

*Les recherches
sur la réduction de
la taille des classes*

Denis MEURET
Université de Bourgogne, IREDU

Janvier 2001

Rapport établi à la demande du
Haut Conseil de l'évaluation de l'école

Les recherches sur la réduction de la taille des classes

LE DÉBAT, LES POLITIQUES	5
QUESTIONS DE MÉTHODE	8
LES RECHERCHES SUR LES ACQUISITIONS ACADÉMIQUES DES ÉLÈVES	12
Les comparaisons internationales	12
Les recherches étrangères	12
Les recherches françaises	19
Enseigne-t-on différemment dans les petites classes ?	21
LES EFFETS NON ACADÉMIQUES	24
LA QUESTION DU COÛT	26
CONCLUSION	29

BIBLIOGRAPHIE	31
PRÉSENTATION SYNTHÉTIQUE DES PRINCIPALES RECHERCHES SUR L'EFFET D'UNE RÉDUCTION DE LA TAILLE DES CLASSES.	34
Méta-analyses	34
Quelques recherches étrangères sur les acquisitions de connaissances	35
Quelques recherches sur d'autres dimensions	37
Recherches françaises	38

Une réduction de la taille des classes (RTC) est susceptible d'avoir de multiples effets, tant en ce qui concerne l'apprentissage et le développement des élèves qu'en ce qui concerne le coût de l'enseignement. Ses partisans en attendent des avantages considérables, tandis que ses adversaires insistent sur la minceur de ceux qu'elle permet d'obtenir pour un coût important. D'autres politiques, selon eux, seraient plus efficaces.

Les recherches sur cette question sont nombreuses et anciennes (depuis les années 20) aux Etats Unis, où elles ont connu récemment un net regain. On en trouve aussi au Royaume Uni, dans les pays scandinaves, et en France¹.

En proposant ici une synthèse, on ne prétendra évidemment pas décider de cette question. Il ne s'agit pas là d'une affectation de modestie scientifique : dans certains domaines, la recherche en éducation permet un avis tranché, sur la nocivité du redoublement par exemple (Crahay, 1996). *Il se trouve seulement que les effets d'une réduction de la taille des classes apparaissent faibles, et dépendants de conditions diverses : leur ampleur, la taille de départ, la population ou encore la nature de l'effet considéré*. Dès lors, une décision sur la taille des classes suppose un choix entre cette politique et d'autres, qui ne peut être instruit par des recherches sur la taille des classes seule. Dès lors aussi, certains chercheurs sont favorables à la RTC, d'autres non, et c'est surtout un débat argumenté que la recherche peut ici proposer au public et aux décideurs. C'est ce débat, *vif surtout aux Etats Unis*, que l'on s'est efforcé de synthétiser.

Nous présenterons d'abord les arguments théoriques échangés sur la taille des classes et l'évolution des politiques en la matière, puis les questions de méthode qui se posent à qui veut mesurer précisément l'effet de la RTC, puis les résultats de diverses recherches sur ces effets : celles qui portent sur les apprentissages (d'abord d'après les comparaisons internationales, puis d'après les recherches étrangères, puis d'après les recherches françaises), puis celles qui portent sur les effets non académiques de la scolarisation. Enfin, on intégrera dans la réflexion quelques politiques alternatives, ce qui permettra de poser la question de l'efficacité.

¹ On en trouve aussi sur les pays en voie de développement, que l'on n'évoquera pas ici.

Le débat, les politiques

Il fut un temps où les pédagogues eux-mêmes privilégiaient les grandes classes. De Peretti (1987) cite Comenius, mort en 1670, “ Plus grand est le nombre d’élèves qu’il voit devant lui, plus grand est l’intérêt que l’enseignant prend à son travail. Pour les élèves, de la même manière, la présence d’un bon nombre de compagnons sera productrice non seulement d’utilité mais aussi de plaisir ”. D’autres arguments étaient avancés : l’enseignement est plus facile dans les grandes classes parce qu’il est obligatoirement plus formel ; ces dernières offrent moins de prise aux “ effets de clique ” entre élèves.

De fait, dans les enseignements primaire et secondaire, on s’accommodait jusqu’au début du XXème siècle, dans les villes, de classes d’une centaine d’élèves. Encore dans les années 60, on pouvait rencontrer des classes de 50 élèves (De Peretti,1987).

Aujourd’hui, non seulement les enseignants, mais aussi les élèves et leurs parents souhaitent des classes plus petites. Certainement, il s’agit d’une facette d’un mouvement général de désaffection vis à vis des grands groupes, associés à une conception rigide des relations sociales. Mais les partisans des petites classes ne manquent pas d’arguments spécifiques : des classes plus petites, disent-ils, permettent un enseignement plus adapté à chaque élève, permettent de faire droit à la dimension affective que réclament tant les enfants que les adolescents, permettent la mise en œuvre du “conflit socio-cognitif”, qui est “ le moteur du véritable apprentissage ”, de faire face à la “ faible capacité d’écoute des élèves, à leur besoin constant de nouvelles activités ”². D’autres arguments sont avancés : dans les petites classes, le temps passé à des tâches bureaucratiques, à maintenir la discipline, est moindre ; la participation et la présence des élèves sont plus grandes ; la communication entre les parents et les enseignants est accrue (cf. par exemple, Texas Education Agency, 2000).

De fait, l’administration a organisé ou laissé faire une baisse de la taille des classes ou des groupes : De 1966 à 1996, en France, la taille des classes a diminué de 47% en Maternelle, de 20% en Élémentaire, de 11% en collège et de 3% dans les lycées d’enseignement général et technique³. Encore ce dernier chiffre est-il trompeur puisque, dans ces lycées, le taux d’encadrement des élèves par les enseignants, lui, a augmenté sur la même période de 14%, en partie parce que, au sein de classes aussi grandes qu’auparavant, les élèves se voient proposer plus souvent des enseignements en petits groupes. De fait, si l’on s’intéresse non à la taille de la classe, mais à la taille moyenne de l’ensemble des groupes dans lesquels un élève suit sa scolarité, pondérée par la durée de chaque groupe, on obtient ce que la Direction de la Programmation et du Développement (DPD) appelle la “ taille moyenne des structures ”, qui est aussi le nombre moyen d’élèves qu’un enseignant a en face de lui, lequel est significativement inférieur à la taille des classes, comme le montre le tableau suivant :

² Je reprends ici les arguments d’une enseignante, H. Eveleigh, dans un article du dossier consacré par les Cahiers Pédagogiques à “ Travailler aussi en grand groupes ” (juin 2000).

³ Cf. MEN-DPD, Notes d’information 96-45, 97-38.

Tableau 1 : Taille des classes et taille des structures à la rentrée 1999		
	Taille des classes	Taille des structures
Collèges hors SEGPA	24.2	22.8
Lycées professionnels	21.1	16.2
Lycées pré-bac	28.8	23.4
Lycées CPGE	33.6	25.8

Source : MEN-DPD, NI 00.12, enseignement public seulement.

Une évolution semblable s'observe dans la plupart des pays développés, à l'exception notable des pays asiatiques. Aux Etats Unis, le nombre d'élèves par enseignant, dans l'enseignement élémentaire et secondaire, est passé de 27 en 1950 à 17 en 1990, puis s'est stabilisé jusqu'en 1994 (Hanushek, 1998). Pour une part, cette évolution est due au développement de l'enseignement spécialisé, en sorte que la taille moyenne des classes est plus élevée que ne le laisserait penser la valeur du taux d'encadrement. Pour l'ensemble du pays, on comptait en 1994, 24 élèves par classe, tant dans l'enseignement élémentaire que dans l'enseignement secondaire.

La liste massive des avantages supposés accompagner la RTC se heurte cependant à des faits tout aussi massifs. Notamment qu'aux Etats Unis, depuis qu'on dispose d'épreuves régulières permettant de mesurer l'évolution des performances des élèves à différents niveaux de l'enseignement (1970), ces performances sont en gros restées stables (elles ont baissé au cours des années 70, ont remonté depuis les années 80), alors que le taux d'encadrement est passé de 22 à 17 pendant cette période (Hanushek, 1998). Divers faits pourraient exonérer la RTC de cette absence apparente d'efficacité (la proportion d'élèves vivant en dessous du seuil de pauvreté a augmenté, ainsi que la proportion d'enfants vivant avec un seul parent) mais d'autres facteurs ont favorisé pendant la même période la performance des élèves (le niveau d'éducation des parents a crû, le nombre d'enfants par famille a diminué, à quoi on pourrait ajouter, depuis les années 80, la mise en place d'une régulation fondée sur la mesure des résultats des élèves⁴). Hanushek montre aussi que l'effet de la croissance du nombre d'élèves dans l'enseignement spécialisé et l'amélioration de leur prise en charge n'explique qu'un tiers de la hausse du taux d'encadrement, en sorte que le taux d'encadrement des élèves "normaux" a bel et bien augmenté sans résultat notable. On voit cependant que des approches longitudinales de ce type, si elles peuvent nourrir le soupçon sur l'efficacité de la RTC, et plus généralement sur celle de l'accroissement du taux d'encadrement, ne peuvent être tenues pour une preuve définitive de leur inefficacité.

