenu des programmes scolaires. Il est aussi essentiel de s'interroger

---

<sup>48</sup> Ces deux derniers éléments pouvant intervenir dans des situations d'apprentissage très variées qui dépassent largement l'approche disciplinaire.

<sup>49</sup> Elèves entrés au CP en 1997.

aussi sur le temps que l'on doit consacrer à l'acquisition des différentes compétences, sachant que le temps alloué est très lié aux progrès des élèves et que celui-ci s'inscrit, pour l'enseignant, dans une logique d'arbitrage entre des activités diverses et concurrentielles (Suchaut, 1996a ; Morlaix, 2000). En outre, certains élèves ont besoin de plus de temps que d'autres pour acquérir certaines notions. Les recommandations pédagogiques doivent donc aussi tenir compte de la diversité des élèves et de leur rythme d'apprentissage spécifique.

## ***II.2 Evolution des compétences au cycle III***

La même méthodologie a été adoptée sur la même cohorte d'élèves en ce qui concerne les évaluations de début 6<sup>ème</sup>, des analyses sous LISREL ont été conduites à partir de l'ensemble des blocs d'items identifiés afin de mettre en évidence les variables latentes qui structurent les corrélations. Nous ne détaillerons pas les analyses qui suivent exactement la même logique que pour les items des épreuves de CE2<sup>50</sup>. On dénombre, à l'entrée au collège 112 variables, soit 49 de plus qu'au niveau du CE2 alors que les épreuves comportent un nombre assez comparable d'items dans les deux niveaux scolaires considérés. Le fait que l'on relève pour la 6<sup>ème</sup> près du double de compétences qu'en CE2 peut s'expliquer de deux manières. La première est de nature statistique ; du fait du plus grand nombre d'items pris en compte dans les corrélations (2,4 fois plus) il n'est pas surprenant que cette augmentation d'items conduise à identifier davantage de compétences. La seconde piste d'explication est davantage de l'ordre de la supposition mais elle relève aussi du bon sens. Les élèves étant plus âgés en 6<sup>ème</sup>, ils ont bénéficié d'une expérience scolaire plus longue, accumulé plus de connaissances et développé également un plus grand nombre de compétences. Dans les documents d'accompagnement du Ministère<sup>51</sup>, le nombre de compétences est de 32 (27 en CE2) en mathématiques et de 20 en français<sup>52</sup> (15 en CE2), il y a donc bien également dans l'approche institutionnelle une tendance à ce que les compétences soient plus nombreuses en sixième qu'au CE2.

Sur l'ensemble des compétences mises à jour, on remarquera des variables latentes qui font intervenir des items successifs appartenant à un même exercice (soit la totalité des items de l'exercice, soit une partie seulement des items), elles correspondent à des compétences répertoriées par les concepteurs des évaluations. On relève aussi évidemment des variables latentes qui traduisent des relations entre items provenant d'exercices différents, voire de disciplines différentes.

---

<sup>50</sup> On pourra se reporter au document qui décrit l'ensemble de la recherche (Morlaix, Suchaut, 2007a).

<sup>51</sup> Pour les évaluations sixième de 2002, les concepteurs distinguent les compétences des composantes, les secondes étant des déclinaisons plus précises des premières, de façon plus marquée en mathématiques qu'en français.

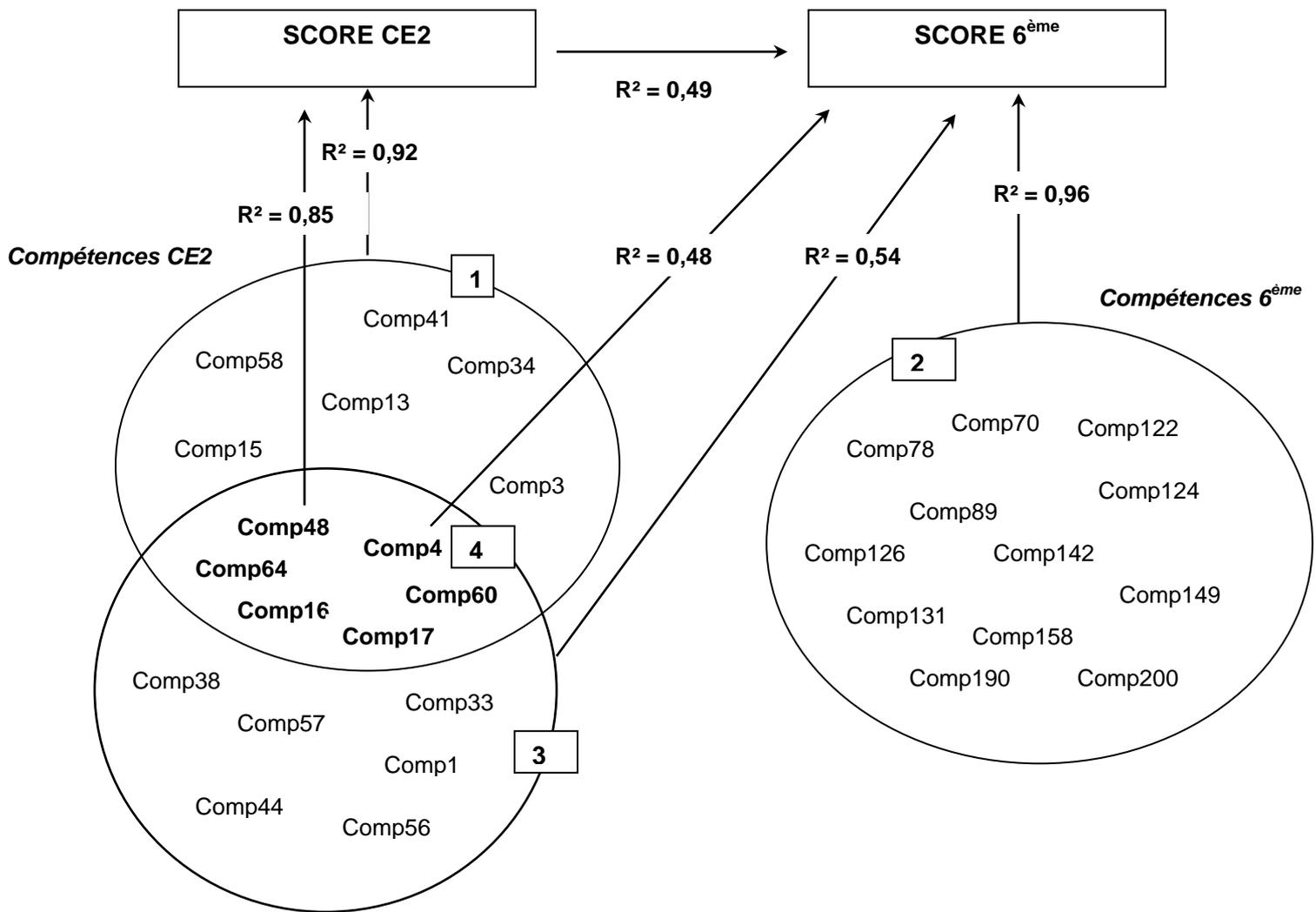
<sup>52</sup> Dans cette discipline, il s'agit du nombre de composantes, c'est en effet ce niveau de définition qui est le plus comparable aux compétences de CE2 dans la même discipline.

De la même manière qu'au CE2, les relations statistiques qui lient ces variables ont été étudiées pour pouvoir dégager des grandes structures dans les apprentissages des élèves à ce niveau de la scolarité. Un modèle de régression pas à pas a été estimé pour déterminer quelles sont les variables latentes les plus prédictives du score global. La variable la plus prédictive explique à elle seule plus des trois-quarts de la variance du score global ( $R^2 = 0,76$ ). Cette variable latente comporte 19 items (6 de français et 13 de mathématiques) et traduit majoritairement des habiletés en calcul mental. Le modèle de régression « pas à pas » retient ensuite une seconde variable (le pourcentage de variance expliquée passe à 83%) qui est composée de 2 items de mathématiques (calcul mental) et de 7 items de français qui proviennent de champs variés. L'ajout successif des autres variables permet très rapidement d'obtenir un pouvoir explicatif très élevé puisque dès la quatrième variable, on explique les neuf dixièmes de la variance du score global. L'ensemble des 12 variables produisant un coefficient de détermination ( $R^2$ ) de 0,95 (soit 95% de la variance totale du score global à l'entrée en sixième).

Le graphique suivant résume les constats concernant la dimension prédictive des compétences de CE2 dans une double perspective. D'un point de vue transversal tout d'abord, figurent sur le schéma les compétences les plus prédictives de chacun des niveaux scolaires étudiés. D'un point de vue longitudinal, le graphique présente les compétences de CE2 les plus prédictives du score de 6<sup>ème</sup><sup>53</sup>. Enfin, sont identifiées en caractères gras les compétences de CE2 qui sont prédictives, à la fois du score de CE2 et de 6<sup>ème</sup>. A partir de ces éléments, on peut poursuivre l'interrogation sur les compétences les plus utiles à développer lors du parcours des élèves à l'école primaire. Pour cela, nous allons décrire en détail les différentes relations observées.

---

<sup>53</sup> Précisons que les analyses longitudinales portent sur un échantillon réduit d'élèves, à savoir les 394 pour lesquels nous disposons à la fois des évaluations de CE2 et de 6<sup>ème</sup>.



Graphique 25 : Relations statistiques entre compétences de CE2 et de 6<sup>ème</sup>

Les analyses ont en premier lieu identifié un ensemble de compétences qui sont les plus pertinentes pour expliquer les différences d'acquisitions entre élèves à l'entrée au cycle III ; il s'agit du premier ensemble noté 1 sur le graphique. Ces douze compétences, nous l'avons vu auparavant, expliquent presque la totalité de la variance du score de CE2 ( $R^2 = 0,92$ ). De même, un ensemble de douze compétences de 6<sup>ème</sup>, noté 2 sur le graphique, rend compte à lui seul de 96% de la variance du score global à ce même niveau d'enseignement. Un premier constat est que pour expliquer les différences de réussite entre élèves aux évaluations nationales (et donc, d'une certaine manière les différences de niveaux d'acquisitions scolaires attendus), il suffit de mesurer la maîtrise de quelques compétences clés. Un troisième ensemble de compétences, noté 3 sur le graphique, traduit en quelque sorte l'évolution des apprentissages des élèves au cours du cycle III : il s'agit des compétences de CE2 qui expliquent le plus les différences de scores entre les élèves à l'entrée en 6<sup>ème</sup>. Le pourcentage

de variance expliquée associé à cette relation est de 54% ( $R^2 = 0,54$ ), ce qui signifie que les 12 compétences rendent compte à elles seules de plus de la moitié de la variance des acquis des élèves au début du collège, soit davantage que l'ensemble des items de l'évaluation de CE2 ( $R^2$  de 0,49). En termes de prédiction de la réussite scolaire au cours du cycle III, ces compétences présentent évidemment un grand intérêt, c'est bien leur maîtrise à l'entrée de ce cycle qui assure, en partie, un bon niveau d'acquisitions à l'entrée au collège.

Mais c'est le sous-ensemble de compétences, noté 4 sur le graphique qui attire particulièrement l'attention. Il s'agit des compétences de CE2 qui sont communes aux deux relations précédentes (prédiction du score de CE2 et prédiction du score de 6<sup>ème</sup>). Ces 6 compétences sont donc, à la fois utiles pour prédire la réussite à l'entrée au CE2 ( $R^2$  de 0,85) et en sixième ( $R^2 = 0,48$ ). On s'interroge alors sur la possibilité qu'ont les élèves à mobiliser les mêmes compétences à moyen terme tout au long d'un cycle scolaire, mais aussi sur le statut de ces compétences. On retrouve 4 compétences (parmi les 6) qui ont été identifiées préalablement comme des compétences (ou capacités) qui structuraient le mieux les apprentissages des élèves au niveau du CE2. Deux d'entre elles (« comp48 » et « comp60 ») concernent le calcul mental, la variable « comp4 » étant associée à des capacités attentionnelles et la variable « comp64 » à des compétences de nature orthographique. Les deux autres variables ajoutent deux dimensions supplémentaires des apprentissages, à savoir : des compétences en matière de reproduction de figures géométriques (« comp16 ») et plus particulièrement la prise de conscience de la symétrie ; des compétences dans le domaine de la production d'écrits (« comp17 ») et plus spécifiquement la capacité à pouvoir prélever des indices dans un court texte pour les réinvestir dans leur production (situation d'induction).

On peut s'interroger plus profondément sur les relations entre les compétences de CE2 les plus prédictives et celles de 6<sup>ème</sup> en estimant un modèle structurel sous LISREL. Le schéma ci-dessous (extrait du précédent), représente les variables à mettre en relation dans les estimations statistiques. Le modèle structurel doit permettre d'obtenir deux types de résultats supplémentaires. En premier lieu, l'analyse fournit un nouveau niveau d'agrégation des variables en les regroupant selon leurs liens statistiques. On obtient ainsi des blocs de compétences pour les deux niveaux scolaires (CE2 et 6<sup>ème</sup>), ce qui va nous permettre, dans un second temps, de mettre en relation ces blocs et de mieux comprendre comment les acquisitions des élèves se structurent dans le temps. Le graphique 26 rend compte des deux étapes de l'analyse.