A tel point d'ailleurs, que, sous l'influence d'une étude à large échelle dont les résultats ont été favorables à la RTC (le projet STAR, dont il sera question plus loin), les Etats Unis se sont lancés dans une politique résolue en ce domaine, dont l'objectif est d'amener en sept ans la taille des classes des trois premières années de l'enseignement primaire à 18 élèves. Des subventions fédérales – 1,2 milliards de \$ en 1999, 1,3 en 2000 – sont allouées aux districts à cet effet. Les districts doivent allouer 80 % de ces fonds aux écoles en fonction du nombre de leurs élèves "vivant dans la pauvreté", et 20 % en fonction de leur nombre total d'élèves (USDE, 2000). Certains états ont anticipé ou complètent cette action par des politiques et des subventions qui leur sont propres. Trois programmes sont souvent cités :

⁴ Voir Meuret, D. , Broccholoichi, S. et Duru-Bellat, M., 2000.

- La politique californienne de RTC, qui date de 1996, a consisté à abaisser la taille maximum des classes de trente à vingt dans les trois premières années de l'enseignement primaire. Cette expérience a mis en lumière deux problèmes dont le programme national tient compte. D'une part, les écoles des zones défavorisées ont moins rapidement mis en œuvre la RTC que les autres (d'où la priorité qui leur est accordée), d'autre part les nouveaux enseignants, recrutés sur un marché du travail tendu, étaient moins qualifiés que les autres (d'où le fait que 25% des crédits alloués peuvent être utilisés pour leur formation) (USDE, 2000 ; CSRRRC, 2000).
- Au Wisconsin, le programme SAGE (Student Achievement Guarantee in Education) vise à augmenter le taux d'encadrement à un enseignant pour 15 élèves dans les zones les plus défavorisées, toujours aux trois premiers niveaux de l'enseignement primaire. En outre, SAGE prévoit un renforcement des liens entre les écoles et la communauté, l'adoption d'un curriculum centré sur les apprentissages fondamentaux et des programmes de formation des enseignants, en particulier à l'évaluation.
- En Caroline du Nord (Burke county) a été testé, puis implanté depuis 1990, un projet de RTC à 15 élèves du CP au CE2, accompagné d'une politique de formation continue des enseignants. Ce projet est à plus petite échelle que les autres, environ 1500 élèves à chaque niveau.

Des politiques de réduction importante de la taille des classes au début de la scolarité sont aussi mises en œuvre dans d'autres pays. C'est le cas aux Pays Bas (Annevelink et al., 2001) et en Angleterre.

Ceux qui ont décidé ces politiques avancent qu'elles sont fondées sur les résultats de recherches favorables à la baisse de la taille des classes (USDE 2000, par exemple). S'il en est bien ainsi est ce que nous allons examiner. Auparavant, cependant, il importe de comprendre les grandes lignes de la façon dont ces recherches procèdent.

Questions de méthode

Anticiper les effets d'une politique de RTC suppose qu'on réponde à deux questions distinctes :

- 1) Apprend-on mieux dans les petites classes, ou, plus généralement, les effets sur les élèves y sont-ils meilleurs ?
- 2) Si, de ce fait, on baisse la taille des classes à l'échelle d'un système éducatif, compte tenu de l'ensemble des effets induits, les effets sur les élèves seront-ils durablement meilleurs ?

Les comparaisons longitudinales, voire les comparaisons internationales des performances des élèves visent à répondre directement à la seconde question. Toutefois, on l'a vu à propos des Etats Unis, elles ne donnent pas de réponse sûre. Si une RTC ne donne pas d'effets tangibles, on peut toujours prétendre que, sans elle, le niveau des élèves aurait évolué de façon pire. De même, si, un pays X où la taille des classes est élevée délivre un enseignement plus efficace qu'un pays Y dont les classes sont plus petites, on peut toujours prétendre que l'écart serait encore plus grand si Y avait des classes plus grandes, soit que les élèves de Y, pour des raisons diverses, soient plus sensibles à la taille des classes que ceux de X, soit, si la taille des classes a les mêmes effets partout, que X subisse des handicaps que Y ne connaît pas.

Pour cette raison, on ne peut se limiter à ce type de comparaisons et de nombreuses recherches se sont efforcées de déterminer si, dans un pays et à un moment donné, on apprend mieux dans les petites classes que dans les grandes. Cet exercice rencontre une série de difficultés.

La taille des classes en début d'année ne donne qu'une idée approchée de la taille des groupes dans lesquels les élèves reçoivent vraiment leur enseignement. Non seulement les classes peuvent être momentanément dédoublées, certains élèves peuvent en être momentanément retirés pour une raison ou une autre, mais aussi certains élèves sont momentanément absents. La taille des classes en début d'année, qu'utilisent la plupart des études, est donc seulement une estimation de la taille des groupes que les élèves vivent réellement et qui est celle dont les effets s'observent réellement. " Cela peut expliquer la difficulté que l'on a à observer des effets substantiels " (Goldstein et Blatchford, 1998).

Les effets d'un écart de taille peuvent différer – et, de fait, diffèrent réellement – selon la taille elle-même : passer de 25 à 22 peut avoir des effets différents que de passer de 20 à 17, ou de 35 à 32. Or, évidemment, les différentes études portent sur des écarts, et des tailles, différents, ce qui explique que leurs résultats puissent différer.

Le problème principal, cependant, est que les élèves ne sont pas affectés dans de grandes ou de petites classes indépendamment de leur capacité à progresser. Il faut donc, pour mesurer le vrai effet de la taille des classes raisonner " toutes choses égales par ailleurs ".

Deux approches sont utilisées à cet effet. Dans les études *expérimentales* – analogues à celles qui sont utilisées pour les traitements médicaux – on tire au hasard les élèves que l'on affectera pour les uns dans de grandes classes, pour les autres dans des petites. Les élèves du groupe

expérimental (ceux des petites classes) et ceux du groupe de contrôle ont d'autant plus de chances d'être identiques à tous égards que la ... taille de ces groupes est élevée. Dans le projet STAR, chaque groupe comporte environ 3500 élèves. Ce type d'études permet à l'expérimentateur de décider des tailles qu'il veut comparer, y compris d'étudier les effets de tailles que l'on ne rencontre pas dans le système. Une expérience extrême à cet égard est celle de Bloom (1984) qui, dans les années 70, compara sur onze jours les progrès des élèves en classes de 30 élèves, en classes de 30 élèves utilisant la Pédagogie de Maîtrise et en situation de préceptorat (un maître pour un à deux élèves)⁵.

Un autre avantage des études expérimentales est qu'il est possible, dans une certaine mesure, de tirer au sort les enseignants qui vont être assignés à chaque classe, de façon à éliminer aussi le biais qui tient à ce que la qualité des enseignants peut n'être pas indépendante de la taille des classes auxquelles on les affecte.

A priori, ce type de recherche est le plus rigoureux et le plus fiable. Pourtant, il n'est pas sans présenter quelques inconvénients, dont les principaux sont (Goldstein et Blatchford, 1998 ; Hanushek, 1998 ; Hoxby, 1998) :

- si l'effet de la taille des classes varie à proportion de la présence dans la classe de tel ou tel type d'élèves, il faut espérer que le hasard aura bien fait les choses et réparti ces élèves également dans les deux catégories de classes. Or, a priori, ces "types" peuvent être de natures très diverses, en sorte qu'on n'est jamais complètement sûr que la répartition de tous les "types" d'élèves pertinents soit équilibrée.
- Lorsque l'étude se passe sur plusieurs années, il est impossible d'éviter les phénomènes dits d' "attrition" : certains élèves quittent la cohorte. Si ceux qui quittent les grandes classes sont meilleurs que ceux qui quittent les petites, les résultats sont biaisés en faveur des petites classes, en faveur des grandes classes dans le cas inverse.
- A la différence des évaluations médicales, qui sont en "double aveugle" (ni les malades, ni les médecins, ne savent qui a reçu le traitement expérimenté et qui a reçu des placebos), les études en éducation sont en "zéro-aveugle" (Hanushek, 1998) : les élèves et leurs enseignants savent s'ils sont dans une petite ou une grande classe. D'où deux risques différents : celui de "l'effet Hawthorne" (les membres du groupe expérimental sont stimulés d'être dans une position originale) ; celui de biais moins involontaires, dont le sens est d'ailleurs indécidable : les enseignants des grandes classes peuvent faire plus d'efforts pour compenser ce qu'ils pensent être un handicap ; ceux des petites classes aussi, parce qu'ils sont conscients de l'enjeu.

⁵ L'effet du préceptorat fut de deux écart types, celui de la Pédagogie de Maîtrise de 1 écart type. Sur la pédagogie de maîtrise, voir Crahay (2000). Ce type de mesure est appelé "effet standard". L'effet standard exprime l'écart entre le groupe expérimental et le groupe de contrôle en fractions de l'écart type de la distribution du groupe de contrôle. En cas (fréquent) de distribution symétrique, un gain de 2 écarts-type signifie que le score de l'élève médian (50%) du groupe scolarisé dans le groupe expérimental est immédiatement supérieur à celui de 98 % des élèves de contrôle. Nous rencontrerons pour la RTC des effets standards compris entre 0.15, signifiant que l'élève médian des petites classes fait mieux que 56% des élèves des grandes, et 0.4 (66%), en passant par 0.2 (58 %). De façon plus imagée, un effet standard de 0.2 signifie donc qu'un élève qui serait 50ème sur 100 après une scolarité dans une grande classe, s'il passe en fait son année dans une petite, atteindra un niveau égal à celui de l'élève qui est 42ème (100-58) après une scolarité dans une grande classe. Plus l'effet standard est positif et élevé, plus l'effet de la RTC est fort et positif.