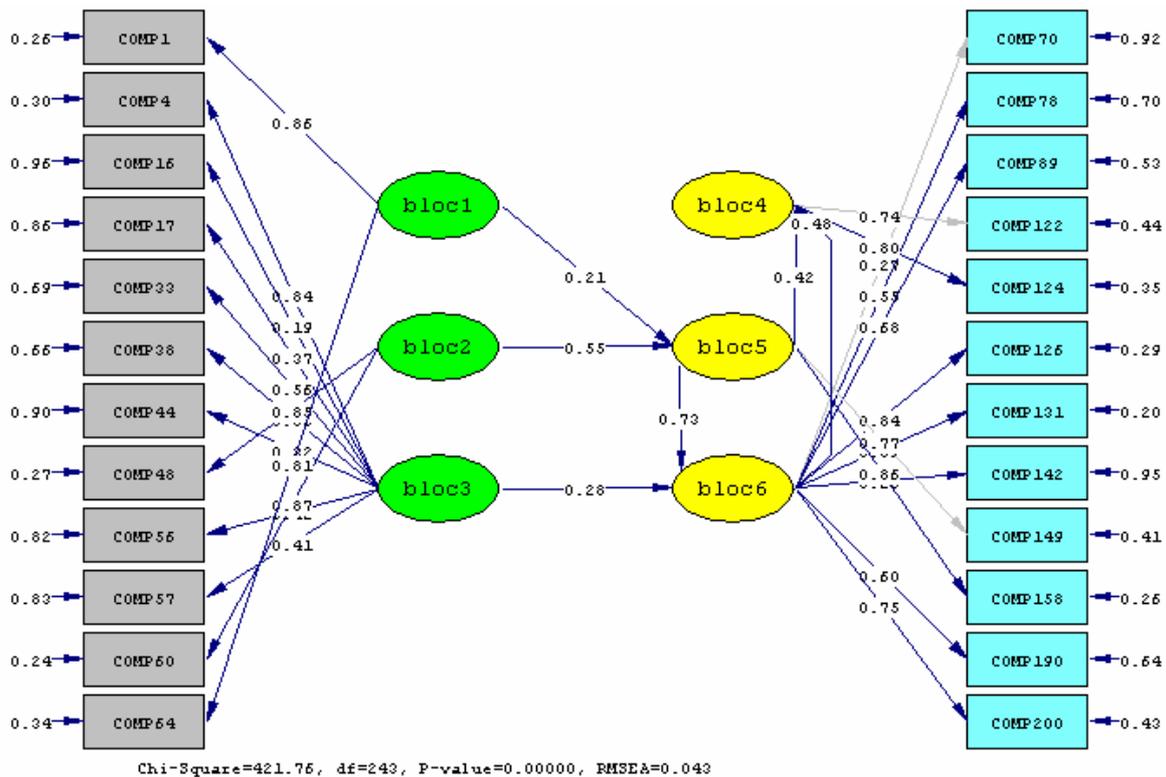
Au niveau des compétences de CE2, les estimations effectuées conduisent à isoler 3 ensembles de compétences qui correspondent à trois nouvelles variables latentes. Le premier (« bloc 1 »), regroupe les variables « comp1 » et « comp64 », il correspond exactement au regroupement effectué précédemment à partir de la matrice de corrélations de l'ensemble des

variables de CE2. Il s'agit, clairement d'un bloc représentant des compétences orthographiques. Le second ensemble (« bloc 2 ») regroupe les variables « comp48 » et « comp60 » mesurant des compétences en calcul mental. On notera que ces deux variables appartiennent au regroupement intitulé « calcul mental » dans les analyses menées au CE2. Le troisième ensemble (« bloc3 ») rassemble les 8 compétences restantes ; c'est un ensemble composite faisant intervenir majoritairement des capacités attentionnelles, la compréhension, la production d'écrits. Dans ce dernier ensemble, les compétences apparaissent comme des indicateurs de la variable latente (« bloc3 ») dont le poids est très variable ; par exemple le coefficient de régression associé à « comp4 » vaut 0,84 (liaison très forte) alors que celui associé à « comp16 » ne vaut que 0,19 (liaison plus faible).

Au niveau de l'évaluation 6<sup>ème</sup>, les analyses mettent également en évidence 3 ensembles de compétences. Le « bloc4 » regroupe les compétences « comp122 » et « comp124 » ; cette dernière est composée très majoritairement d'items faisant intervenir du calcul (la plupart du temps du calcul mental) ; la première comporte des items de calcul et de géométrie. Notons que la liaison statistique entre cette variable (« comp122 ») et la variable latente n'est pas significative. Un second ensemble « bloc5 » associe les deux variables « comp149 » et « comp158 » qui toutes les deux rassemblent des items de calcul mental et de connaissance de la numération. On relèvera par ailleurs que le coefficient associé à « comp149 » n'est pas statistiquement significatif<sup>54</sup>. Le troisième ensemble « bloc6 » regroupe un très grand nombre d'items, via la présence de 8 variables. Il n'est pas possible de dégager une dimension précise des acquis des élèves à travers cet ensemble composite puisque sont présents des items très variés mesurant aussi bien la compréhension, la capacité à effectuer des calculs ou encore la production d'écrits.

---

<sup>54</sup> Certaines relations entre « compétences » et variables latentes deviennent non significatives dans le modèle structurel d'ensemble, alors qu'elles l'étaient dans les modèles de mesure estimés séparément (respectivement en CE2, puis en 6<sup>ème</sup>).



Graphique 26: Analyse structurelle entre compétences de CE2 et compétences de 6<sup>ème</sup>

Les 6 ensembles dégagés par cette analyse structurelle n'ont pas tous la même configuration dans la mesure où certains apparaissent comme très homogènes avec peu d'indicateurs (« bloc1 », « bloc2 », « bloc4 », « bloc5 »), alors que deux autres (« bloc3 » et « bloc6 ») témoignent d'une plus grande hétérogénéité dans les compétences et/ou habiletés mobilisées par les élèves. On précisera que ces deux derniers blocs ont été constitués en dernier dans les analyses, une fois que les liaisons les plus fortes ont été dans un premier temps isolées ; ce sont donc en quelque sorte des regroupements « par défaut » des variables restantes, même si des corrélations entre les variables composant ces blocs existent. Compte tenu de ce contexte lié à la logique de l'analyse, l'interprétation du contenu de ces deux blocs (« bloc3 » et « bloc6 ») est rendue d'autant plus délicate du fait de la plus grande hétérogénéité des compétences.

Le modèle structurel représenté sur le graphique 14 permet également de visualiser, grâce aux flèches, les relations entre les blocs de compétences identifiés aux deux niveaux scolaires selon une perspective longitudinale et ceci est d'une importance capitale pour notre questionnement. Une remarque d'ordre général tout d'abord concerne le degré de dépendance des blocs de compétences ; on relève à cet égard des configurations différentes entre le CE2 et la 6<sup>ème</sup>. Les blocs de CE2 n'entretiennent pas de relations directes entre eux, ce qui conduit à

penser que les compétences les plus représentatives des acquis des élèves, sont à ce niveau de la scolarité, relativement indépendantes les unes des autres. Ce serait ainsi le cas des acquisitions en orthographe et des habiletés en calcul mental. Ce constat n'est plus valable pour les acquisitions à l'entrée en 6<sup>ème</sup> puisque les 3 blocs de compétences essentielles entretiennent entre eux des fortes relations statistiques. La variable latente « bloc 4 » (compétences liées au calcul) dépend statistiquement, des deux autres regroupements « bloc5 » (calcul mental et numération) et « bloc6 » (regroupement composite) ; ces deux blocs de compétences apportant chacun une contribution presque équivalente à la variable « bloc5 ». Cette même variable (« bloc5 ») détermine fortement la variable latente « bloc6 » (coefficient de régression de 0,73). On peut donc, au niveau des acquisitions à l'entrée au collège, hiérarchiser les blocs de compétences identifiés par l'analyse sous LISREL ; il apparaît que les habiletés en calcul (calcul mental et connaissance de la numération) expliquent fortement les différences de compétences entre les élèves dans les autres domaines d'apprentissage. Cette affirmation est confortée quand on examine les liens entre les blocs de CE2 et de 6<sup>ème</sup> puisque la variable « bloc5 » dépend, à des degrés divers, des variables « bloc1 » (orthographe) et « bloc2 » (calcul mental).

En résumé, l'analyse structurelle des acquisitions des élèves au cycle III permet déjà de dégager les constats suivants. Au cours du cycle III, les acquisitions des élèves entretiennent entre elles des liens plus importants, liens matérialisés par les fortes corrélations entre les items, compétences et regroupement de compétences des épreuves d'évaluation de 6<sup>ème</sup>. Les compétences identifiées au niveau des évaluations de 6<sup>ème</sup> sont plus nombreuses qu'au CE2 ; ceci témoignerait de la diversification des acquis des élèves au fil du temps. Certaines compétences, capacités ou habiletés comme calculer mentalement semblent déterminer fortement l'évolution des acquis des élèves dans le temps dans des dimensions diverses.

C'est donc un double phénomène qui est observé : diversification et renforcement des relations entre les acquisitions des élèves, avec toutes les conséquences qui en découlent quant à la réussite ou à l'échec des élèves à l'école élémentaire. Dans une certaine mesure la réussite (ou l'échec) prend, au fil du temps, une plus grande cohérence et limite ainsi la possibilité que certains élèves en échec dans un domaine puissent compenser par la réussite dans un autre domaine. Ce constat est confirmé par des analyses complémentaires effectuées sur le panel 1997 pour lesquelles on observe ce même phénomène puisque les évaluations à l'entrée au CP ne permettent d'identifier qu'un nombre réduit de compétences (qui correspondent quasiment aux grands domaines définis par les concepteurs des épreuves. EN outre les corrélations entre ces domaines apparaissent bien plus faibles que celles relevées dans les épreuves du CE2 et à fortiori de 6<sup>ème</sup> .

Nous allons aborder à présent la dernière étape de notre analyse énoncée précédemment en examinant quels liens existent entre les acquisitions des élèves, leurs capacités cognitives et leur origine sociale.

### **III Les liens entre capacités cognitives, compétences scolaires et origine sociale**

Les recherches françaises qui ont examiné les déterminants de la réussite scolaire à l'école élémentaire distinguent généralement deux catégories de facteurs. La première a trait aux caractéristiques personnelles des élèves, celles-ci étant représentées par des variables socio-démographiques et scolaires. Outre le fait de pouvoir intégrer dans une même analyse un grand nombre de facteurs, l'intérêt de ces recherches est de mobiliser une approche longitudinale qui permet de raisonner en termes de progrès réalisés par les élèves au cours d'une période donnée, généralement une année scolaire. Les acquisitions des élèves sont toutefois appréhendées de façon globale, avec au mieux des découpages par discipline ou par domaine et ont surtout une valeur instrumentale pour les analyses. Or, ces acquisitions englobent simultanément des compétences scolaires et des habiletés cognitives sans que l'on puisse clairement les distinguer dans les épreuves administrées aux élèves. L'objectif de cette section est précisément de se centrer sur les capacités cognitives des élèves et d'étudier les relations qu'elles entretiennent avec les facteurs les plus prédictifs de la réussite scolaire, à savoir le niveau de compétences initial et l'origine sociale. Les questions abordées portent plus précisément sur le rôle des capacités cognitives dans les apprentissages des élèves à l'école élémentaire, cette problématique ayant peu été abordée dans le champ des sciences de l'éducation. Ces questions se déclinent globalement en deux points.

Un premier point est de s'interroger sur le lien entre l'évolution des performances scolaires et les capacités cognitives des enfants. Il est intéressant de savoir si le niveau de performance scolaire dépend de la même manière des capacités cognitives à différents niveaux de la scolarité. Il est en effet possible que les apprentissages fondamentaux dépendent de ces capacités de manière plus étroite que les acquisitions ultérieures qui, elles, pourraient davantage dépendre de l'acquisition de règles ou de connaissances factuelles plus sensible aux effets des milieux éducatifs, scolaires et sociaux. Une fois la contribution spécifique des capacités cognitives aux progressions des élèves établie, un second point portera sur la place relative des facteurs sociaux, l'origine sociale de l'élève plus exactement, comparée à celles des capacités cognitives pour expliquer les différences de progressions entre élèves. Les résultats d'une analyse portant sur la relation entre origine sociale, capacités cognitives et progrès des élèves seront alors proposés. Après avoir défini la manière dont les capacités cognitives ont été appréhendées et mesurées dans notre travail, nous présenterons les données mobilisées avant d'exposer les résultats relevant des deux points énoncés précédemment.

### *III.1 Mémoire de travail et capacités cognitives*

La mémoire humaine est traditionnellement conçue comme constituée de la mémoire à long terme, dont la capacité semble quasi illimitée et qui permet de maintenir des informations sur des durées extrêmement prolongées, et de la mémoire à court terme, dont la capacité est limitée, et qui permet de maintenir des traces actives pendant une durée brève. Nous nous centrerons ici sur la mémoire à court terme. Elle a longtemps été considérée comme un réceptacle passif, uniquement utilisé pour stocker un nombre limité d'informations. Cependant, un certain nombre de résultats expérimentaux ont montré qu'elle servait aussi de lieu de traitement de l'information. Les tâches cognitives quotidiennes (lire un article dans le journal, calculer la somme due pour un repas au restaurant...) impliquent souvent le maintien temporaire d'information alors même que des opérations mentales sont effectuées simultanément. Ainsi, le concept de mémoire de travail (Atkinson & Shiffrin, 1968 ; Baddeley & Hitch, 1974) a peu à peu remplacé celui de mémoire à court terme. Il est aujourd'hui utilisé pour rendre compte du système ou des mécanismes qui sous-tendent le maintien de l'information pertinente durant la réalisation de tâches cognitives.

Les premiers modèles de mémoire de travail (Daneman & Carpenter, 1980 ; Case, 1985) reposaient sur une métaphore d' « espace de travail ». D'après ces modèles, la mémoire de travail aurait deux limitations : spatiale (plus la quantité d'informations à retenir augmente et plus la charge cognitive sera importante), et temporelle (plus les traitements sont longs, et plus l'information sera difficile à retenir). Cependant, les modèles les plus récents (Engle, Cantor & Carullo, 1992 ; Just & Carpenter, 1992 ; Cowan, 1995) relient le fonctionnement de la mémoire de travail à l'attention. Ainsi, la capacité de la mémoire de travail n'est plus considérée comme dépendante de l'espace nécessaire aux traitements, mais comme la quantité de ressources attentionnelles disponibles pour maintenir actives les informations utiles au traitement.