Dans les études *corrélationnelles* (dites parfois aussi observationnelles, parfois aussi économétriques, si elles utilisent des modèles statistiques) on tire parti de la diversité des tailles de classe observables dans le système scolaire tel qu'il est, en corrigeant, au moyen de techniques statistiques, les différences brutes de performances que l'on observe entre les élèves à l'issue d'une classe plus ou moins grande, de celles qui proviennent de facteurs autres que la taille elle-même.

Les principaux facteurs individuels qui, au début d'une année scolaire, permettent de prédire les progrès d'un élève sont bien connus : son niveau dans la matière considérée, son sexe, sa situation sociale, qui est appréhendée de façon diverse selon les pays.

D'autres facteurs de biais se situent, on l'a dit, au niveau des enseignants, mais aussi de la classe elle-même, ou de l'école. Par exemple, une grande classe composée de bons élèves pourra être efficace à cause du second caractère et non du premier. Un établissement efficace, donc réputé, et donc demandé, aura de plus grandes classes qui seront la conséquence et non la cause de son efficacité. Pour corriger ce type de biais, il faudra introduire des variables qui se situent au niveau de la classe ou de l'école, et non au niveau de l'élève, et cela dans une fonction statistique qui recherche le lien entre la performance de *chaque élève* et la taille de la classe qui l'accueille. Cela peut se faire en affectant à chaque élève les variables décrivant sa classe, ou son école, mais cette façon de faire occasionne des biais, et une sous estimation de l'incertitude des résultats, qui sont fortement atténués par l'usage de modèles dits " multiniveaux " qu'utilisent les études les plus récentes (Goldstein et Yang, 2000 ; Nye et al., 2000 par exemple).

Toutefois, avec ce type d'approche, on est jamais sûr d'avoir pris en compte *tous* les facteurs parasites, d'autant qu'on est limité aux facteurs mesurables. Or, par exemple, peu de facteurs aisément mesurables rendent compte de l'efficacité des enseignants, en sorte qu'il est difficile de corriger vraiment de ce facteur là, pourtant essentiel. De plus, on ne peut ainsi corriger de facteurs liés aux stratégies des acteurs. Si toutes choses égales, les parents plus soucieux de la réussite de leurs enfants s'arrangent pour qu'ils soient dans de petites classes⁶, et puisque ce souci est effectivement un facteur de réussite, cette approche, même dans sa forme multiniveau, attribuera à la taille des classes les effets des comportements qui accompagnent ce souci.

D'où une autre forme d'étude corrélacionnelle (Hoxby, 2000) qui consiste, toujours par des techniques statistiques, à ne retenir parmi toutes les différences de tailles des classes, que celles qui sont dues au hasard, par exemple au fait que, s'il existe un seuil maximum de 30 élèves par classe, l'effectif à répartir est passé, d'une année à l'autre de 58 à 63, et qu'on a dû passer de deux classes de 29 à trois classes de 21. On se demandera si les élèves de cette école réussissent mieux la seconde année que la première, évidemment en continuant à corriger des variables individuelles pertinentes. Cette méthode est dite de la " variation des populations naturelles ".

Aussi bien la méthode corrélacionnelle que la méthode expérimentale sont, on le comprend, coûteuses à mettre en œuvre, ce qui explique qu'aucune ne puisse être sans défaut, et aussi que leurs résultats puissent diverger. D'où l'intérêt des *synthèses des recherches*, qui recensent le nombre d'études concluant dans tel ou tel sens, et des *méta-analyses*. Cette technique consiste à

⁶ Il ne s'agit pas d'un cas d'école. L'auteur de ces lignes a étudié l'allemand en sixième pour cette raison.

compenser la nature forcément imparfaite des recherches en agrégeant les résultats de plusieurs d'entre elles. Il s'agit de retenir un certain nombre d'études, sur des critères de rigueur et de similitude, puis de remonter à leurs données elles mêmes, et de calculer ainsi un effet moyen de la taille des classes selon l'ensemble des études sélectionnées.

Les recherches sur les acquisitions académiques des élèves

Les comparaisons internationales

La plus récente évaluation internationale du niveau des élèves en mathématiques (TIMSS, 1995) montre que, sur 17 pays, il existe une relation négative, significative seulement au seuil de 10%, entre le taux d'encadrement en primaire et le niveau en mathématiques en 4^{ème} : ceux dont les élèves réussissent le mieux ont moins de professeurs par élèves. Lorsqu'on enlève la Corée, qui réussit très bien avec de très grandes classes, cette relation persiste, mais devient non significative. Sur les six évaluations internationales précédentes en mathématiques, aucune relation significative n'est trouvée entre le taux d'encadrement et le niveau en mathématiques, la durée d'études des parents étant tenue sous contrôle (Hanushek, 1998). L'auteur lui-même avertit qu'il y a par ailleurs trop de différences entre ces pays pour que ces résultats aient valeur de preuve, mais, à tout le moins, ils montrent que l'on peut aussi apprendre bien dans de grandes classes et que la taille des classes n'est qu'un des nombreux déterminants de la réussite scolaire.

Pong et Pallas (1999) ont aussi analysé les données de TIMSS *au niveau de la quatrième* dans 10 pays dont la France, dans cette perspective. Ils ne comparent pas, comme le faisait Hanushek, les scores nationaux et les tailles moyennes des classes dans chaque pays. En utilisant une méthode multiniveau, ils s'interrogent sur la forme de la relation taille/performance dans chaque pays. La taille des classes dans ces pays va de 16 (Islande) à 50 (Corée). Leur analyse présente des limites compte tenu des données recueillies, notamment parce qu'ils ne peuvent neutraliser que de façon très imparfaite la relation entre le niveau des élèves et la taille de la classe qu'ils fréquentent : ils utilisent à cette fin le niveau social moyen des élèves de la classe. Pour trois pays, ils trouvent un lien positif entre taille des classes et niveau des élèves, pour six pays, dont la France, ils trouvent une absence de lien, et pour un seul (les USA) ils trouvent un lien négatif. Toutefois, ils observent que, compte tenu de la taille de l'effet, réduire la taille des classes aux Etats Unis de 24 à 18 ne procurerait qu'un gain égal au septième de l'écart actuel entre ce pays et Singapour, le pays qui a les plus fortes performances.

Les recherches étrangères

C'est aux USA, comme souvent s'agissant de la recherche en éducation, que nous trouverons le débat le plus approfondi sur l'effet de la RTC. Dans les années 60, la cause semblait entendue, non seulement à cause de résultats non concluants obtenus précédemment (certains faiblement positifs, d'autres faiblement négatifs), mais aussi parce que plusieurs études à large échelle et à méthodologie rigoureuse venaient de conclure contre la baisse de la taille des classes, le rapport Coleman aux USA, mais aussi l'étude de Marklund en Suède et une réanalyse des données recueillies pour le rapport Plowden en Angleterre.

Puis, deux chercheurs (Glass et Smith, 1979) ont analysé systématiquement 77 études issues de 70 ans de recherche dans 12 pays et ont observé que :

- Sous la diversité des résultats, il y avait bien une unité : “ Sur 725 comparaisons, 435, soit 60% concluait en faveur des plus petites classes ”. Cependant, l’effet moyen était très faible.
- Que, mesuré sur les 14 études qu’ils ont estimées “ bien contrôlées ”, l’effet de la réduction n’était pas linéaire : quasi nul pour une diminution de 40 à 20 élèves, moyen de 20 à 10, fort en dessous de 10.
- Cependant, elles reconnaissaient que le coût était élevé : “ Obtenir un accroissement de 10% des performances des élèves demanderait que l’on diminue la taille des classes d’un tiers à la moitié... et que l’on multiplie donc le coût des écoles dans la même proportion⁷ ”.

Commença alors un des plus vifs débats opposant des chercheurs aux USA à propos d’éducation. Il a mobilisé des chercheurs en éducation, mais aussi des économistes, dans le cadre de la question plus générale du lien entre les ressources dépensées pour l’éducation et son efficacité. On a reproché à Glass et Smith d’avoir conservé des études de qualité douteuse et le fait que leurs résultats étaient fortement influencés par ceux d’une seule étude, extrêmement positive pour les petites classes.

Plusieurs synthèses, qui ont sélectionné les recherches de façon plus rigoureuse, ont conclu soit à l’absence d’effet (notamment Tomlinson (1986)) soit à un effet plus faible que celui que Glass et Smith avaient avancé, soit à un effet réservé aux premières années de l’enseignement. .

Cependant, le balancier est reparti dans un sens favorable à la baisse des classes sous l’influence de deux études, rigoureuses et à large échelle.