A partir de la métaphore spatiale des premiers modèles de mémoire de travail, des tâches ont été créées visant à mesurer la taille de l'espace que constitue la mémoire de travail. Ces tâches ont été appelées tâches d'empan de mémoire de travail. Elles ont la particularité de nécessiter le maintien en mémoire de listes de mots, de chiffres, ou de lettres tout en effectuant une tâche secondaire. On parle alors de tâches d'empan complexe car deux activités cognitives sont à effectuer en parallèle. Un exemple de tâche de mémoire de travail est l'empan d'opérations (« operation span » de Turner et Engle, 1989). Dans cette tâche, le sujet doit mémoriser des listes de mots tout en effectuant des opérations arithmétiques intercalées entre chaque mot présenté (i.e., MAISON,  $(4 \times 3) / 2 = ?$ , CHIEN,  $(5 + 7) \times 3 = ?$ , rappel correct : MAISON, CHIEN). L'empan correspond à la plus longue liste dont le sujet peut rappeler tous les mots dans l'ordre de présentation sans avoir commis d'erreur sur les opérations. Un autre test très

utilisé est l'empan de lecture (« reading span test » de Daneman et Carpenter, 1980). Il consiste à faire lire des séries de phrases présentées une à une au sujet, celui-ci devant mémoriser le dernier mot de chacune des phrases lues. L'empan est ici le nombre maximal de phrases dont le sujet peut rappeler le dernier mot après leur lecture. Dans ces tâches, il s'agit donc de maintenir des informations en mémoire afin de pouvoir les rappeler ultérieurement, tout en réalisant en même temps des activités cognitives complexes (résolution d'opérations, lecture de phrases).

Il a été montré qu'il existe une corrélation élevée entre l'empan de mémoire de travail mesuré grâce à ces tâches classiques et les activités cognitives complexes. Ainsi, l'empan de mémoire de lecture, mais aussi l'empan d'opérations, prédisent les performances en compréhension en lecture mieux que ne le font les tâches d'empan simple en mémoire à court terme<sup>55</sup> (Daneman, Carpenter, 1983). Les empan de mémoire de travail sont aussi de bons prédicteurs des performances en arithmétique (Hitch, Towse, Hutton, 2001) ou encore en raisonnement (Barrouillet, Lecas, 1999). De manière générale, les tâches de mémoire de travail se sont avérées d'excellents prédicteurs des performances scolaires chez l'enfant et chez l'adolescent (Gathercole, Pickering, 2000). Il a donc été suggéré que les tâches d'empan de mémoire de travail évaluent les ressources cognitives des sujets, ressources qui seraient particulièrement sollicitées dans la réalisation des tâches cognitives de haut niveau telles que la compréhension, l'apprentissage d'habiletés complexes, ou le raisonnement (Gavens, Camos, 2006). Cependant, l'existence d'un lien entre les mesures d'empan obtenues grâce aux tâches classiques de mémoire de travail et les acquisitions scolaires apparaît quelque peu triviale. En effet, toutes les tâches classiques de mesure d'empan de mémoire de travail impliquent des activités elles-mêmes complexes (lecture, calcul) et fortement liées aux performances qu'elles prédisent (compréhension de textes, arithmétique, etc...). Il n'est par exemple pas très étonnant que l'empan de lecture soit un bon prédicteur des performances en compréhension de lecture, puisque les deux doivent nécessiter des processus cognitifs communs.

Récemment, des tâches de mesure d'empan en mémoire de travail qui impliquent des activités beaucoup plus simples ont été créées (Barrouillet, Camos, 2001 ; Lépine, Barrouillet, Camos, 2004). Pour construire ces tâches, les activités complexes impliquées dans les tâches classiques ont été remplacées par des activités élémentaires réputées pour leur facilité, voire leur caractère automatique, telles que parcourir la chaîne numérique ou lire des chiffres. Ainsi, dans la tâche baptisée « continuous operation span » (empan d'opérations continues), les opérations complexes utilisées dans l'empan d'opérations classiques ont été remplacées

---

<sup>55</sup> Dans les tâches d'empan simple, les sujets doivent répéter une série d'items (chiffres, lettres ou mots) qui leur a été présentée visuellement ou auditivement. Contrairement aux tâches d'empan complexes, aucune activité secondaire n'est à mener en parallèle. Cela demande une simple mémorisation à court terme.

par une succession d'opérations élémentaires (ajouter ou soustraire 1). Par exemple, le sujet doit effectuer à haute voix la série d'opérations « 4, +1, -1, +1, +1 », les opérandes étant présentés successivement à l'écran, tout en maintenant en mémoire des lettres. Une autre tâche consiste simplement à faire lire au sujet des listes de chiffres (e.g., 1, 5, 3, 7, 4, 6, 2) présentés un à un pendant qu'il retient des lettres. Ainsi, contrairement aux tâches classiques de mesure d'empan, ces nouvelles tâches impliquent des activités cognitives élémentaires réputées peu coûteuses. Il a cependant été montré que ces tâches sont en fait extrêmement difficiles, et donnent lieu à des performances de rappel plus faibles que les tâches d'empan classiques (Lépine, Barrouillet, Camos, 2004 ; Lépine, Bernardin, Barrouillet, 2005).

Cette difficulté est probablement due au fait que ces nouvelles tâches entravent la mise en œuvre de stratégies sophistiquées de maintien en mémoire des listes à rappeler alors même que la tâche secondaire est effectuée. En effet, dans les tâches classiques, le sujet dispose du temps qu'il désire pour réaliser la tâche secondaire. Par exemple, dans l'empan d'opérations, le sujet peut interrompre momentanément la vérification des équations complexes qui lui sont présentées pour se remémorer mentalement les mots déjà lus, puis reprendre sans dommage le traitement et le conduire à terme. Il est probable que les individus diffèrent dans leurs habiletés à mettre en œuvre de telles stratégies d'alternance entre les diverses activités à mener de front, la mémorisation et le traitement. À l'inverse, les tâches que nous avons créées entravent la mise en œuvre de ces stratégies dans la mesure où le déroulement temporel des traitements est déterminé par le rythme de présentation des stimuli à l'écran (2 secondes par opération dans l'empan d'opérations continues et 1 seconde par chiffre dans la tâche de lecture de chiffres). Ainsi, les tâches conçues permettent des mesures plus pures des capacités cognitives des sujets que les tâches classiques de mémoire de travail pour lesquelles la performance dépend à la fois des ressources cognitives du sujet et de facteurs stratégiques. Ces nouvelles tâches permettent donc de prolonger l'analyse des liens entre capacités en mémoire de travail et performances scolaires, et de répondre à notre questionnement.

### ***III.2 La description des données***

Les données mobilisées ici proviennent de la même base que celle qui vient d'être analysée dans les deux points précédents (cohorte CE2 / 6<sup>ème</sup>). Un sous échantillon représentatif de 87 élèves<sup>56</sup> a été extrait de la population de départ. La sélection de ce sous échantillon a été réalisée en prenant comme critère principal de représentativité les scores des élèves aux évaluations de CE2, la distribution des scores de CE2 étant totalement comparable entre

---

<sup>56</sup> La taille de l'échantillon peut apparaître faible mais elle résulte d'une contrainte temporelle liée à aux passations des tests. En effet, les tests ont nécessité une administration individuelle pouvant varier de 45 à 90 minutes par élève.

l'échantillon initial et ce sous-échantillon. Nous disposons donc pour ces élèves des résultats aux évaluations nationales de CE2 et de 6<sup>ème</sup> ainsi que de leurs caractéristiques socio-démographiques. Les tests cognitifs ont été administrés individuellement et rendent compte de deux dimensions principales des capacités cognitives. La première est la mémoire de travail mesurée par deux types d'épreuves : « reading letter span » (nommée « MT langage ») et « continuous operation span » (nommée « MT nombre »). La seconde est la vitesse des processus cognitifs appréhendée également avec deux types d'épreuves (vitesse de traitement et vitesse d'articulation). La vitesse de traitement a été elle-même mesurée à l'aide de deux tâches : une tâche de barrage où les enfants devaient le plus rapidement possible retrouver et barrer des cibles parmi un ensemble de petits dessins et une tâche informatisée de comparaison de nombres où ils devaient décider lequel de deux nombres présentés était le plus grand. La vitesse d'articulation est évaluée par le temps mis pour réciter le plus rapidement possible l'alphabet ; 4 mesures sont moyennées.

Le tableau 40 présente les corrélations entre les scores cognitifs et scolaires des différentes épreuves. Sans surprise, la corrélation la plus forte est relevée entre les deux scores mesurant les acquis scolaires des élèves. Parmi les différentes mesures de capacités cognitives, seuls les scores en mémoire de travail sont corrélés aux acquisitions scolaires. On ne relève en effet aucun lien entre les résultats en CE2 et en 6ème et les scores de vitesse que ce soit pour la vitesse de traitement ou celle d'articulation<sup>57</sup>. Cette observation s'explique par le fait que les épreuves des évaluations nationales ne cherchent pas à discriminer les élèves sur des critères de vitesse de réalisation des tâches associées aux différents items. Quand on examine plus en détail les corrélations entre mémoire de travail et niveau scolaire, on observe que la mémoire de travail portant sur des épreuves de langage est davantage corrélée au score de 6ème (+0,45) qu'au score de CE2 (+0,37). Le constat inverse est fait pour les épreuves de mémoire de travail portant sur les nombres où cette fois c'est le score de 6ème qui est le moins corrélé à cette mesure (+0,35 et +0,50 pour le CE2).

Tableau 40 : Corrélations entre scores cognitifs et scores aux évaluations nationales (N = 87)

	MT Langage	MT Nombre	Vitesse Traitement	Vitesse Articulation	Score 6 <sup>ème</sup>
MT Nombre	+0,44 ***				
Vitesse Traitement	+0,22 **	+0,37 ***			
Vitesse Articulation	-0,32 ***	-0,26 **	-0,26 **		
Score 6 <sup>ème</sup>	+0,45 ***	+0,35 ***	+0,05 n.s.	-0,12 n.s.	
Score CE2	+0,37 ***	+0,50 ***	+0,13 n.s.	-0,05 n.s.	+0,71 ***

n.s. : non significatif, \*\* : significatif au seuil de 5%, \*\*\* : significatif au seuil de 1%.

<sup>57</sup> Les deux scores rendant compte de la vitesse de traitement sont corrélés négativement, compte tenu de leur échelle de mesure.

En ce qui concerne les variables socio-démographiques, nous nous centrerons uniquement sur l'origine sociale qui pourra être croisée aux niveaux scolaires et cognitifs des élèves. Les données de base font apparaître un classement des catégories sociales en 9 postes mais, compte tenu de nos faibles effectifs, un regroupement en deux catégories a été effectué. La première catégorie regroupe les 29 enfants dont le père est ouvrier (qualifié ou non) ou agriculteur. Les autres élèves, au nombre de 58, relèvent des catégories moyennes et favorisées (professions intermédiaires et supérieures). Au niveau des écarts de performances scolaires entre ces deux catégories sociales, à l'entrée au CE2 les élèves de milieu social défavorisé affichent un score moyen de 69% de réussite tandis que les autres élèves obtiennent un score de 77%, soit une différence de 8 points. A l'entrée en 6<sup>ème</sup> cet écart est proche de 9 points : 68% pour les élèves de milieu défavorisé et 77,5% pour les autres. Des écarts existent également pour les scores mesurant les capacités cognitives (tableau 41) mais ceux-ci sont plus faibles. Ainsi, pour les performances en mémoire de travail dans la dimension du langage, l'écart est de 0,52 (+0,18 - (-0,34)), soit un demi écart-type de la distribution des score puisque l'échelle de mesure des scores cognitifs utilise la variable centrée réduite (moyenne de 0 et écart-type de 1).

*Tableau 41 : Scores moyens aux tests cognitifs et origine sociale*  
(N = 87)

	MT langage	MT nombre	Vitesse traitement	Vitesse articulation
Milieu social moyen et favorisé	+0,18	+0,06	+0,11	+0,09
Milieu social défavorisé	-0,34	-0,13	-0,22	-0,18

Cette description des données suggère fortement que les acquisitions scolaires sont dépendantes à la fois de l'origine sociale et des capacités cognitives des élèves, ces deux dimensions étant par ailleurs corrélées. La section suivante examine plus en détail cette relation de dépendance en rendant compte plus particulièrement du poids des capacités cognitives sur les progressions des élèves, puis en isolant le poids de l'origine sociale dans cette relation.

### ***III.3 Le poids des capacités cognitives et de l'origine sociale sur les progressions scolaires***

Le tableau 42 présente deux modèles avec comme variable dépendante les scores globaux à l'entrée en 6<sup>ème</sup><sup>58</sup>. Le premier modèle (A) traduit à nouveau le lien entre les deux scores (CE2 et 6<sup>ème</sup>) et nous informe que la variance du score de 6<sup>ème</sup> est expliquée pour moitié ( $R^2 = .50$ ) par les différences de scores à l'entrée en CE2. L'ajout des scores cognitifs à ce modèle

<sup>58</sup> Les scores de vitesse (traitement et articulation) n'ont pas été pris en compte dans les analyses qui vont suivre du fait de leur non corrélation avec les scores d'acquisitions scolaires

apporte un gain de variance, le  $R^2$  du modèle B étant de 0,54<sup>59</sup>. Le coefficient du modèle B indique qu'un point de plus au score de mémoire de travail langage (ce qui équivaut, compte tenu de l'échelle de mesure, à un écart-type) procure un accroissement du score de 6<sup>ème</sup> de 3,10 points, à niveau scolaire de CE2 donné.