When money matters (Wenglenski, 1997) est une étude sur l’effet des ressources sur l’efficacité de l’enseignement, menée sur 203 districts au niveau du CM1 et sur 182 districts en 4eme en utilisant trois bases de données nationales. La recherche conclut à l’inefficacité des dépenses en capital (matériel, bâtiments) et des dépenses pour l’administration de l’école. En revanche, elle trouve un lien entre la baisse du nombre d’élèves par enseignant et l’efficacité de l’enseignement, particulièrement fort pour les élèves des minorités. Au CM1, la RTC agit directement et fortement. En 4^{ème}, l’effet est moins fort et indirect, il transite par le climat social de l’école (moins de violence,...).

Plus décisive encore a été l’influence du projet *STAR* (Student/teacher achievement ratio), une étude expérimentale, menée à une échelle rarement observée pour ce type d’étude : 11 000 élèves, répartis pendant 5 ans (1985-1990) pour un tiers dans des classes de 13 à 17 élèves, pour un tiers dans des classes de 22 à 26 élèves, et pour un tiers dans des classes de 22 à 26, mais dans lesquelles un aide était adjoint à l’instituteur, depuis le Kindergarten⁸ (notre Grande Section) jusqu’en fin de CE2 (Finn et Achilles, EEPA, 1999). Il est apparu que :

- La présence de l’aide n’avait aucune influence sur l’efficacité, mais que...

⁷ Ce qui n’est pas tout à fait vrai, dans la mesure où les dépenses en personnel enseignant, et même celles qu’elle induisent directement, ne sont pas tout à fait les seules du système éducatif.

⁸ Le Kindergarten est fréquenté par les enfants de cinq ans. Il peut donc être assimilé à notre Grande Section de maternelle, à ceci près que, dans une partie des écoles, les apprentissages y sont davantage “ scolaires ”.

- ... les élèves des petites classes réussissaient mieux que ceux des grandes classes. Les élèves des petites classes surpassent ceux des grandes de 0.15 à 0.26 écart-type selon la discipline ou le niveau considéré. Réanalysant ces données, Krueger (1997) observe que l'écart entre élèves des petites et grandes classes représente 64% de l'écart de performances entre blancs et noirs en Grande Section et 82 % au CE2.
- Pour les élèves des minorités, l'écart est très substantiel. En grande section, il est de même ampleur que pour les blancs, mais ensuite il se situe entre 0.3 et 0.4 écart type contre 0.15 à 0.2 pour les blancs. Krueger (1997) observe que l'effet des petites classes est particulièrement bénéfique aussi pour les élèves éligibles aux repas gratuits.
- Ce gain s'observe dès la première année, se maintient ensuite, y compris, ce qui est essentiel, cinq ans après que les élèves des petites classes aient rejoint, à partir du " CM1 " des classes de même taille que les autres. Par exemple, en maths, les élèves autrefois en petites classes du CM1 jusqu'à la fin de la cinquième, surpassent, de façon stable, les élèves autrefois en grandes classes, d'environ 0.15 écart-type. En " 5ème " ceux qui étaient autrefois dans de petites classes surpassent aussi les autres dans des matières qui n'étaient pas enseignées dans les années de l'expérience (social studies, par exemple), ce qui fait penser que l'effet peut, pour partie, être dû à des habitudes " d'engagement dans la tâche " prises dans les petites classes, et pas seulement au plus grand nombre de connaissances acquises ces années là (Nye et al., in EEPA, 1999).

Les résultats de STAR ont été abondamment *discutés*.

Plusieurs auteurs ont souligné le caractère " zéro aveugle " de l'étude : enseignants et élèves étaient conscients de leur situation et de l'enjeu, et de plus, les groupes étaient en contact, puisque les petites et les grandes classes appartenaient aux mêmes écoles (ce qui, cependant, avait l'avantage de neutraliser l'effet-école). Par ailleurs, élèves et enseignants ont été tirés au hasard dans les écoles participantes, mais celles-ci ont accepté de participer à l'expérience, ce qui a pu créer un biais de sélection au niveau des écoles.

Hanushek(1998) observe que l'essentiel des effets bénéfiques a été obtenu dès la grande section et conclut que STAR plaide pour une réduction de la taille des classes seulement " en grande section et peut être au CP ", mais pas après.

Golstein et Blatchford (1998) ont mis en évidence les problème de l'attrition et de la réallocation de certains élèves d'un groupe à l'autre. Ceux qui sont sortis de la cohorte étudiée (attrition) avaient un score inférieur aux autres. Or, ils ont été plus nombreux dans ce cas dans les grandes que dans les petites classes. Par ailleurs, en lecture, le niveau de ceux qui ont été réalloués des petites aux grandes classes est un peu plus faible que le niveau de ceux qui ont été réalloués des grandes aux petites classes. Ces deux phénomènes sont susceptible de biaiser les résultats en faveur des petites classes.

Cependant, les *réanalyses* qui ont été faites des données de STAR en utilisant des techniques plus sophistiquées ont plutôt confirmé l'effet positif des petites classes.

Krueger (1997) a utilisé des techniques qui lui permettent de neutraliser les effets de l'attrition et de la réallocation en estimant des valeurs pour les données manquantes. Il neutralise aussi le biais de sélection des écoles. Il confirme, on l'a vu ci dessus, les résultats initiaux, en rejoignant cependant Hanushek sur un point, mais avec une différence importante : les effets positifs sont effectivement obtenus au bout de seulement un an dans une petite classe, même si cette classe n'est pas la grande section⁹. Selon lui, cela signifie que " fréquenter une petite classe dans les niveaux initiaux de l'école produit un effet de socialisation à l'école (school-socialization effect) qui augmente la performance d'un niveau uniforme, sans affecter grandement la trajectoire (de ces performances) ". Autrement dit, l'effet serait obtenu la première année, se maintiendrait ensuite sans augmenter, ce qui aurait été le cas si l'effet de la petite classe avait été d'augmenter la capacité à apprendre de l'élève¹⁰.

Nye et al. (2000) ont réanalysé les données en utilisant un modèle multiniveau, et en tenant sous contrôle un grand nombre de variables décrivant les élèves, leurs enseignants et leurs écoles. Ils confirment les résultats initiaux, et, à la différence de Krueger, mettent en évidence un léger effet cumulatif : au CE2, par exemple en maths, l'écart avec un élève identique qui a été tout le temps dans des grandes classes est de 0.15 écart-type ; pour les élèves qui ont été pendant un an dans une petite classe, 0.19 ; pour les élèves qui les ont fréquentées pendant deux ans ou plus, 0.26 ; pour ceux qui les ont fréquentées pendant trois ans ou plus, 0,30, et 0.35 pour ceux qui les ont fréquentées pendant les quatre années de l'étude¹¹.

Sous contrôle du niveau en fin de grande section, Goldstein et Blatchford (1998) trouvent un effet positif de la RTC au CP (effet que Hanushek n'acceptait qu'avec réticence), au moins pour les maths. Ils mesurent un effet de 0.18 écart-type pour les blancs et de 0.35 pour les noirs. En lecture l'effet est quasi nul pour les élèves blancs (0.04) et plus substantiel pour les noirs (0.21).

Sous contrôle du niveau en fin d'année précédente, Goldstein et Yang (1999) trouvent pour la lecture un effet moyen quasi nul pour le CE1 et le CE2 (ils n'ont pas travaillé sur les scores de maths).

Surtout, les résultats de STAR ont incité à entreprendre de *nouvelles études*, corrélationnelles, dont une hors des Etats Unis. Souvent elles utilisent les grandes bases de données recueillies par le ministère ou par des Etats, de sorte qu'elles approchent la taille des classes parfois par des voies moins directes que le projet STAR.

Akerhielm (1995), sur 24000 élèves de 4ème, trouve une liaison positive de la taille de la classe avec le niveau des élèves en maths, anglais, histoire et sciences. Cet effet disparaît lorsque l'auteur tient sous contrôle le niveau initial des élèves, pour faire place à un effet négatif, significatif pour les sciences et l'histoire, mais non significatif pour les maths et l'anglais. Elle insiste sur le fait que, même en histoire et en sciences, cet effet est tel qu'une baisse massive de la

⁹ Krueger a pu obtenir ce résultat parce que, de même que certains élèves ont quitté la cohorte en cours de route, d'autres y sont entrés.

¹⁰ Pour une raison dont je dois avouer qu'elle me reste hermétique, vu la complexité de l'analyse, Krueger soutient aussi que ses calculs lui permettent d'éliminer l'hypothèse d'un " effet Hawthorne ".

¹¹ Evidemment, la population qui les a fréquentées pendant deux ans et plus contient celle qui les a fréquentées pendant trois ans et plus.

taille (de 25 à 15) n'accroît que de 0.8 point, sur une échelle de 25, le score de sciences et de 0.7 point, sur une échelle de 30, le score d'histoire.