*Tableau 42 : Modèles de régression expliquant les progressions des élèves au cycle III, en fonction des acquisitions initiales et des capacités en mémoire de travail.*

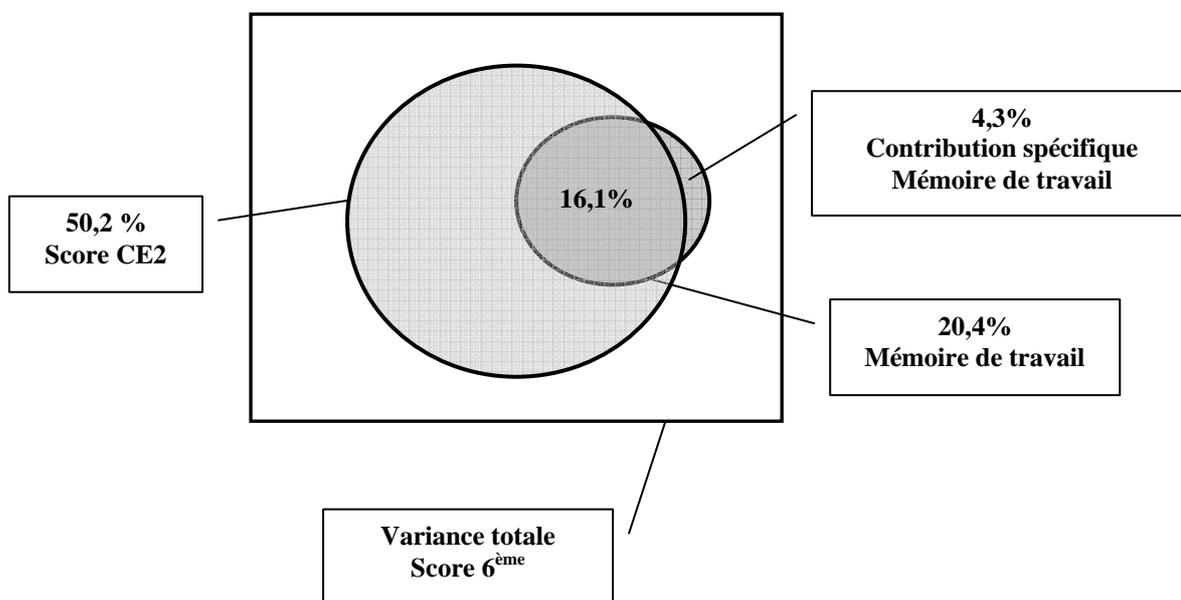
(N=87)

Variables explicatives	A	B
Score CE2	+0,73 ***	+0,64 ***
Score Mémoire de travail		+3,10 ***
Constante	19,97 ***	26,19 ***
$R^2$	0,502	0,544

n.s. : non significatif, \*\* : significatif au seuil de 5%,

\*\*\* : significatif au seuil de 1%.

Le graphique suivant permet, sur la base des modèles A et B, de compléter les premiers résultats par une image de la décomposition de la variance du score de 6<sup>ème</sup> en fonction du niveau d'acquisitions à l'entrée au CE2 et des performances en mémoire de travail.



*Graphique 27: Décomposition de la variance du score de 6ème en fonction du score de CE2 et du score en mémoire de travail*

<sup>59</sup> Seule la dimension « langage » de la mémoire de travail joue sur les progressions des élèves, le coefficient associé à la seconde dimension (nombre) n'étant pas significatif quand les deux variables sont introduites simultanément dans le même

Le score en mémoire de travail explique à lui seul 20,4% de la variance du score de 6<sup>ème</sup> (modèle non présenté dans le tableau précédent). Comme nous venons de le noter, le modèle B, qui intègre à la fois le score de CE2 et le score en mémoire de travail, explique 54,4% de la variance du score de 6<sup>ème</sup>. Par conséquent, la contribution spécifique de la mémoire de travail à l'explication des différences de réussite des élèves à l'entrée au collège est donc de l'ordre de 4,2 points (54,4 – 50,2) et les 16,1 % restants sont donc le fait d'une part de variance conjointe entre les deux mesures de performances des élèves.

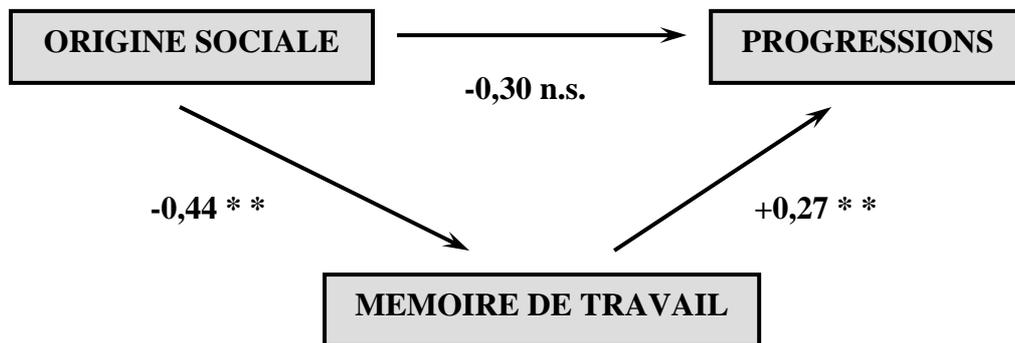
Ces analyses ont été complétées en introduisant une information supplémentaire : l'origine sociale de l'élève (tableau 43).

*Tableau 43: Modèles de régression expliquant les progressions des élèves au cycle III, en fonction des acquisitions initiales, des capacités en mémoire de travail, et de l'origine sociale*

Variables explicatives		C	D
Score CE2		+0,69 ***	+0,62 ***
Score Mémoire de travail			+2,83 ***
Autres professions	Milieu social défavorisé	-3,99 **	-3,21 n.s.
Constante		24,32 ***	29,18 ***
R <sup>2</sup>		0,52	0,56

Le modèle C met en évidence l'influence de l'origine sociale sur l'évolution des acquis au cours du cycle III, puisqu'à score de début CE2 comparable, les élèves de milieu social défavorisé ont des progressions inférieures de quasiment 4 points aux autres élèves. Le modèle D intègre, à la fois, l'origine sociale et les capacités cognitives dans l'équation de régression et fournit des résultats particulièrement intéressants. En effet, on observe que l'effet de l'origine sociale est totalement absorbé par l'influence des capacités cognitives. Ainsi, la variable « milieu social » n'affecte plus significativement les scores alors que le coefficient associé à la variable « mémoire de travail » reste très proche de celui relevé dans le modèle B précédent. Ceci nous amène à nous interroger plus finement sur l'articulation entre l'origine sociale de l'élève, ses performances en mémoire de travail et ses progressions entre le CE2 et l'entrée en 6<sup>ème</sup>.

Dans cette perspective, une analyse en pistes causales (« path analysis ») a été conduite intégrant ces trois variables. Les progrès des élèves ont été appréhendés ici par les résidus de la régression entre score de 6<sup>ème</sup> et de CE2 ; une valeur positive et élevée de cet indicateur impliquant une forte progression et inversement. Le graphique 28 permet de visualiser les résultats de l'analyse en pistes causales.



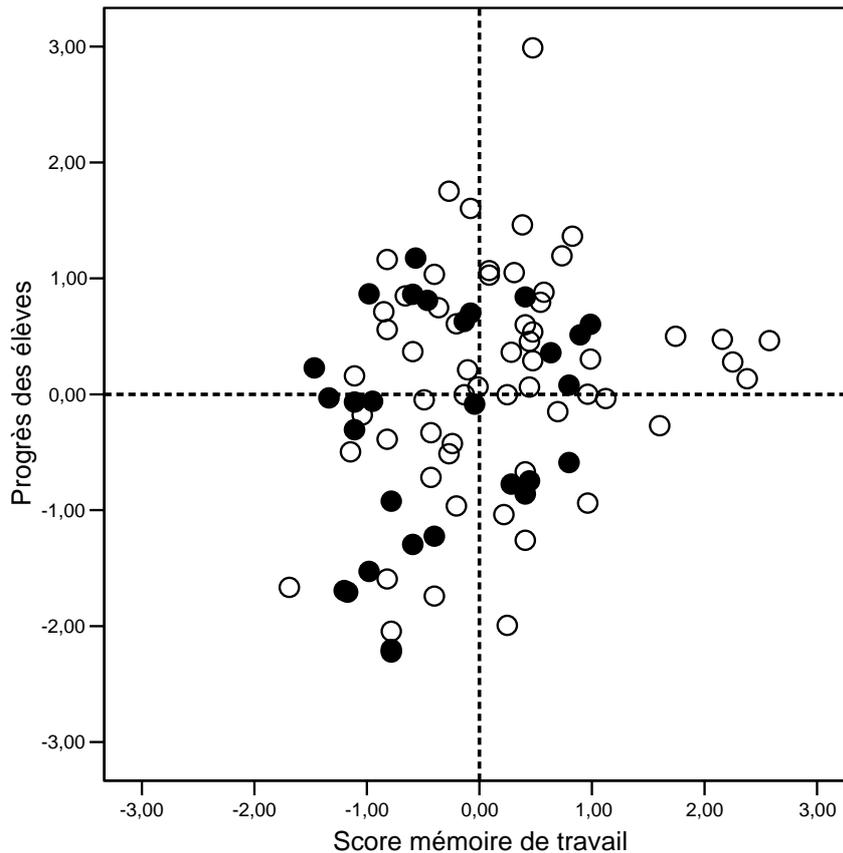
*Graphique 28 : Résultats de l'analyse en pistes causales entre origine sociale, mémoire de travail et progrès scolaires au cycle III*

Le principe de cette analyse en pistes causales est de décomposer l'effet de l'origine sociale sur les progressions des élèves en un effet direct et un effet indirect. L'effet direct de l'origine est nul (coefficient de régression de -0,30, non significatif) et confirme les résultats précédents du modèle D. L'effet indirect de l'origine sociale sur les progrès transite par les performances en mémoire de travail. Cet effet indirect se décompose donc en deux parties. La première est symbolisée sur le graphique par la flèche qui relie l'origine sociale au score en mémoire de travail (coefficient de -0,44, significatif au seuil de 5%). Ceci traduit l'effet négatif d'une origine sociale défavorisée sur les performances en mémoire de travail. La seconde partie est représentée par la flèche reliant ce même score cognitif aux progrès des élèves. Dans ce cas, le coefficient positif relevé (+0,27, significatif au seuil de 5%) traduit un effet positif de la mémoire de travail sur les progrès réalisés au cycle III. L'effet indirect correspond donc au produit des deux relations précédentes, soit  $-0,44 \times 0,27 = -0,12$ .

Ces résultats fournissent des éléments sur le rôle central de la mémoire de travail dans l'explication des différences de progressions entre élèves. C'est principalement par cette dimension des capacités cognitives que l'origine sociale exerce son influence sur les progressions scolaires, sachant que par ailleurs, le rôle de la vitesse de traitement n'a pas été mis en évidence sur les données mobilisées dans cette recherche.

Tout cela peut donner une image assez déterministe de la réussite scolaire et il serait tentant d'avancer que les possibilités de progrès des élèves issus des milieux défavorisés seraient limitées du fait de leurs plus faibles performances en mémoire de travail. Il convient de relativiser les résultats précédents en les situant dans le contexte statistique d'ensemble des analyses. On notera en premier lieu que l'origine sociale n'explique qu'une faible part des

différences de capacités cognitives, la part de variance du score de mémoire de travail expliquée par le milieu social de l'élève est proche de 7%, ce qui est un chiffre modeste. On rappellera en second lieu que si les progrès des élèves au cycle III dépendent en partie de leurs capacités cognitives, cette liaison est elle-même limitée, puisque les scores en mémoire de travail expliquent également 7% des différences de progressions scolaires entre le CE2 et la 6<sup>ème</sup>. Le graphique suivant présente la répartition des élèves de l'échantillon en fonction des trois éléments qui viennent d'être analysés et commentés.



*Graphique 29: Relation entre score en mémoire de travail, progrès des élèves au cycle III et origine sociale*

Les élèves sont tout d'abord distingués en fonction de leur origine sociale, les cercles noirs représentant les élèves de milieu social défavorisé. Les deux axes du graphique combinent les performances en mémoire de travail (axe horizontal) et les progrès réalisés entre le CE2 et la 6<sup>ème</sup> sur la base des scores aux évaluations. Les deux échelles étant centrées réduites (moyenne de 0 et écart-type de 1), les traits en pointillés sur le graphique matérialisent la moyenne des deux échelles. Cette représentation graphique a surtout pour objectif de montrer, que même si les élèves les plus défavorisés socialement ont, dans l'ensemble, une position moins avantageuse que les autres, la plupart partagent leur situation avec nombre d'élèves issus de milieux social plus favorisés.

Les principaux résultats de cette dernière section confirment le lien entre les capacités cognitives des élèves et leur aptitude à progresser dans leurs acquisitions scolaires au cours de l'école élémentaire. Les dimensions des acquis scolaires les plus liées aux capacités cognitives ont été mises en évidence à partir de compétences issues des évaluations nationales de CE2 et de 6<sup>ème</sup> identifiées précédemment. L'analyse des relations statistiques entre les scores en mémoire de travail et les compétences scolaires donne lieu à des corrélations d'intensité variée (de 0 à +0,45). Les 10 compétences qui affichent les coefficients les plus élevés ont trait à l'évaluation de 6<sup>ème</sup> et 8 d'entre elles évaluent des habiletés en calcul mental. Deux autres compétences ont trait à la production d'écrits. Cette proximité entre calcul mental et mémoire de travail n'est pas surprenante car les activités scolaires de cette nature font « naturellement » appel à cet aspect des capacités cognitives (McLean, Hitch, 1999). A l'opposé, certaines compétences scolaires entretiennent des liens très faibles avec la mémoire de travail. Parmi les 10 compétences les moins corrélées aux capacités cognitives la répartition entre les deux niveaux scolaires (CE2 et 6<sup>ème</sup>) est équilibrée. Les compétences évaluées concernent la production de texte, la compréhension simple, le repérage spatial (mise en page d'un texte, reproduction d'une figure).