Figlio (1999) étudie une cohorte de 5800 élèves sur trois ans, de la quatrième à la seconde. Il compare l'effet sur les progrès des élèves en maths, sciences, lecture, et "social studies" de quatre types de ressources : le taux d'encadrement, le salaire des enseignants en début de carrière¹², le nombre d'heures d'enseignement reçues par les élèves, la proportion d'enseignants qualifiés¹³. Il observe que la qualification n'a pas d'effet, que les trois autres facteurs en ont un léger. Un accroissement de 10% du taux d'encadrement conduit, sous contrôle du niveau en début de 4^{ème}, ainsi que de plusieurs caractéristiques des élèves et de leur environnement, à un accroissement de 1,4% du score de sciences en fin de seconde, soit un résultat faible pour un coût important. Un accroissement de 10% du nombre d'heures enseignées conduit, lui, à une amélioration de 5% du score de fin de seconde, et s'avère donc une démarche plus efficiente.

Hoxby (2000) propose, on l'a vu, une approche par les "variations naturelles", en fait les variations dues au hasard, de la taille des classes, dans laquelle elle ne tient pas compte des classes hors d'une fourchette 15-30. Elle étudie, sur 11 années (1986-97) si, dans les districts du Connecticut où le hasard conduit à une baisse de la taille des classes, les élèves progressent davantage que dans les autres¹⁴, ceci au CM1, en sixième et en quatrième. Alors que sa méthode lui permettrait de détecter des effets aussi faibles que 0.03 écart-type, elle ne trouve aucun effet significatif à 5% de la RTC, et ceux qui sont significatifs à 10% sont en faveur des grandes classes. A la différence des autres recherches, celle ci ne conclut pas à l'existence d'effets plus forts pour les élèves pauvres ou appartenant à des minorités.

L'ensemble des recherches américaines converge donc vers ceci que, dans le cadre de ce système éducatif, il existe bien, dans les petites classes de l'enseignement primaire, au moins en Grande section ou au CP, un effet positif d'une réduction forte de la taille des classes, un effet vraiment substantiel pour les élèves des minorités et pour les enfants de famille défavorisées, un effet durable même après que les élèves ont rejoint de grandes classes. Nye et al. (2000) soulignent la proximité de ces résultats avec ceux de la méta-analyse de Glass et Smith (1979) : alors que ces derniers prédisaient, pour une baisse de 25 à 15, un effet standard de .22, leur propre réanalyse des données de STAR conduit, pour un écart de taille proche, à des effets compris entre .15 et .30 selon les disciplines ou les niveaux¹⁵. La question de savoir si cet effet est obtenu en un an ou si la fréquentation de petites classes pendant plusieurs années l'amplifie reste controversée (Krueger et Hanushek vs Nye et al.).

En revanche, les recherches dans le second degré sont contradictoires : deux concluent à un effet positif sensible des petites classes (Wenglenski et l'étude expérimentale du district de San Juan (cf. la liste des études en annexe), deux autres, on vient de le voir, concluent à des effets soit nuls,

¹² Ce salaire varie aux Etats Unis d'un district à l'autre.

¹³ "with masters degree".

¹⁴ Cette approche est possible parce que la plupart des districts de cet état sont très petits, ce qui fait qu'aucune politique ne peut lisser les effets des variations naturelles en modifiant les zones de recrutement des écoles.

¹⁵ On peut noter que 0,2 écart-type est une grandeur qui apparaît dans d'autres réformes éducatives de grande ampleur. Dans le Kentucky, a été mise en place une réforme importante de l'enseignement fondée sur la mesure systématique des résultats des élèves, l'autonomie des établissements, le soutien scolaire et la formation continue des enseignants. Les résultats, réanalysés minutieusement en comparant la réussite des élèves aux épreuves de l'État et à des épreuves nationales, montrent un progrès, en maths pour des élèves de quatrième, de 0,26 écart-type (Koretz, D.M. et Barron, S.L., 1998).

soit bien trop ténus pour être, pour employer une expression américaine, “ significatifs au niveau éducatif ”. Cependant, dans la dernière en date de ses synthèses des recherches, Hanushek (EEPA, 1999) ne trouve pas de différences entre les effets dans le premier et dans le second degrés.

On l’a vu, le gouvernement fédéral en a tiré la conclusion qu’il valait la peine de financer des programmes onéreux de baisse de la taille des classes dans les premières années de l’enseignement primaire. Certains états l’ont précédé dans cette politique, d’autres – une vingtaine – l’ont suivi, ajoutant leurs propres subventions aux subventions fédérales. Un vif débat existe quant aux *résultats de ces politiques*, entre ceux qui pointent certaines améliorations obtenues, et ceux qui les trouvent trop faibles pour le coût consenti. Ces derniers avancent qu’accroître l’efficacité des enseignants serait une politique plus efficace. Il est exact que “ l’effet enseignant ” est bien plus fort que l’effet “ taille des classes ”. Par exemple, *selon une étude française, au CP, en français, l’écart moyen d’efficacité entre enseignants est 2,5 fois plus fort que l’effet d’une réduction de la taille de la classe de 27 à 17 élèves* (Mingat, 1991). Mais cet argument suppose tout de même que l’on connaisse des moyens infaillibles pour accroître l’efficacité des enseignants les moins efficaces, ce qui n’est pas le cas.

En Californie, où un programme autonome avait été lancé avant la politique fédérale, les autorités ont mandaté un consortium de chercheurs pour évaluer ce programme. L’évaluation souligne que l’effet du programme de RTC a été affaibli par la moindre qualité des enseignants qui ont été recrutés pour satisfaire les besoins en postes que la réduction avait créés, et que les enseignants les moins qualifiés ont été assignés dans les classes situées en zones populaires . Ils notent malgré cela un (très) léger effet positif au CE2 : la proportion d’élèves réussissant au dessus de la médiane de l’État serait, selon les disciplines, de 1.4 à 3.6 points plus forte dans les zones où la taille a été réduite que dans les autres (CSRRC, 2000). Toutefois, comme ils reconnaissent aussi un biais dans l’exécution du programme qui a fait que les classes les moins favorisées en ont bénéficié moins que les autres, il n’est pas impossible que ces écarts, bruts, expriment en fait un effet nul.

Au Wisconsin, les effets du projet SAGE semblent plus positifs. Un écart statistiquement significatif a été observé avec des élèves semblables hors programme, dans plusieurs disciplines dont les maths et la lecture, en CP, lors d’une évaluation rigoureuse du programme (Molnar et al., in EEPA, 1999). Cet écart se maintient, mais ne s’accroît pas, en fin de CE1, ce qui est conforme aux résultats de Krueger sur STAR. Par ailleurs, l’écart de performances blancs/noirs s’est réduit dans les zones SAGE alors qu’il a augmenté dans les zones témoins (USDE, 2000, Pritchard, 1999, Molnar et al., 1999). A New York, un panel indépendant a aussi évalué la première année du programme de réduction, trouvant qu’il était trop tôt pour observer des résultats sur les tests mais que “ les élèves semblaient apprendre plus vite dans les petites classes que dans des classes plus grandes ” (USDE, 2000).

En Caroline du Nord, les élèves dont la classe a été réduite dépassent les élèves comparables des classes non réduites en lecture et en maths, pour toutes les années du CP au CE2 (Pritchard, 1999).

Nous avons évoqué ci dessus une étude hors des Etats Unis. On pourra regretter qu’il s’agisse d’une *étude anglaise*, donc sur un pays dont le système éducatif est sans doute plus proche du

système américain que du notre. *Blatchford et Goldstein (2000)* sont en train de terminer l'analyse d'une étude sur les trois premières années de l'enseignement primaire (de la "Reception year"¹⁶ au grade 2). Deux cohortes, de 8000 élèves chacune, ont été suivies sur ces trois niveaux, l'une commençant en 1996, l'autre en 1997. Il s'agit d'une analyse corrélationnelle. Les auteurs restreignent leurs analyses aux classes de 10 à 35 élèves. Seuls les résultats de la "Reception year" sont actuellement disponibles. Ils témoignent d'un effet négatif de la taille des classes, en lecture comme en mathématiques, aussi bien lorsque le niveau initial des élèves est tenu sous contrôle que lorsqu'il ne l'est pas. Dans le cas de la lecture, cet effet est plus fort pour les élèves éligibles aux repas gratuits et pour les élèves académiquement faible. Les résultats observés dans le Tennessee semblent donc résister à la fois à un changement de système éducatif et à un changement d'approche, corrélationnelle au lieu d'expérimentale. L'étude de Goldstein et Blatchford montre par ailleurs non seulement que l'effet de la taille des classes n'est pas linéaire – ce qu'on sait depuis Glass et Smith (1979) – mais que cette non linéarité diffère selon les disciplines : en lecture, l'effet est à peu près linéaire entre 15 et 30 élèves, tandis qu'en maths, l'effet est net entre 15 et 25, quasi nul ensuite.

Outre ces études et ces évaluations des politiques récentes de RTC, de nouvelles *méta-analyses* ont été entreprises depuis le projet STAR, qui l'incluent d'ailleurs parmi les études recensées.

Slavin (1990, cité par Goldstein et Yang, 2000) a réanalysé neuf études expérimentales. Il conclut cette fois à un effet moyen de 0.17 écart-type pour un passage de 25-30 à 15-16 élèves par classe, soit un effet réel mais pour une quasi division par deux de la taille des classes.

La méta-analyse de Goldstein et Yang (2000) utilise l'approche multiniveau (les différentes études sont un des niveaux de regroupement les données) ce qui leur permet une précision accrue. Ils analysent, comme Slavin, neuf études (dont sept déjà réanalysées par lui, mais pas au moyen de l'approche multiniveau, et dont STAR), toutes expérimentales, toutes comportant une mesure des performances initiales, et couvrant une plage de 15 à plus de 30 élèves par classe. Ils présentent d'abord les résultats sans STAR. L'inclusion de STAR ne change pas les résultats, ceci bien que l'analyse mette en évidence des différences entre les résultats des huit premières études. L'effet estimé de la RTC est un gain moyen de 0.02 écart-type par élève en moins, d'un ordre de grandeur comparable aux 0.2 écart-type que nous avons plusieurs fois observé dans des réductions de 10 élèves. Cette méta-analyse met aussi en évidence un effet positif de la taille des classes au delà de 30 élèves, mais les auteurs signalent que ce résultat est dû à deux études seulement, probablement celles de Mazareas (1981) sur le CP et de Shapson (1980) sur le CM1 (voir ces études en annexe).

Ces deux méta-analyses d'études expérimentales donnent une image plus positive de la RTC que la synthèse des recherches, en général corrélationnelles, proposée par Hanushek in EEPA, 1999. Sur 90 recherches conduites avant 1994, proposant 277 mesures différentes du lien entre taille des classes et performance des élèves, 28 % seulement donnent un résultat significatif, dont 15 concluent à un lien négatif et 13 à un lien positif, ce qui lui permet de conclure qu'il y a sans doute des conditions sous lesquelles la RTC produit des résultats positifs mais que " nous ignorons lesquelles ". Il est possible cependant que les effets de la RTC soient assez faibles, assez

¹⁶ Les élèves commencent en général la "Reception year" à 4 ans. Elle offre un curriculum davantage orienté vers l'apprentissage que les "Nursery schools".

limités aux réductions drastiques, pour que les études corrélationnelles ne puissent les appréhender rigoureusement, et donnent de ce fait cette distribution aléatoire.

Les recherches françaises

Il n'est pas sûr que les résultats obtenus à l'étranger soient transférables sans précaution en France. On pourrait penser, par exemple, que l'enseignement américain est moins frontal que l'enseignement français et pâtirait pour cette raison davantage de la taille des classes. Il semble que, en maths dans le second degré, ce ne soit pas le cas (la proportion d'enseignement qui se passe devant la classe entière n'est pas moins forte aux Etats unis qu'en France (Pong et Pallas, 1999), mais d'autres facteurs peuvent intervenir.

En France, la première étude sur le sujet fut réalisée par un IEN, M. Auvinet. Il observe que les résultats au certificat d'études sont meilleurs dans les classes supérieures à 30 qu'à moins de 30. Un peu plus tard, A. Legrand, puis une étude du SIGES en 1978, ont mis en évidence que le pourcentage d'élèves en retard baisse quand la taille des classes augmente, la même relation ayant d'ailleurs été observée en sens inverse à Genève par W. Hutmacher beaucoup plus récemment¹⁷.

Dans la période récente, aucune étude n'a été consacrée spécifiquement à ce problème. Les mesures de l'effet de la taille des classes interviennent dans des études générales sur les déterminants de la progression des élèves (Mingat, 1991), dans des études sur l'école rurale (Leroy-Audoine et Mingat, 1995) ou sur des méthodes pédagogiques (Suchaut et Le Bastard, 2000, par exemple). Il n'est pas sûr que ce soit un inconvénient, puisque cela permet parfois de comparer l'effet de la taille des classes à celui d'autres facteurs. Par ailleurs, on l'a dit, ce sont toutes des recherches corrélationnelles, ce qui a l'avantage et l'inconvénient du réalisme. Ces études mesurent l'effet de la taille des classes sous contrôle du niveau initial des élèves et de leurs caractéristiques sociales, mais pas sous celui de la qualité des enseignants, des établissements ou, bien sûr, des effets d'éventuelles stratégies volontaires des parents les plus soucieux de la réussite de leurs enfants, qui s'arrangeraient pour les mettre dans de petites classes. Une autre limite est qu'elles ne distinguent pas l'ampleur des effets en fonction des caractéristiques des élèves.

Suchaut (1996) étudie un échantillon de 900 élèves de grande section de Maternelle, scolarisés dans des classes de 17 à 33 élèves, en mesurant les compétences des élèves, en début comme en fin d'année, sur une série d'exercices portant sur les discriminations phonologiques et visuelles, la mémoire, l'organisation spatiale, etc. Ces exercices sont utilisés par les psychologues pour tester la maîtrise des pré-requis pour l'apprentissage au CP. Il ne trouve pas d'effet significatif de la taille de la classe.

Mingat (1991) étudie un échantillon de classes de CP de 17 à 27 élèves, en se limitant à des classes à cours simples. Il trouve que, toutes choses égales par ailleurs, les acquisitions en français et en mathématiques au CP sont un peu meilleures dans les petites classes (- 0,01 écart-

¹⁷ Ce résultat, obtenu dans les trois cas de façon longitudinale, ne peut être biaisé du fait que les meilleurs élèves seraient dans de plus grandes classes. Il contredit l'argument parfois avancé, en particulier aux Etats Unis, que le coût d'une baisse de la taille des classes serait compensé en partie par le coût des redoublements qu'elle éviterait. Les trois études françaises sont présentées dans de Peretti, A., 1987.

type par élève supplémentaire). En maths, l'effet est non significatif ; en français, il est de l'ordre de grandeur de celui que nous avons vu dans les études expérimentales américaines (+ 0,18 écart type pour une réduction de 27 à 17). Cette étude est celle dont les résultats sont le plus favorable à la RTC. Toujours au CP, sur des échantillons un peu moins importants (900 élèves contre 2200 pour Mingat), Suchaut (1996) trouve un effet positif de la taille des classes sur les apprentissages en maths (significatif à 5% seulement), et ne trouve pas d'effet significatif en français tandis que Suchaut et Le Bastard (2000) trouvent en français un effet positif de la taille des classes, significatif à 1%, lorsqu'on passe de 18 à 23, puis un effet négatif au delà de 23. Suchaut (1998) porte aussi sur l'apprentissage du français au CP, sur un échantillon 750 élèves répartis dans des classes de 15 à 26 élèves. Cette étude porte sur des écoles scolarisant une population en majorité défavorisée ou d'origine étrangère dans le département de la Seine-Saint-Denis. L'effet de la taille des classes y apparaît non significatif.

Leroy-Audoin et Mingat, (1995) ont réalisé en Saône et Loire, sur un millier d'élèves de zones rurales, une étude sur la progression en français et en mathématiques pendant le CE2. Ils ne distinguent pas ces deux disciplines dans la présentation de leurs résultats. Bien que limités en toute rigueur aux zones rurales, les résultats valent par le fait que les auteurs ont étudié les interactions entre l'effet de la RTC et, d'une part la taille de départ, d'autre part la structure du cours. Globalement, ils trouvent que l'on apprend significativement mieux dans les classes de 18-22 élèves que dans les classes inférieures à 18, mais un peu mieux dans les classes de 18-22 que dans les classes supérieures à 23 élèves. Dans ce dernier cas de figure, l'effet négatif de la taille des classes est de -0.01 écart-type par élève supplémentaire, très voisin de celui trouvé par Mingat au CP. Comme Suchaut et le Bastard, donc, ils trouvent un optimum vers 23 élèves par classe. Cependant, l'effet positif de la taille semble limité aux classes à cours multiples. Dans leur cas, il est de l'ordre de 0,2 écart-type pour 10 élèves supplémentaires. Lorsque, comme dans l'étude de Mingat sur le CP, ils se limitent aux classes à cours simples, ils ne trouvent pas d'effet de la taille des classes.

Bressoux (1993) étudie les progrès en lecture après le CE1 ; il ne trouve pas d'effet de la taille des classes sur les progressions des élèves, sauf un très léger effet – négatif, cette fois – pour les élèves des classes à cours multiples (significatif seulement à 10%).

En bref, et si l'on excepte les classes à cours multiples, on trouve une étude concluant en faveur de classes plus petites (le français au CP, Mingat (1991), deux études concluant en faveur de classes plus grandes, jusque vers 23 élèves (le français au CP, Suchaut et Le Bastard, 2000 ; les maths au CP, Suchaut, 1996), et cinq études concluant à l'absence d'effet significatif de la taille des classes. Ces résultats sont donc moins positifs que ceux des études expérimentales qui ont relancé aux Etats Unis la politique de RTC¹⁸.

Cela vient-il de la différence entre les deux systèmes éducatifs ou de la différence entre les méthodes de recherche ? Nous penchons plutôt pour la seconde solution, compte tenu de la similitude de ces résultats avec ceux des synthèses de Hanushek portant surtout sur des études corrélationnelles, savoir des effets tellement ténus que la distribution des résultats entre études

¹⁸ Rappelons cependant que l'efficacité des enseignants n'est pas ou mal contrôlée dans ces études, alors qu'elle l'est dans les études expérimentales, que les vraiment petites classes (de taille inférieure à 16 élèves) sont peu fréquentes dans leurs échantillons, et que le fait qu'elles ne recourent pas à des modèles multiniveau surestime la significativité des résultats.

concluant dans un sens ou dans l'autre est proche d'une distribution aléatoire. Mais ceci comporte un enseignement : l'effet positif de la RTC semble se manifester surtout en cas de baisse brusque, sans doute plus favorable à un changement des pratiques pédagogiques des enseignants, et dans une situation où les enseignants savent que les compétences de leurs élèves vont être mesurées.