Les résultats permettent aussi de s'interroger sur les mécanismes d'apprentissages des élèves à l'école élémentaire et peuvent déboucher sur des préconisations pertinentes en terme de politique éducative. Si le niveau d'acquisition des élèves est déterminé à hauteur de 50% par le niveau obtenu à l'entrée au CE2, les variables sociales et les capacités cognitives impriment elles aussi leur marque sur les progressions au cours du cycle III. L'originalité de notre démarche est d'avoir mobilisé des données longitudinales pour mettre en évidence le rôle de la mémoire de travail sur les résultats scolaires, tout en contrôlant l'effet de l'origine sociale. Le résultat le plus important est sans aucun doute la disparition de l'effet du milieu social de l'élève sur les progressions quand on intègre dans l'analyse les capacités en mémoire de travail. En outre, certaines compétences scolaires, comme les habiletés en calcul mental, sont très liées à ces capacités cognitives. On peut donc penser qu'il peut être pertinent de développer certaines activités d'enseignement permettant de limiter l'influence des deux variables principales (origine sociale et capacités cognitives) sur les écarts de progressions entre élèves. Les activités mobilisant les habiletés en calcul mental sont donc les premières concernées et leur pratique systématique pourrait permettre de réduire le coût cognitif des activités d'apprentissage en automatisant certains processus. Dans la mesure où le niveau de compétences des élèves à l'entrée en CE2 est fortement prédictif des progrès, c'est donc principalement avant le cycle III que doivent être mises en place ces activités systématiques afin de mieux armer les élèves dans les dimensions des acquisitions les plus prédictives de la réussite ultérieure.

## **Conclusion générale**

Cette conclusion s'attachera modestement à situer les résultats exposés dans notre travail dans une perspective de politique éducative en évoquant quelques points qui nous paraissent essentiels pour l'amélioration de l'école primaire française.

Notre travail avait pour objectif de porter un regard pluridisciplinaire sur l'évaluation de l'école primaire française ; il s'agissait de mettre en relation une partie de nos travaux les plus récents et en cours, tout en soulignant l'intérêt de l'aspect méthodologique de l'évaluation. D'un point de vue plus général, il était aussi question de montrer que la recherche empirique en éducation pouvait servir à éclairer certains aspects du fonctionnement de notre système éducatif. Il ne s'agissait en aucun cas de traiter de manière exhaustive l'ensemble de la problématique liée à l'école primaire française et à sa qualité. La variété des travaux exposés a été volontairement ciblée sur la dimension méthodologique qui mobilise différents niveaux d'analyse statistique, de l'approche macro avec les comparaisons internationales, à l'approche micro qui se focalise sur les apprentissages des élèves. La complexité de l'organisation des systèmes éducatifs, la multiplicité des questions de politique éducative, les contraintes statistiques liées à l'utilisation de la modélisation, sont autant d'éléments qui justifient notre perspective méthodologique.

Les facteurs qui expliquent le niveau d'acquisition des élèves interviennent, de fait, à des degrés différents et emboîtés, intégrant l'ensemble des éléments du contexte scolaire. Nous avons plus particulièrement exploré dans ce texte des variables associées au niveau du pays, de la classe et de l'élève lui-même, alors que l'établissement scolaire et l'enseignant n'ont pas été ciblés. Ceci correspond à un choix découlant directement de la thématique des recherches mobilisées qui ne porte pas en priorité sur ces deux dimensions du contexte, même si celles-ci contribuent bien évidemment à expliquer la qualité de l'école.

Les analyses présentées dans ce travail permettent tout d'abord de s'interroger globalement sur l'efficience et l'efficacité de notre école primaire. Sur le plan des acquisitions des élèves, la position de la France dans le contexte international est beaucoup moins favorable qu'il y a une dizaine d'années. On constate, en outre, une diminution sensible de l'efficience qualitative et quantitative de notre école au cours de cette dernière décennie. Pourtant, l'augmentation des coûts unitaires a plutôt avantagé l'école primaire eu égard aux accroissements des effectifs d'élèves aux différents niveaux d'enseignement. Les conditions d'enseignement se sont donc progressivement améliorées au cours de ces trente dernières années (meilleurs taux d'encadrement, augmentation des moyens, revalorisation des salaires des enseignants du primaire...) sans pour autant que le niveau d'acquisition des élèves

connaisse une progression notable. Il y aurait donc une incapacité de notre système éducatif à transformer efficacement les ressources en résultats.

Bien que la recherche des causes de ce dysfonctionnement ne soit pas simple, chacun s'accorde aujourd'hui pour mettre en avant la faiblesse des mécanismes de pilotage, de gestion pédagogique et d'évaluation (H.C.E., 2007). La politique des cycles constitue un exemple illustratif de la difficulté que peut avoir l'institution à agir pour la mise en place d'une réforme qui portait, non pas sur les structures, mais sur les pratiques des équipes enseignantes au sein des écoles et des classes. Cette réforme, associée à la loi d'orientation de 1989 constituait pourtant un cadre législatif pertinent et adapté à la prise en compte de la diversité des élèves et une solution à la gestion de l'hétérogénéité des parcours scolaires. Les causes de l'échec de la mise en place effective des cycles sont vraisemblablement le fait de l'insuffisance de mesures d'accompagnement et d'un manque de continuité dans la réalisation de cette réforme au fil des années. Dans un même temps, certaines pratiques sont restées persistantes alors que l'on connaît leur inefficacité : le recours au redoublement en est une illustration caractéristique. Cette insuffisance du pilotage de proximité remet en cause les missions attribuées aux Inspecteurs de l'Education nationale et l'échelon de pilotage pédagogique que constitue la circonscription primaire. La question d'une gestion pédagogique plus proche du terrain ne peut alors faire l'économie d'une réflexion sur le statut des établissements scolaires du premier degré et sur celui des directeurs d'école (H.C.E., 2007).

Au-delà des résultats d'ensemble de notre école primaire, notre démarche visait à mobiliser des recherches capables d'apporter des éléments de réflexion plus précis sur le fonctionnement de notre école. La centration sur le contexte de la classe a permis de s'interroger sur les procédures de constitution des groupes d'élèves et leurs incidences sur les acquisitions des élèves. De façon complémentaire, l'influence de la composition sociale de la classe a été examinée sous l'angle de l'efficacité pédagogique. Enfin, une centration sur l'élève et ses apprentissages a été réalisée par une analyse fine de la structure des acquisitions des élèves au cours de l'école élémentaire. La question est alors de savoir comment les travaux scientifiques peuvent fournir des pistes concrètes à la politique éducative, sachant que, par ailleurs, l'utilisation de ces résultats par les responsables est loin d'être une démarche habituelle dans le contexte français. Une illustration classique pouvant appuyer cette affirmation concerne la pratique du redoublement qui, malgré l'unanimité des constats scientifiques (Crahay, 1996, Paul, 1996), a donné lieu à des annonces parfois singulières de la part des responsables au plus haut niveau du système éducatif ces dernières années. L'exemple récurrent de l'apprentissage de la lecture où se mêlent au débat, à la fois des arguments scientifiques et des convictions personnelles, est lui aussi caractéristique de la faible portée politique des travaux de recherche français sur l'éducation. L'analyse de cette

question nécessite d'articuler enjeux politiques, processus de décision et rapports de force entre les acteurs.

Le paradoxe est réel dans la mesure où l'ambition des décideurs est bien de mettre le système éducatif, au moins partiellement, au service de l'économie et du social en répondant à des besoins externes clairement identifiés. Force est pourtant de constater que les décisions sont davantage fondées sur l'arbitrage entre les intérêts contradictoires véhiculés par les acteurs internes du système. Tout semble se passer comme si le système éducatif fonctionnait en autonomie relative par rapport aux contraintes économiques et sociales (Solaux, 2002). C'est donc dans ce cadre national, peu favorable à la prise en compte des conclusions des recherches, que nos résultats prennent place. On se limitera ici à extraire des travaux présentés quelques éléments qui nous semblent avoir une implication importante en termes de politique éducative.

D'un point de vue global tout d'abord, les analyses conduites à partir des études internationales comme Pisa, montrent bien que la structure des systèmes d'enseignement n'est pas indépendante du niveau moyen d'acquisitions des élèves et de leurs disparités. Les systèmes qui ont opté pour des structures scolaires indifférenciées jusqu'à la fin du premier cycle secondaire apparaissent plus efficaces et équitables que les autres. Mais c'est principalement dans la dimension relative aux inégalités sociales de réussite que l'avantage d'un cursus indifférencié est le plus marqué. Ce constat a des répercussions évidentes sur des choix de politique éducative qui sont débattus actuellement dans notre pays avec la question du collège unique et de la carte scolaire. C'est donc non seulement la structure même du système éducatif qui est en jeu mais également la ségrégation entre établissements, celle-ci pouvant faire l'objet de politiques plus ou moins volontaires. Les travaux scientifiques plaident assez largement pour la mixité scolaire et sociale au sein des établissements et des classes, non seulement au niveau des analyses établies au niveau macro, mais aussi à travers les travaux conduits au niveau micro.

Les analyses présentées précédemment confirment l'influence de la composition des classes sur le niveau d'acquisition des élèves. Si, à l'école primaire, les enseignants ont, lors des procédures de constitution des classes, la possibilité d'agir sur la mixité scolaire, cette action demeure limitée. Les résultats tendent toutefois à montrer que les choix réalisés ont bien une incidence sur les acquisitions des élèves et cela se vérifie particulièrement quand ces choix portent sur l'affectation des élèves dans des contextes pédagogiques particuliers (classes à cours multiple notamment). Plus largement, il a été également souligné que la composition sociale des classes au sein des écoles n'était pas sans incidence sur les progressions des élèves.

De façon beaucoup plus précise, l'analyse des apprentissages des élèves dans une perspective dynamique montre que les clés de la réussite transitent réellement par les contenus d'activités et les compétences qui y sont associées. Les résultats actuels de l'école primaire laissent penser que l'école maternelle joue un rôle essentiel dans la construction de ces compétences. Certaines compétences évaluées à l'entrée à l'école élémentaire sont en effet très prédictives de la maîtrise ultérieure d'autres compétences. Il s'agit alors d'identifier comment ces compétences se construisent chez l'enfant, comment les contenus de programmes peuvent les prendre en compte à travers les activités pratiquées dans les classes. Le rôle de la grande section de maternelle est à ce titre fondamental pour assurer les bases nécessaires, d'autant plus que la réussite scolaire, comme nous l'avons souligné, est largement dépendante des variables sociales et des facteurs cognitifs.

Les recherches exposées dans ce texte laissent entrevoir que les causes des problèmes réels de l'école primaire française sont à rechercher au niveau le plus fin de la politique éducative : le fonctionnement pédagogique des écoles et des classes. Les inégalités de réussite entre élèves ne peuvent se réduire que par une action volontariste sur des questions comme celles de la différenciation pédagogique, du respect des rythmes d'apprentissage, du développement précoce des compétences essentielles. C'est donc sur l'organisation même des contenus d'enseignement et de leurs modalités de transmission qu'il serait urgent d'agir efficacement. Notre système éducatif s'est pourtant doté d'outils législatifs (loi d'orientation) et pédagogiques (évaluations diagnostiques et leurs documents d'accompagnement) pertinents, sans que les pratiques d'enseignement soient suffisamment modifiées pour laisser des traces visibles au niveau de la réussite de tous les élèves.

Dans un contexte social et éducatif national en évolution, à l'heure où une consultation s'engage sur le métier d'enseignant, la politique éducative à tout à gagner pour qu'une place plus large soit accordée aux travaux empiriques de recherche en éducation dans le but d'alimenter la réflexion, d'appuyer la prise de décision, et, in fine, de participer à l'amélioration de notre école primaire.

## **Bibliographie**

Aish-Van Vaerenbergh, AM.(1997), « Modèles statistiques et inférences causales : analyse de structures de covariances avec LISREL ». In Faut-il chercher aux causes une raison ? L'explication causale en sciences humaines. Aish-Van Vaerenbergh, AM. et al. Librairie philosophique Vrin, pp106-130.

Altet M. (1994). Comment interagissent enseignant et élèves en classe? *Revue Française de Pédagogie*, N°107, pp. 123-139.

Altet M., Bressoux P., Bru M., Lambert C. (1994). Étude exploratoire des pratiques d'enseignement en classe de CE2. Les dossiers d'éducation et formations, 44.

Altet M., Bressoux P., Bru M., Lambert C. (1996). Étude exploratoire des pratiques d'enseignement en classe de CE2 : Deuxième phase. Les dossiers d'éducation et formations, 70.

Ardoino J., Berger G. (1986). L'évaluation, comme interprétation, in Pour, n°107, éd. Privat.

Arnoux M. (2004). Une étude sur l'activité de l'enseignant en situation interactive : observation de la gestion et de l'organisation du temps dans des classes de CM1. Thèse de Doctorat, Université Pierre Mendès France, Grenoble.

Attali A., Bressoux P. (2002). L'évaluation des pratiques éducatives dans les premiers et second degrés. Rapport pour le Haut Conseil de l'Évaluation de l'École. Paris, HCÉÉ.

Atkinson, J.R., Shiffrin, R.M. (1968), "Human memory : A proposed system and its control processes". In K.W. Spence (Ed.), The psychology of learning and motivation : Advances in research and theory, Vol. 2 (pp. 89-195). New York : Academic Press.

Aubriet-Morlaix S. (1999), *Essai sur l'allocation et l'optimisation du temps scolaire*, Thèse de Doctorat en Sciences économiques, Université de Bourgogne, Dijon.

Baddeley, A.D., Hitch, G.J. (1974), "Working memory". In G.H. Bower (Ed.), Recent advances in learning and motivation, Vol. 8 (pp. 647-667). New York : Academic Press.

Barrouillet, P., Camos, V. (2001), "Developmental increase in Working Memory Span : Resource sharing or temporal decay ?", Journal of Memory and Language, 45 (1), 1-20.

Barrouillet P., Lecas. JF, (1999), "Mental Models in conditionnal reasoning and working memory". Thinking and Reasoning, 5 (4), 289-302.