On dispose aussi de quelques études sur l'effet de la taille des classes au début du collège. Grisay (1993) trouve que la petite taille du collège et des classes est un facteur explicatif de la performance des établissements en français et en maths. Outre que cette variable apparaît ici dans l'explication des performances du collège, et non de chaque élève, son mode de construction ne permet pas de distinguer entre taille des classes et taille du collège, et il est probable que cette dernière est plus "active" en l'affaire que la seconde. C'est ce qu'inclinent à penser trois études sur les progressions des élèves en sixième.

Schmitt-Rolland et Thauvel-Richard (1997) n'observent pas d'effet de la taille des classes sur la progression des élèves en français, dans une étude portant sur un large échantillon d'élèves scolarisés dans des classes de 13 à 31 élèves et dans le cadre d'une analyse contrôlant l'effet d'un grand nombre d'autres variables caractérisant les classes (leur niveau moyen, les caractéristiques sociales et scolaires de leur public). Verdon et Thauvel-Richard (1997), dans une analyse semblable, n'en observent pas non plus sur la progression des élèves en mathématiques. Génelot (1997) n'en trouve pas non plus sur la progression des élèves en anglais.

Il ne semble donc pas que la taille des classes, dans la limite des tailles actuellement observables, ait un effet sur les élèves au début du collège. On ne dispose pas, à ma connaissance, d'étude française sur les effets de la taille des classes en quatrième, un niveau souvent étudié aux Etats Unis, et où, on l'a vu, un effet positif était observé dans certaines études à travers une diminution des attitudes négatives des élèves (Wenglenski).

Enseigne-t-on différemment dans les petites classes ?

Les arguments théoriques en faveur des petites classes énoncent seulement que l'on *peut* enseigner différemment dans les plus petites classes, être plus attentif aux élèves, par exemple. La question de savoir si cela se produit réellement est différente. Certains économistes (Hoxby, 2000) ne sont pas mécontents de pouvoir avancer que, si la baisse de la taille a des effets si faibles comparés à ceux que les arguments théoriques laissent espérer, c'est que rien n'incite les enseignants à se saisir de cette occasion d'enseigner de façon plus efficace, puisqu'ils n'en tirent aucun bénéfice.

Et, de fait, certaines études déjà anciennes observent que les enseignants ne changent pas leur façon d'enseigner quand la classe est plus petite (notamment Yeany, 1976, voir en annexe). Dans leur comparaison internationale, Pong et Pallas (1999) disent "trouver peu de preuves que les pratiques des enseignants (d'après leurs réponses à ce sujet au questionnaire enseignant de TIMSS, note D.M.), différent beaucoup selon la taille de leur classe". Par exemple, la proportion de temps consacré à enseigner la classe entière ne change pas aux Etats Unis selon la taille de la

classe¹⁹. De même, on ne couvre pas une partie plus importante du programme dans les petites classes.

Inversement, certaines évaluations récentes des programmes américains de réduction mettent en évidence une réduction des problèmes de discipline, un centrage sur la prévention plutôt que sur la remédiation, une meilleure participation des élèves, un “moral” des enseignants meilleur²⁰ (New York, USDE, 2000), une atmosphère plus positive dans les classes, une meilleure connaissance de chaque enfant par les enseignants, davantage de temps disponible pour l’instruction, une individualisation de l’instruction (SAGE, Molnar et al., in EEPA, 1999 ; USDE, 2000). Dans les écoles SAGE, le temps consacré à l’instruction est passé de 80 % à 86% du temps total (Pritchard, 2000). D’une évaluation de la politique néerlandaise de RTC, il ressort que le temps consacré au travail n’est pas plus grand dans les classes plus petites, mais que davantage d’attention y est portée aux élèves en difficulté, et que l’enseignement est plus adapté à chaque enfant (Annevelink et al., 2001).

Les questionnaires remplis par les enseignants pendant le projet STAR indiquent, de la part des élèves des petites classes, un comportement “davantage centré sur la tâche” et davantage d’engagement (EEPA, 1999).

Deux études ont été conduites sur cette question dans l’enseignement secondaire. Betts et Skolnick (EEPA, 1999) comparent les réponses à un questionnaire sur leurs pratiques, l’année n et l’année (n-1), d’enseignants de maths qui enseignent l’année n dans une classe plus petite qu’en (n-1). Les enseignants dont la classe s’est réduite consacrent effectivement plus de temps à l’instruction parce qu’ils rencontrent moins de problèmes de discipline, mais ne couvrent pas pour autant une plus grande partie du programme ; ils consacrent plus de temps à revenir sur des éléments appris, à vérifier les acquis des élèves et à une instruction individualisée. Cependant les auteurs soulignent que ces effets sont d’une ampleur très faible, même lorsque la diminution de la taille de la classe est substantielle. Ils sont selon eux trop ténus pour avoir une influence réelle sur l’efficacité de l’enseignement. Rice (EEPA, 1999) trouve que, dans les classes plus petites, les élèves travaillent davantage en petits groupes, les enseignants utilisent davantage des pratiques innovantes, sont engagés plus souvent dans des discussions avec toute la classe. Elle observe aussi que ce sont les enseignants qui consacrent le plus de temps à la préparation de leur cours qui changent le plus leur façon d’enseigner.

Une tendance semble se dégager, celle d’un enseignement plus individualisé. Or, elle pose problème. Crahay (2000) cite plusieurs études qui concluent à l’inefficacité des modes d’enseignement qui reposent beaucoup sur le travail individuel des élèves (pas sur le travail en petits groupes). Il semble qu’on en ait tiré les conséquences en Angleterre où le ministère promeut “l’enseignement interactif en classe entière”. Une évaluation récente de ce type d’enseignement conclut à son efficacité et confirme l’inefficacité des professeurs qui pratiquent “l’individual monitoring” (Muijs et Reynolds, 2000). Il apparaît donc que toutes les formes d’enseignement que favorise une petite classe ne sont pas également efficaces, et que certaines le sont particulièrement peu. Cela aussi peut expliquer la faiblesse des effets de la RTC.

¹⁹ Il ressort d’ailleurs de cette étude que les élèves n’apprennent pas plus mal dans les classes où beaucoup de temps est consacré à enseigner la classe entière.

²⁰ Les documents consultés ne permettent pas de savoir comment ces résultats ont été obtenus.

De l'avis général des chercheurs, il reste beaucoup à comprendre de la façon dont la RTC agit sur les conditions de l'enseignement. En particulier quant à la compréhension de la durabilité des effets : tiennent-ils à un meilleur apprentissage du métier d'élève, auquel cas il faut cibler la RTC sur les premiers niveaux de l'enseignement ; tiennent-ils à un changement dans les méthodes pédagogiques, auquel cas il importerait d'accompagner la RTC de formations sur les méthodes d'enseignement les plus efficaces dans les petites classes ; tiennent-ils simplement au fait que, dans les petites classes, la probabilité diminue qu'un élève gêne, par ses questions ou son comportement perturbant, l'apprentissage des autres, auquel cas il s'agirait d'un effet quasi mécanique ?

Cette dernière thèse est soutenue par un économiste (Lazear, 1999) : le temps du professeur est considéré comme un bien public affecté "d'effets de congestion", c'est à dire que le comportement de certains usagers peut gêner la consommation du bien par les autres. Sa force est d'expliquer à la fois que la RTC ait un effet, mais aussi que les effets de la RTC sont d'autant plus forts que l'on part déjà d'une petite taille (une réduction d'un élève diminue plus fortement la probabilité de perturbation si l'on passe de 16 à 15 élèves que de 30 à 29), qu'ils sont plus forts pour les élèves défavorisés (qui ont davantage besoin d'accès à l'enseignant), et aussi un autre phénomène qui n'a rien à voir avec la taille des classes, savoir que l'arrivée de plusieurs élèves nouveaux dans une classe gêne leur apprentissage ainsi que celui des élèves qu'ils ont rejoint. Sa faiblesse est de ne pas rendre compte de la durabilité de l'effet, sinon à travers le fait que l'élève disposera d'un bagage supérieur à l'issue de sa petite classe.

Les effets non académiques

Les études en ce domaine sont infiniment moins nombreuses, mais semblent indiquer qu'il ne faut pas considérer a priori que les effets de la RTC seront favorables pour le développement personnel et social des élèves, comme pour leur bien-être.

Glass et Smith ont publié en 1982 une synthèse des recherches sur les effets de la taille des classes sur diverses dimensions non académiques, surtout sur l'attitude des élèves envers l'école (créativité, motivation...). Une grande majorité des études recensées conclut à un effet positif des petites classes. Mais des études plus récentes, portant il est vrai sur des dimensions légèrement différentes, sont moins favorables à la RTC.

Meuret et Marivain (1997) ne trouvent pas d'effet significatif de la taille des classes et de celle du collège sur le bien-être des élèves.