Barrouillet, P., Bernardin, S., Camos, V. (2004), "Time constraints and resource sharing in adults' working memory span". Journal of Experimental Psychology : General, 133, 83-100.

Barrouillet P., Camos V., Morlaix S., Suchaut B. (2007). Compétences scolaires, capacités cognitives et origine sociale : quels liens à l'école élémentaire ? Soumis à la Revue Française de Pédagogie.

Becker, G. S. (1964). Human Capital : A Theoretical and Empirical Analysis. New-York: Columbia University Press.187 p.

Bloom, B.S. (1979) Caractéristiques individuelles et apprentissages scolaires, Bruxelles, Labor, Paris, Nathan.

Bossert S.T., Dwyer D.C., Rowan B., Lee G.V. (1982). The Instructional Management Role of the Principal, Educational Administration Quaterly. 18 (3), pp. 34-64.

Bottani N., Vrignaud P. (2004). La France et les évaluations internationales des acquis des élèves, Paris, HCEE, OCDE.

Bressoux P. (1994). Les recherches sur les effets-école et les effets-maître, Revue française de pédagogie, 108, pp. 91-137.

Bressoux P. (1995). Les effets du contexte scolaire sur les acquisitions des élèves : effets-écoles et effets-classes en lecture. Revue Française de Sociologie, 36, 273-294.

Bressoux P. (2000). Modélisation et évaluation des environnements et des pratiques d'enseignement. Habilitation à diriger les recherches. Université Pierre Mendès France. Grenoble.

Bressoux P. dir. (2002). Les stratégies de l'enseignant en situation d'interaction. Note de synthèse pour cognitive. Programme Ecole et sciences cognitives. Février 2002. 231 p.

Bressoux, P., Pansu, P., (2003). Quand les enseignants jugent leurs élèves. Paris : PUF.

Bressoux P., Coustère P., Leroy-Audouin C.(1997) Les modèles multiniveau dans l'analyse écologique: le cas de la recherche en éducation, in Revue Française de Sociologie, XXXVIII, pp. 67-96.

Briquet-Duhazé S. (2005). Ecoute furtive en lecture des élèves du préscolaire dans les classes à plusieurs niveaux, in Education et francophonie, vol. XXXIII : 2, 259-275.

Bru M. (1991). Les variations didactiques dans l'organisation des conditions d'apprentissage. Toulouse : Editions universitaires du sud.

Bru, M. (2002). Pratiques enseignantes : des recherches à conforter et à développer. Revue Française de Pédagogie, N°138, pp. 63-74.

Burns R., Mason D. (1998). Class Formation and Composition in Elementary Schools, American Educational Research Journal, Vol. 35, N°4, 739-772.

Burns R., Mason D. (2002). Class Composition and Students Achievement in Elementary Schools, American Educational Research Journal, Vol. 39, N°1, 207-233.

Caille JP., Rosenwald F., (2006), “Les inégalités de réussite à l'école élémentaire : Construction et évolution ”. In INSEE (dir.). France, portrait social. Paris : Institut national de la statistique et des études économiques, p. 115–137.

Carleton Board of Education (1991). Summary of observations of a selected group of exemplary multigrade classes in the CBE Nepean. Ontario, Canada.

Case, R. (1985). Intellectual development: Birth to adulthood. New York: Academic Press.

Coleman, J., McDill E. L. (1965). Family and Peer Influences in College Plans of High School Students. Sociology of Education, 38, 2, pp. 112-126.

Cosnefroy O., Gibert F. (2006), Comment mieux apprécier l'évolution des compétences et connaissances au cours du temps ? Une étude chez des élèves de cours préparatoire. 18ème colloque international de l'ADMEE : Comment évaluer ? Outils, dispositifs et acteurs/ (24, 25 et 26 octobre 2005, Reims.

Cousin O. (1993). L'effet établissement. Construction d'une problématique, Revue française de sociologie, 34, 395-419.

Cousin O. (1998), L'efficacité des collèges. Sociologie de l'effet établissement, PUF, Paris.

Cousin O. (2000), Politiques et effets-établissements dans l'enseignement secondaire, in Van Zanten A. (sous la dir. de), L'école, l'état des savoirs. Ed. La découverte, pp. 139-148.

Cowan, N. (1995). Attention and memory: An integrated framework. New York: Oxford University Press.

Crahay M. (1989). Contraintes de situation et interactions maîtres-élèves : changer sa façon d'enseigner, est-ce possible ? Revue française de pédagogie. N°88. pp.67-94.

Crahay M. (1996). Peut-on lutter contre l'échec scolaire ? Bruxelles, De Boeck.

Crahay M. (2000). L'école peut-elle être juste et efficace ? Bruxelles, De Boeck

Crahay M. (2006), Dangers, incertitudes et incomplétude de la logique de la compétence en éducation. Revue française de pédagogie, N°154, pp. 97-110.

Crahay M., Delhaxe A. (2003). L'enseignement secondaire inférieur : entre culture de l'intégration et culture de la différenciation, Rapport pour le PIREF-Ministère de la recherche.

Cronbach, L. (1976), Research on Classrooms and Schools : Formulation of questions, design and analysis, Stanford, CA : Stanford Evaluation Consortium.

Damoiselet N., Lévy-Garboua L., 2001, Comparaison des systèmes éducatifs : une approche économique, in Boudon R., Bulle N., Cherkaoui M. (eds), Ecole et société, Paris : PUF, pp. 241-268.

Daneman, M., Carpenter, P.A. (1980). "Individual differences in working memory and reading". Journal of Verbal Learning and Verbal Behavior, 19, 450-466.

Daneman M., Carpenter, P.A., (1983), "Individual Differences in integrating information between and within sentences". Journal of Experimental Psychology : learning, Memory, and Cognition, 9, 561-584.

Danner M., Duru-Bellat M., Le Bastard-Landrier S., Suchaut B. (2001). L'aide individualisée en seconde, mise en route et premiers effets d'une innovation pédagogique . Education et Formations, N°60, Juillet-septembre 2001. pp. 55-65.

Davezies L. (2005). Influence des caractéristiques du groupe des pairs sur la scolarité élémentaire. Education et formations, n°72, 171-199.

De Ketele J.M. (1993). L'évaluation conjuguée en paradigme. Revue française de pédagogie. N°103 pp.59-80.

De Ketele J.M., Gerard F.M. (2005), La validation des épreuves d'évaluation selon l'approche par les compétences, Mesure et Évaluation en Éducation, À paraître.

Demailly L. Dir. (2000a). Évaluer les politiques éducatives : sens, enjeux, pratiques. De Boeck, 2000.

Demailly L. (2000b), Management et évaluation des établissements, in VAN ZANTEN A. (sous la dir. de), L'école, l'état des savoirs. Ed. La découverte, pp. 130-138.

Demeuse M., Crahay M., Monseur C. (2001). Efficiency and Equity, in Hutmacher W. et al.(eds), In Pursuit of Equity in Education, Dordrecht : Kluwer Academic Publishers.

Demeuse M., Henry G (2004), « La théorie classique des tests », in Demeuse M. (2004), Introduction aux théories et aux méthodes de la mesure en sciences psychologiques et en sciences de l'éducation. Les Editions de l'université de Liège.

Derouet J.L. (1987), Contrat et établissement scolaire, in Burguière et al., Contrat et éducation. La pédagogie du contrat, le contrat en éducation, Paris : L'Harmattan/INRP, coll. CRESAS n°6.

Dreeben R., Barr R., (1988). Classroom composition and the design of instruction, Sociology of Education, vol.61, 129-142.

Dupriez V. (2003). De l'isolement des enseignants au travail en équipe : les différentes voies de construction de l'accord dans les établissements, Cahier de Recherche du GIRSEF, n° 23, 20 p.

Durkheim E. (1897) Le Suicide, Paris : PUF.

Duru-Bellat M. (1990). L'école des filles. Quelle formation pour quels rôles sociaux ? L'Harmattan.

Duru-Bellat M., 1998, La mixité et le "curriculum caché" des élèves, Enfance et Psy, n°3. pp. 73-78.

Duru-Bellat M. (2002). Les inégalités sociales à l'école : genèse et mythes, Paris, PUF.

Duru-Bellat M. (2003a). Inégalités sociales à l'école et politiques éducatives, Paris, UNESCO.

Duru-Bellat M. (2003b), Les apprentissages dans leur contexte : les effets de la composition de l'environnement scolaire. Carrefours de l'éducation, 16, 182-206.

Duru-Bellat M., Van Zanten A. (1999), Sociologie de l'école. A. Colin, Paris.

Duru-Bellat M., Landrier-Le Bastard S., Piquée C., Suchaut B. (2004). Tonalité sociale du contexte et expérience scolaire des élèves au lycée et à l'école primaire, Revue française de sociologie, 45, n°3, 441-468.

Duru-Bellat M., Le Bastard-Landrier S., Piquée C., Suchaut B. (2004). Tonalité sociale du contexte et expérience scolaire des élèves au lycée et à l'école primaire. Revue Française de Sociologie, 45, n°3, pp. 441-468.

Duru-Bellat M., Mingat, A. (1988). Le déroulement de la scolarité au collège : Le contexte « fait des différences ». Revue Française de Sociologie, XIX, pp. 649-666.

Duru-Bellat M., Mingat, A. (1997). La constitution de classes de niveau par les collèges : Les effets pervers d'une pratique à visée égalisatrice. Revue Française de Sociologie, N° 38, pp. 759-790.

Duru-Bellat M., Mingat A. (1993), Pour une approche analytique du fonctionnement du système éducatif. PUF.

Duru-Bellat M., Mingat A. (1998a). Vérification et falsification dans la recherche en éducation, in Hadji C, Baillé J. (éds), Recherche et éducation. Vers une « nouvelle alliance ». La démarche de la preuve en 10 questions. Ed. De Boeck Université, coll. Pédagogies en développement.

Duru-Bellat M., Mingat A. (1998b). L'opacité du fonctionnement du système éducatif français : quelles pistes pour une meilleure régulation ? Policy Options/Options politiques, Vol. 19, N°6, pp. 46-49.

Duru-Bellat M., Mons N., Suchaut B. (2003). Caractéristiques des systèmes éducatifs et compétences des jeunes de 15 ans : l'éclairage des comparaisons entre pays. Cahiers de l'Irédu, N°66, Dijon, IREDU, février 2004, 158 p.

Duru-Bellat M., Suchaut B. (2005a). Organisation and context, efficiency and equity of educational systems. What PISA tells us. European Educational Research Journal, 4 (3), 2005. pp. 181-194

Duru-Bellat M., Suchaut B. (2005b). L'approche sociologique des effets du contexte scolaire. Méthodes et difficultés. Revue Internationale de Psychologie Sociale, 2005, n°3, pp 5-42

Doyle W. (1986). Paradigmes de recherche sur l'efficacité des enseignants. In Crahay M., Lafontaine D. (Eds.), L'art et la science de l'enseignement. (pp. 435-483). Bruxelles, Ed. Labor.

Engle, R.W., Cantor, J., Carullo, J.J. (1992), "Individual differences in working memory and comprehension: A test of four hypotheses". Journal of Experimental Psychology: Learning, Memory, and Cognition, 5, 972-992.

Erikson, R., Jonsson, J. eds. (1996). Can Education be Equalized ? Oxford : Westview Press.

Eurydice (2005). Chiffres clés de l'éducation en Europe 2005. In Eurydice [En ligne]. [http://www.eurydice.org/ressources/eurydice/pdf/0\\_integral/052FR.pdf](http://www.eurydice.org/ressources/eurydice/pdf/0_integral/052FR.pdf).

Fayol M. (1989). Psychologie cognitive et instruction. In Monteil J.M., Fayol M. (Eds.). La psychologie scientifique et ses applications. Grenoble, PUG.

Felouzis G. (2003). La ségrégation ethnique au collège et ses conséquences. Revue Française de Sociologie. 44, 3, 413-448.

Ferrier J. (1998). Améliorer l'efficacité de l'école primaire, Rapport remis à Ségolène Royal, Ministre déléguée chargée de l'enseignement scolaire, MENRT.

Florin A., Cosnefroy O., Guimard P. (2004). Trimestre de naissance et parcours scolaire. Revue européenne de psychologie appliquée, Vol. 54, n°4, pp. 237-246.

Forestier C., Emin J.C., Thélot C. (2007). Que vaut l'enseignement en France ? Les conclusions du Haut Conseil de l'évaluation de l'école. Stock, 2007.

Fraine (de) B., van Damme J., van Landeghem G., Opdenakker M-C., Onghena P., (2003). The effects of schools and classes on language achievement, British Educational Research Journal, 29, n°6, 841-859.

Friedberg E. (1993). Le pouvoir et la règle, Paris : Le Seuil.

Fuller B., Wood K., Rapoport T., Dornbush S.M. (1982). The organizational context of individual efficacy. Review of Educational Research, 52 (1), pp. 7-30.

Gamoran A., Mare R.D., (1989). Secondary school tracking and educational inequality: compensation, reinforcement, or neutrality ? American Journal of Sociology, n°94, 1146-1183.

Gather Thurler M. (1994), Relations professionnelles et culture des établissements scolaires : au-delà du culte de l'individualisme ? Revue Française de Pédagogie, n° 109, pp. 19-39.

Gathercole, S.E., Pickering, S.J. (2000), "Working memory deficits in children with low achievements in the national curriculum at 7 years of age", British Journal of Educational Psychology, 70, 177-194.

Gauthier C., éd. (1997). Pour une théorie de la pédagogie. Recherches contemporaines sur le savoir enseignant. Bruxelles : De Boeck.

Gavens, N., Camos, V. (2006), "La mémoire de travail : une place centrale dans les apprentissages scolaires fondamentaux". In E. Gentaz & P. Dessus (Eds.), Apprentissages et enseignement : Sciences cognitives et éducation (pp. 91-106). Paris : Dunod.

Genlot S., Suchaut B. (2000). L'évaluation des innovations pédagogiques : quelles modalités de coopération entre les différents acteurs ?). Mesure et évaluation en éducation, 23 (1), 2000. pp. 23-42.

Gerard F.M. (2005), L'évaluation des compétences à travers des situations complexes. Actes du Colloque de l'Admee-Europe, IUFM Champagne-Ardenne, Reims, 24-26 octobre 2005.

Gillet P. [éd.] (1991), Construire la formation : outils pour les enseignants et les formateurs. Paris, ESF.

Glasman N.S., Heck R.H. (1987). Evaluation in decision making: the case of assigning teachers to classrooms, Administrator's Notebook, 32 (5), 1-4.

Goldhammer K., Becker G. (1971). Elementary school principals and their schools, Eugene, Center for the Advanced Study of Educational Administration.

Gorard S., Taylor C. (2002). What is segregation ? Sociology, vol.36, n°4, 875-895.

Goux D., Maurin E. (1997). Destinées sociales : le rôle de l'école et du milieu d'origine, Economie et Statistique, n°306, 13-26.

Goux D., Maurin E. (2003). Neighborhood Effects on Performance at School, Paris, CREST (miméo).

Grisay A. (1997). L'évolution des acquis cognitifs et socio-affectifs des élèves au cours de années de collège, Les Dossiers d'Education et Formations, n°88.

Haut Conseil de l'Education (2007). L'école primaire. Bilan des résultats de l'école 2007.

Heck R., Marcoulides G., Glasman N. (1989). The application of causal modeling techniques to administrative decision making : The case of teacher allocation. Educational Administration Quarterly, 25, pp. 253-267.

Heyneman S. (1986)., Les facteurs de la réussite scolaire dans les pays en développement in M.Crahay et D. Lafontaine, L'art et la science de l'enseignement, Bruxelles : Labor.

Hitch, G., Towse, J.N., Hutton, U. (2001), "What limits children's working memory span? Theoretical accounts and applications for scholastic development". Journal of Experimental Psychology: General, 130 (2), 184-198.

Jarousse J. P., Mingat A. (1993). Les disparités d'acquisitions scolaires en CE2 : caractéristiques individuelles, contexte scolaire et social de scolarisation, effet d'école et de circonscription. Rapport à la Direction de l'Évaluation et de la Prospective, Ministère de l'Éducation Nationale, 14 p.

Jarousse J.P., Leroy-Audouin C., Suchaut B. (1999). Les vacances d'été nuisent-elles aux acquisitions scolaires ? Carrefours de l'éducation, N°8, pp. 2-25.

Joutard P., Thélot C. (1999), Réussir l'école. Le seuil, 280 p.

Just M.A., Carpenter P.A. (1992); "A capacity theory of comprehension : Individual differences in working memory". Psychological Review, 99, 122-149.

Lafontaine D., Baye A., Matoul A., Demeuse M., Demonty I. (2001). PISA, Rapport final sur la campagne 2000 en communauté française de Belgique. Université de Liège. Service de Pédagogie Expérimentale.

Lapostolle G., Suchaut B. (2007). Les résultats des recherches en Education et leur prise en compte dans les politiques éducatives depuis les années 1980. Revue Eduquer. A paraître.

Laveault D., Grégoire J. (1997), Introduction aux théories des tests en sciences humaines. De Boeck Université. 336 p.

Leamer E. (1983), Let's take the con out of Econometrics, The American Economic Review, vol. 73, N°1, 31-43.

Le Bastard-Landrier S., Suchaut B. (2000). L'élève comme objet d'évaluation au sein d'une démarche d'apprentissage de la lecture. Mesure et évaluation en éducation, 23 (2-3), 2000. pp. 21-36.

Lee V.E., Dedrick R.F., Smith J.B. (1991). The effect of the social organization of schools on teachers' efficacy and satisfaction, Sociology of Education, 64, pp. 190-208.

Lépine, R., Barrouillet, P., Camos, V. (2005). "What makes working memory spans so predictive of high level cognition ?" Psychonomic Bulletin and Review, 12, 165-170.

Lépine, R., Bernardin, S., Barrouillet, P. (2005). "Attention switching and working memory spans". European Journal of Cognitive Psychology, 17, 329-345.

Leroy-Audouin C., Mingat A. (1995), L'école primaire rurale en France : structure des classes, efficacité pédagogique et intégration au collège. Rapport pour la direction de la prévision, Ministère de l'économie. 51 p. Paris

Leroy-Audouin C., Suchaut B. (1994), Modes de groupements et apprentissages des élèves à l'école maternelle, communication à la Deuxième Biennale de l'Education et de la Formation, 9-12 avril, Paris, la Sorbonne.

Leroy-Audouin C., Suchaut B. (2005). A chaque classe ses élèves. Procédures et critères d'affectation à l'école élémentaire. Revue française de pédagogie, 2005, n° 152, pp. 89-105.

Leroy-Audouin C., Suchaut B. (2007a). L'attribution des classes aux enseignants : le cas des écoles primaires. Carrefours de l'éducation, N°23, janvier-juin 2007, pp. 71-84.

Leroy-Audouin C., Suchaut B. (2007b). Revisiter l'efficacité pédagogique des classes à plusieurs cours. Revue Française de Pédagogie, N°160, juillet-août-septembre 2007. A paraître.

Leroy-Audouin C., Suchaut B. (2007c). Effects of combination classes on student achievement in French elementary schools. School Effectiveness and School Improvement. A paraître.

McLean, J.F., Hitch, G.J. (1999), "Working memory impairments in children with specific arithmetic learning difficulties", Journal of Experimental Child Psychology, 74, 240-260.

Marsh H.W., Hau K-T. (2003), Big-Fish-Little-Pond Effect on Academic Self-Concept, American Psychologist, 58, n°5, 364-376.

Maroy C., Dupriez V. (2000). La régulation dans les systèmes scolaires, proposition théorique et analyse du cadre structurel en Belgique francophone, Revue Française de Pédagogie, n° 130, pp. 73-87.

Mason D., Burns R. (1996), « Simply no worse and simply no better » may simply be wrong: a critique of Veenman's conclusion about multigrade classes. Review of Educational research, 66, 307-322.

Mason D., Burns R. (1997), Reassessing the effects of combination classes, Educational Research and Evaluation, 3, 1-53.

Maurice M., Sellier F., Silvestre J.J. (1982). Politique d'éducation et organisation industrielle en France et en Allemagne. Paris : PUF.

MEN (1975), Tableaux des enseignements et de la formation : édition 1975. Paris, Ministère de l'Éducation Nationale

M.E.N. (1982), Les classes à plusieurs cours dans le premier degré, présentation statistique, Document de travail n°280.

MEN (1985), Repères et références statistiques sur les enseignements et la formation, Paris, Ministère de l'Éducation Nationale

M.E.N. (1996), Enquête sur le fonctionnement des écoles élémentaires. Note d'Information, 96-43.

M.E.N. (1997). Enquête internationale sur le fonctionnement des écoles élémentaires. Les dossiers d'Education et Formations, n° 85.

MEN DEP (1990), Les dossiers Education et Formations : le compte de l'éducation années 1974 - 1988, 8.

MEN DEP (1998), Le compte de l'éducation et le compte de l'éducation supérieure, années 1994 – 1997, Les dossiers, 105.

M.E.N. D.P.D. (1998). La lecture en CM2 : comparaison des résultats en lecture des élèves en fin de CM2 à dix ans d'intervalle (1987-1997), Les dossiers, 102, 143 p.

M.E.N.R.T. (1999), Enquête dans les écoles maternelles, élémentaires et spéciales, Tableaux statistiques, 6623, DPD C1/1CJ.

M.E.N Direction de l'Enseignement Scolaire (2000). L'exploitation de l'évaluation nationale en CE2 : la lecture. Programme national de pilotage : actes du séminaire national. C.R.D.P de l'Académie de Versailles).

MEN DEP (2002), Le compte de l'éducation et le compte de l'éducation supérieure, années 1998 – 2001, Les dossiers, 140

MEN DEP (2005), Le compte de l'éducation 2005, version électronique fournie par MEN DEPP A3

MEN (2006), Repères et références statistiques sur les enseignements, la formation et la recherche. Ministère de l'Éducation Nationale.

Merton R.K (1938) "Social structure and anomie", American Sociological Review, 3, 672-682.

Meuret D. (1994). L'efficacité de la politique des zones d'éducation prioritaire dans les collèges, Revue française de pédagogie, N°109, pp. 41-64.

Meuret D. (2001). Les recherches sur la réduction de la taille des classes. Paris, Haut Conseil de l'Evaluation de l'école, coll. « Les rapports », N°1, 39 p.

Meuret D. (2003). Efficacité et équité des collèges. L'effet établissement. In Derouet J.L. (sous la dir. de), Le collège unique en question, Paris : PUF.

Miller B. (1990), A review of the quantitative research on multigrade instruction, Research on Rural Education, 7 (1), 1-8.

Mingat A. (1983). Evaluation analytique d'une action Z.E.P. au Cours Préparatoire, Les Cahiers de l'IREDU, 37, 126 p.

Mingat A. (1984). Les acquisitions scolaires des élèves au cours préparatoire : les origines des différences. Revue française de pédagogie. N°69. pp.49-62.

Mingat A., Suchaut B. (2000). Les systèmes éducatifs africains : une analyse économique comparative. Bruxelles, De Boeck Université, 2000, 312 p. (Coll. Pédagogies en développement).

Mingat A., Suchaut B. (1996). Incidences des activités musicales en grande section de maternelle sur les apprentissages au cours préparatoire. Les sciences de l'éducation pour l'ère nouvelle, Vol. 29, N°3/1996, pp. 49-76

M.J.E.N.R. (2003), L'évolution du réseau des écoles primaires, Rapport de l'Inspection Générale à M. le Ministre de la Jeunesse, de l'Education Nationale et de la Recherche.

Moissan C., Simon J. (1997). Les déterminants de la réussite scolaire en zone d'éducation prioritaire, Rapport pour le Ministère de l'Education nationale.

Monk D. (1992), Educational productivity research : An update and assessment of its role in education finance reform, Educational Evaluation and Policy Analysis, 14, pp. 307-332.

Monteil J-M., Huguet P. (2002). Réussir ou échouer à l'école : une question de contexte ? Grenoble : PUG.

Morlaix S. (2000). Rechercher une meilleure répartition du temps scolaire en primaire pour favoriser la réussite au collège. Revue française de pédagogie. N°130. pp.121-131

Morlaix S. (2002). Intérêts et apports de l'analyse des variables latentes pour le chercheur en sciences sociales : exemple d'application à l'économie de l'éducation. Orientation scolaire et professionnelle, vol 31., N°1. pp.117-138.

Morlaix S., Suchaut B. (2007a). Evolution et structure des compétences des élèves à l'école élémentaire et au collège : une analyse empirique des évaluations nationales. Cahiers de l'Irédu, N°68 IREDU, mai 2007, 259 p.

Morlaix S., Suchaut B. (2007b). Identification des compétences à l'école élémentaire : une approche empirique à partir des évaluations institutionnelles. Mesure et évaluation en éducation. A paraître. Juin 2007.

OCDE (2001a), Knowledge and skills for life. First results from PISA 2000, Paris : OCDE.

OCDE (2001b), Choice of assessment tasks and the relative standing of countries in PISA 2000. Document pour la 13ème réunion du conseil des pays participants, Paris.

OCDE (2001c), Teachers for Tomorrow's Schools, Paris : OCDE.

OCDE (2003), Literacy Skills for the World of Tomorrow, Paris : OCDE.

Ouvrard F. (1990), Les petits établissements scolaires, Education et Formations, 43, 113-116.

Opdenakker M-C., Van Damme J. (et el.) (2002). "The Effects of Schools and Classes on Mathematics Achievement", School Effectiveness and School Improvement, 13, n°4, 399-428.

Passeron J-C. (1991). Le raisonnement sociologique, Paris : Nathan.

Paul J. J. (1996), Le redoublement: pour ou contre ? Editions ESF, Collection Pratiques et enjeux pédagogiques.

Piketty T., (2004). L'impact de la taille des classes et de la ségrégation sociale sur la réussite scolaire dans les écoles françaises. texte disponible sur le site pythie.cepremap.ens.fr.

Piquée, Suchaut (2002). Les opinions des acteurs confrontées à l'évaluation externe : le cas d'une action d'accompagnement scolaire au CP. Les Sciences de l'Education pour l'ère nouvelle, 35 (3), 2002. pp. 103-126

Piquée C., Suchaut B. (2004). Un maître supplémentaire dans la classe. Quels effets sur les progressions au cycle III ? Revue Française de Pédagogie, N°146, janvier-mars 2004, pp. 91-103.

Plassard J.M., Larré F. (2005). L'enseignant dans tous ses états. Note n°416. LIRHE. Septembre 2005, 24 p.

Pratt D. (1986), On the merits of multiage classrooms. Research on Rural Education, 3 (3), 111-115.

Reynaud J.D. (1993). Les règles du jeu. L'action collective et la régulation sociale, Paris : A. Colin.

Robin I., Rocher T., 2002, La compétence en lecture des jeunes de 15 ans : une comparaison internationale, Données sociales, INSEE, pp.93-101.

Singer J.D. (1998), "Using SAS PROC MIXED to fit Multilevel Models, Hierarchical Models, and Individual Growth Models". Journal of educational and Behavioral Statistics. Winter 1998, Vol. 24, N°4, pp 323-355

Slavin R., 1990, Achievement Effects of Ability Grouping in Secondary Schools : A Best-Evidence Synthesis. Review of Educational Research, 60, 3, 471-499.

Solau G. Dir. (2000a). L'évaluation des politiques d'éducation. CRDP Dijon, 2000.

Solau G. (2000b). De la complexité particulière de l'évaluation des politiques d'éducation dans l'univers de l'évaluation des politiques In L'évaluation des politiques d'éducation, pp. 7-31, Solau dir. Dijon : CRDP, 2000 - Coll. "Documents, actes et rapports pour l'éducation", 189 p.

Solau G. (2002). Les modalités de changements dans le système éducatif français. Politiques et Management Public, Vol 20, N°2, juin 2002, pp. 81-96.

Sorensen A.B., Morgan S.L. (2000). School Effects. Theoretical and Methodological Issues. In M.T.Hallinan (ed), Handbook of the Sociology of Education, NY, Kluwer Academic, 137-160

Suchaut B. (1996a). Le temps scolaire : allocation et effets sur les acquisitions des élèves en grande section de maternelle et au cours préparatoire. Thèse de doctorat en Sciences de l'Éducation, Université de Bourgogne, Dijon.

Suchaut B. (1996b). La gestion du temps à l'école maternelle primaire : diversité des pratiques et effets sur les acquisitions des élèves. L'année de la recherche en sciences de l'éducation, 1996, pp. 123-153

Suchaut B. (1998). L'individualisation de l'enseignement de la lecture : éléments d'évaluation d'un dispositif pédagogique au cours préparatoire. Les sciences de l'éducation pour l'ère nouvelle, vol. 31, n°4, pp. 41-67.

Suchaut B. (2003a). La lecture au CP. Les déterminants de la réussite in : Vers la maîtrise de l'écrit, par tous. Association pour Favoriser une Ecole Efficace (APFEE). Lyon, Editions Aléas, 2003, (pp. 29-39).

Suchaut B. (2003b). De la nécessité d'évaluer les pratiques enseignantes : entre enjeux sociaux et obstacles méthodologiques. Les Dossiers des sciences de l'éducation, Presses Universitaires du Mirail. N°10, 2003. pp. 17-30.

Suchaut B. (2004). Les différences et les inégalités de réussite à l'école primaire : enseignements, portée et utilité des résultats de la recherche en éducation. 4ème Université d'automne du S.N.U.I.P.P., La Londe-les-Maures, 22/23/24 octobre 2004, SNUIPP. 15 p.

Suchaut B. (2005a). L'évaluation des dispositifs éducatifs et de formation. Table ronde au colloque ADMEE, 24-26 octobre 2005 "Comment évaluer ? Outils, dispositifs, acteurs", IUFM Champagne-Ardenne, Reims, 5 p.

Suchaut B. (2005b). Regard sur l'efficacité des dispositifs de lutte contre l'échec scolaire. Regards sur l'Actualité, N°310, avril 2005, pp. 51-58

Suchaut B. (2006). Les évaluations nationales en France : outils pédagogiques et outils de pilotage : intérêts, limites et nouvelles perspectives. 19<sup>ème</sup> colloque de l'ADMEE, Luxembourg, 11-13 septembre 2006, 17 p.

Swidler A. (1979). Organization without authority: Dilemmas of Social Control in Free Schools. Cambridge : Harvard University Press.

Thélot C. (1993). L'évaluation du système éducatif. Paris, Nathan Université.

Thrupp M., (1999). Schools Making a Difference. Let's be Realistic ! Buckingham : Open University Press.

Toczek M.C., Martinot D. ed. (2004). Le défi éducatif. Armand Colin Editeur, Paris 2004, 351 p.

Turner, M. L., Engle, R.W. (1989), "Is working memory capacity task dependent?" Journal of Memory and Language, 28, 127-154.

Vandenberghe V. (2001). Efficacité-équité dans l'enseignement secondaire : quelle relation ?, Reflets et Perspectives de la vie économiques, vol.XL, n°4 (reproduit in Problèmes économiques, n°2.753, 20 mars 2002, 12-16).

Zanten (van) A., 2001, L'école de la périphérie, Paris : PUF.

Zanten (van) A., (1996). Fabrication et effets de la ségrégation scolaire, in S.Paugam (dir.), L'exclusion, l'état des savoirs, Paris : La Découverte.

Veenman S. (1995), Cognitive and noncognitive effects of multigrade and multi-age classes: A best-evidence synthesis. Review of Educational Research and Evaluation, 65, 319-381.

Veenman S. (1997), Combination Classrooms Revisited, Educational Research and Evaluation, Vol. 3, n°3, 262-276.

Vial M. (1997), Les modèles de l'évaluation. Textes fondateurs et commentaires, éd. De Boeck.

Vogler J., Bouissou B. (1987), Evaluation pédagogique dans les écoles. Ecoles rurales et écoles urbaines, Education et Formations, 10, 3-9.

Walsh J. M. (1989), Attitudes of students, parents and teachers toward multigrading. Unpublished doctoral dissertation, University of Connecticut, Storrs.

Warren R. (1975). Context and isolation : The teaching experience in the elementary school, Human Organization, 32 (2), pp. 139-148.

Wilkinson I. (1998), Dealing with diversity: Achievements gaps in reading literacy among New Zealand students. Reading Research Quarterly, 33, 144-167.

Wilkinson I., Hamilton R. (2003), Learning to read in composite (multigrade) classes in New Zealand : teachers make the difference. Teaching and Teacher Education, 19, 221-235.

Willms J.D., Somers M-A. (2001). Family, Classroom and School Effects on Children's Education outcomes in Latin America. School Effectiveness and School Improvement, 12, n°4, 409-445.

Zachary M-D. (2001). Structures scolaires et équité des systèmes d'enseignement en Europe, Wallonie, n°66, 93-102.

## **Liste des tableaux**

Tableau 1 : Comparaison des performances de la France dans les enquêtes internationales en compréhension en lecture.....	p. 13
Tableau 2 : Comparaisons internationales des dépenses publiques d'éducation (2003) .....	p. 20
Tableau 3 : Effectifs d'élèves (en milliers) par niveaux d'enseignement entre 1974 et 2005 .....	p. 24
Tableau 4: Structure de la DIE par niveaux d'enseignement (millions d'euros prix 2005) .....	p. 25
Tableau 5 : Evolution comparée des ressources et des effectifs par niveau d'études (1974-2005) .....	p. 25
Tableau 6 : Coûts unitaires de scolarisation par niveau d'études après ventilation des dépenses non affectées (1974 – 2005) .....	p. 26
Tableau 7: Dépenses des administrations publiques du premier degré par type de financeur (1974-2005 .....	p. 28
Tableau 8 : Synthèse du poids des facteurs ayant influencé l'évolution des coûts unitaires dans le premier degré, 1974– 2005.....	p. 29
Tableau 9 : Décomposition de l'augmentation du coût unitaire de scolarisation du premier degré entre 1974 et 2005.....	p. 31
Tableau 10 : Eléments de comparaison des coûts unitaires de scolarisation (2003) .....	p. 33
Tableau 11 : Modèle expliquant les inégalités sociales par la différenciation des systèmes et la dispersion des résultats des élèves en écrit.....	p. 57
Tableau 12 : Type de classe occupé selon que l'enseignant est le dernier arrivé ou pas dans l'école.....	p. 71
Tableau 13 : Description des caractéristiques des élèves selon le type de cours fréquenté.....	p. 81
Tableau 14: Modèles multiniveau de progression en CE1, score de français, de mathématiques et global (échantillon complet).....	p. 82
Tableau 15 : Effets des différentes configurations de type de cours sur les progressions au CE1 (échantillon complet) .....	p. 83

Tableau 16 : Modèles multiniveau de progression en CM1, score de français, de mathématiques et global (échantillon complet).....	p. 84
Tableau 17 : Effets des différentes configurations de type de cours sur les progressions au CM1 (échantillon complet) .....	p. 85
Tableau 18 : Modèles multiniveau de progression en CE1, score de français, de mathématiques et global (écoles qui ne comptent qu'une classe de CE1) .....	p. 87
Tableau 19: Effets des différentes configurations de type de cours sur les progressions au CE1 (écoles qui ne comptent qu'une classe de CE1) .....	p. 87
Tableau 20 : Modèles multiniveau de progression en CE1, score de français, de mathématiques et global (écoles qui offrent le choix, dont un cours double).....	p. 88
Tableau 21 : Effets des différentes configurations de type de cours sur les progressions au CE1 (écoles qui offrent le choix, dont un cours double).....	p. 89
Tableau 22: Modèles multiniveau de progression en CM1, score de français, de mathématiques et global (écoles qui ne comptent qu'une classe de CM1) .....	p. 89
Tableau 23 : Effets des différentes configurations de type de cours sur les progressions au CM1 (écoles qui ne comptent qu'une classe de CM1).....	p. 90
Tableau 24 : Modèles multiniveau de progression en CM1, score de français, de mathématiques et global (écoles qui offrent le choix, dont un cours double)..	p. 91
Tableau 25 : Effets des différentes configurations de type de cours sur les progressions au CM1 (écoles qui offrent le choix, dont un cours double) .....	p. 92
Tableau 26 : Relation entre score initial et tonalité sociale de la classe au niveau agrégé..	p. 98
Tableau 27 : Progressions des élèves (explication du score final par le score initial) au niveau agrégé.....	p. 101
Tableau 28 : Modèles agrégés de progressions intégrant la tonalité sociale de la classe..	p. 101
Tableau 29 : Modèles agrégés de progressions intégrant la tonalité sociale de la classe, le niveau moyen et l'hétérogénéité.....	p. 102
Tableau 30 : Modèles multiniveaux de progression en CE1.....	p. 104
Tableau 31 : Modèles multiniveaux de progression en CM1.....	p. 104

Tableau 32 : Effets de la non linéarité de la tonalité sociale sur les progressions au CE1 .....	p. 105
Tableau 33 : Effets des variables d'interaction entre variables d'origine sociale et de tonalité sociale de la classe dans les modèles de progression au CE1.....	p. 107
Tableau 34 : Scores moyens aux épreuves de CP, CE2 et 6 <sup>ème</sup> (panel 1997) .....	p. 117
Tableau 35 : Modèles de progression globale des élèves (panel 1997) .....	p. 118
Tableau 36 : Modèle analysant la variance du score de 6 <sup>ème</sup> en fonction des différentes dimensions de l'évaluation de début CP.....	p. 121
Tableau 37 : Exemple d'un regroupement d'items sur la base des corrélations (évaluation CE2, 1999) .....	p. 127
Tableau 38 : Modèle analysant la prédictivité des variables latentes sur le score global de CE2 (régression pas à pas) .....	p. 134
Tableau 39 : Relations entre l'échec et la réussite aux compétences de CE2.....	p. 135
Tableau 40 : Corrélations entre scores cognitifs et scores aux évaluations nationales....	p. 149
Tableau 41 : Scores moyens aux tests cognitifs et origine sociale .....	p. 150
Tableau 42 : Modèles de régression expliquant les progressions des élèves au cycle III, en fonction des acquisitions initiales et des capacités en mémoire de travail.....	p. 151
Tableau 43: Modèles de régression expliquant les progressions des élèves au cycle III, en fonction des acquisitions initiales, des capacités en mémoire de travail, et de l'origine sociale.....	p. 152

## **Liste des graphiques**

Graphique 1 : Evolution du niveau moyen des élèves de 4 <sup>ème</sup> année primaire en compréhension en lecture entre 1990 et 2001.....	14
Graphiques 2 et 3 : Evolution de l'efficience qualitative de l'école primaire de 1990 à 2001	16
Graphique 4 : Evolution temporelle des résultats des évaluations nationales de français au CE2 entre 1996 et 2005.....	p. 18
Graphique 5 : Espérance de Vie Scolaire selon le PIB par habitant.....	p. 22
Graphique 6 : Espérance de vie scolaire selon la part des dépenses publiques d'éducation dans le PIB.....	23
Graphique 7 : Evolution (base 100 en 1974) des effectifs par ordre d'enseignement (1974-2005) .....	p. 24
Graphique 8 : Niveau moyen et dispersion en compréhension de l'écrit.....	p. 44
Graphique 9: Pourcentages d'élèves en dessous du niveau 1 et au-dessus du niveau 5 de compétences en compréhension de l'écrit.....	p. 47
Graphique 10 : Niveau moyen et dispersion en compréhension de l'écrit par grandes régions.....	p.48
Graphique 11 : Indice global d'inégalités sociales et niveau moyen de performance des pays.....	p. 51
Graphique 12 : Inégalités sociales de performances et différenciation des systèmes éducatifs .....	p.56
Graphique 13 : Analyse en composantes principales des caractéristiques des pays.....	p. 58
Graphique 14 : Répartition des différentes configurations de classes.....	p. 69
Graphique 15 : Niveau initial des classes et proportion d'élèves de milieu favorisé, en CE1.....	p. 97
Graphique 16 : Niveau initial moyen des classes et degré d'hétérogénéité, au niveau CE1.....	p.99
Graphique 17 : Progressions différentielles des élèves selon la tonalité sociale de la classe au CE1.....	p. 108
Graphique 18 : Structure temporelle de l'évaluation des élèves (panel 1997) .....	p. 116

Graphique 19 : Relation entre les scores de CP et de 6 <sup>ème</sup> (panel 1997) .....	p. 118
Graphique 20 : Modèle de mesure entre deux variables latentes et leurs indicateurs.....	p. 129
Graphique 21: Modèle de mesure initial.....	p.130
Graphique 22 : Modèle de mesure amélioré.....	p. 131
Graphique 23 : Blocs de compétences des évaluations CE2.....	p. 133
Graphique 24: Structure pyramidale des compétences des élèves au CE2.....	p. 136
Graphique 25 : Relations statistiques entres compétences de CE2 et de 6 <sup>ème</sup> .....	p.140
Graphique 26: Analyse structurelle entre compétences de CE2 et compétences de 6 <sup>ème</sup> ...	p.143
Graphique 27: Décomposition de la variance du score de 6ème en fonction du score de CE2 et du score en mémoire de travail.....	p. 151
Graphique 28 : Résultats de l'analyse en pistes causales entre origine sociale, mémoire de travail et progrès scolaires au cycle III.....	p. 153
Graphique 29: Relation entre score en mémoire de travail, progrès des élèves au cycle III et origine sociale.....	p.154