Blatchford et Goldstein (2000) observent, en Angleterre, un effet négatif des petites classes sur le comportement des élèves. Les enseignants des petites classes rapportent plus souvent que les autres que les élèves sont agressifs entre eux, ou que certains sont rejetés par les autres. Cela confirmerait l'idée que les grandes classes évitent les effets de clique, et aussi les résultats obtenus au Danemark par Egelund (1979) sur des élèves plus âgés.

Un effet non académique mieux documenté est celui des effets sur les salaires des élèves une fois adultes. Certains auteurs, en effet, considèrent que les épreuves scolaires mesurent de façon partielle des effets de l'école sur les élèves, puisque l'école ne transmet pas seulement des connaissances. Ceux ci sont, selon eux, mieux mesurés par les revenus de l'élève devenu adulte, dans la mesure où le salaire représente la valeur que la société attribue à la contribution professionnelle de l'individu, laquelle représente une part importante de sa contribution globale. Ils se demandent donc si des élèves qui ont bénéficié de classes plus petites, ou d'un taux d'encadrement plus grand, ou plus généralement d'un enseignement plus coûteux, tirent un revenu plus important que les autres de leur scolarité. Autrement dit, sachant qu'un des résultats les plus sûrs de la recherche économique en éducation est que ceux qui ont fait une meilleure carrière scolaire ont des salaires plus élevés, que donc on peut affecter à une année supplémentaire d'éducation un certain "rendement", ils se demandent si ce rendement d'une année scolaire supplémentaire est plus fort pour les élèves qui ont bénéficié de plus petites classes.

Card et Krueger (1990) ont comparé les revenus des individus scolarisés selon le taux d'encadrement dans les états où ils avaient fait leurs études, à carrière scolaire identique. Ils ont trouvé un léger effet positif du taux d'encadrement, ainsi que de la longueur de l'année scolaire et du salaire moyen des enseignants. Une étude ultérieure (Altonji et Dunn, citée par Card et Krueger, 1996) établit que des dépenses d'éducation par élève plus élevées de 10 % se traduisent par une augmentation de 1.3% des revenus tirés d'une scolarité donnée. Card et Krueger (1996) citent cependant d'autres études qui ne constatent pas cet effet ou qui en minimisent l'ampleur (Heckman et al., 1996 ; Betts, 1995 trouve un accroissement des gains de seulement 0.4% pour une hausse de 10% du taux d'encadrement) et la discussion continue entre les économistes à ce sujet. Discussion à laquelle il semble possible d'ajouter qu'en la matière, il n'est

pas sûr que l'on puisse déduire de l'effet d'un écart de rendement entre deux Etats le fait que, si tous les Etats s'alignaient sur le plus coûteux, le rendement des études au niveau national en serait augmenté dans les proportions correspondantes.

La question du coût

Même les chercheurs les plus convaincus des effets positifs de la RTC mettent en garde contre le coût élevé de cette politique. Non par un souci *per se* de ne pas augmenter les dépenses de l'État, mais par souci de recommander les politiques les plus efficaces. Ils savent que les dépenses ne sont pas extensibles et que toute dépense consacrée à baisser la taille des classes ne pourra l'être à des politiques peut être plus efficaces. Au delà, donc, de la discussion sur les effets de la RTC, que nous venons de voir, il est nécessaire d'en engager une autre sur la comparaison du coût et des effets de la RTC par rapport à ceux d'autres politiques.

Or, la RTC coûte cher. Glass et Smith (1979), eux mêmes, on l'a vu, le font observer. Aux Etats-Unis les différences de taille des classes expliquent presque la moitié des différences de coût par élève entre les districts (Card et Krueger, 1996). De plus, la RTC coûte d'autant plus cher qu'elle se produit à partir d'une taille réduite : son coût est proportionnel à la réduction en pourcentage de la taille, pas à sa réduction en valeur absolue²¹ (Hoxby, 2000). Par ailleurs, ce coût peut varier avec les modalités de la politique : l'instauration de seuils maximaux rigides coûte particulièrement cher, puisqu'elle peut conduire à des classes largement en dessous du seuil.

Cependant, ce coût n'est pas forcément prohibitif si les baisses sont ciblées sur des populations particulières. A titre d'exemple, et d'une façon grossière, on peut estimer, sauf erreur de notre part, que passer de 21,3 à 18 élèves les (environ) 6400 classes de CP d'éducation prioritaire (ZEP+REP), coûterait, sur la base d'un coût salarial de 23 100 F²² par mois, environ 300 millions de francs par an, sans tenir compte des frais annexes (classes à construire, formation des nouveaux enseignants, formation continue associée à la baisse, etc.), ni des effets de seuil.

Toutefois, s'intéresser au coût ne se conçoit qu'en rapport avec l'efficacité. Rechercher des politiques "économiques" n'a pas plus de sens que de rechercher des politiques efficaces sans tenir compte de leur coût.

Or, du point de vue de l'efficience, la RTC se discute aussi. On l'a vu, Pong et Pallas (1999) ont calculé que les gains en maths que l'on pourrait espérer aux Etats Unis d'une RTC de 24 à 18 représenterait seulement le septième de l'écart entre ce pays et Singapour.

On a fait remarquer aussi que l'effet de la RTC était bien inférieur à celui d'autres politiques éducatives moins coûteuses. Par exemple, Fraser (1987) a analysé les recherches existant sur une vingtaine de variables agissant sur l'efficacité de l'enseignement. Il a calculé ensuite pour chaque variable l'effet moyen indiqué par l'ensemble des études qui lui étaient consacrées. La RTC est

²¹ Pour répartir 10 000 élèves dans des classes de 29 au lieu de 30, il faut 12 classes supplémentaires. Pour répartir ces 10 000 élèves dans des classes de 19 au lieu de 20, il faut 26 classes supplémentaires. Cela signifie aussi qu'un programme de réduction de 8 élèves, de 24 à 16 par exemple, coûte plus que 133% d'un programme de réduction de 6 élèves, de 24 à 18 (Brewer et al., in EEPA, 1999).

²² Il s'agit du coût, en 1998, d'un professeur des écoles de 15 ans d'ancienneté, incluant des charges sociales estimées à 53% du salaire brut, incluant aussi la prime ZEP (renseignements communiqués par le bureau DPD C6). Ce coût est donc une estimation haute du coût *immédiat* en personnel d'une politique de RTC, puisque les enseignants recrutés pour mettre en œuvre cette politique seraient des enseignants débutants, qui coûteraient moins cher. 21,3 est l'effectif moyen des classes élémentaires de l'éducation prioritaire (ZEP+REP) (État de l'école, 2000), chiffre qui est peut être supérieur à l'effectif des CP si l'on a dans ces zones déjà accordé une certaine priorité au CP à cet égard.

l'avant dernière quand on classe ces variables par ordre d'efficacité décroissante. La RTC apparaît dans cette analyse quatre fois moins efficace que le tutorat entre élèves, huit fois moins que le fait de donner du travail à la maison et de le corriger rapidement, sept fois moins que le " cooperative learning ", c'est-à-dire l'apprentissage au sein de petits groupes hétérogènes, quatre fois moins que l'apprentissage en groupes homogènes dans une classe hétérogène, etc. Toutefois, Fraser compare la RTC à des dispositifs expérimentés sur peu de temps, à petite échelle, dont il n'est pas sûr que l'effet soit permanent, ou se maintienne au niveau observé dans des conditions expérimentales. De plus, il n'est pas juste de comparer la RTC à des dispositifs comme la correction des copies, puisqu'il s'agit là d'un comportement directement lié à l'apprentissage ; il faudrait la comparer à un dispositif incitant les enseignants à améliorer le mode de correction de leur copies ou à évaluer plus fréquemment leurs élèves²³.

Si l'on compare la RTC à des politiques de même nature et de coût comparable, on peut avancer que certaines sont probablement plus efficaces, d'autres moins.

Parmi les politiques plus efficaces, on trouverait sans nul doute quelques actions visant à diminuer le nombre d'enseignants en grande difficulté, dont une sélection plus sévère en fin de seconde année d'IUFM. Macbeath et al. (2001) viennent de terminer une évaluation de la politique britannique de " Study support ", c'est à dire d'un ensemble de dispositifs visant à proposer des enseignements aux élèves les plus faibles en dehors des heures normales de cours (non seulement des activités de soutien après les cours, comme en France, mais aussi des activités pendant le week-end, avant les cours, des " écoles d'été ", etc.), cette évaluation est positive, et, bien qu'il s'agisse aussi d'une politique coûteuse, il est probable qu'elle est plus efficace que la RTC. Bien qu'elle aille à rebours des orientations pédagogiques actuelles, la Pédagogie de Maîtrise, il faut le rappeler, paraît plus efficace que la RTC pour un coût bien moindre. L'évaluation de Bloom a été revue à la baisse lorsque l'on a évalué des programmes implantés en vraie grandeur et sur la durée, mais Crahay (2000) conclut tout de même à un effet de l'ordre de 0.4 écart-type.

Pour d'autres politiques, des investigations supplémentaires seraient clairement nécessaires. Les adversaires américains de la RTC proposent à sa place une politique d'incitation financière des enseignants à l'efficacité académique, mais on manque d'études établissant l'efficacité réelle de ce type de politique, et les expériences dans d'autres services publics que l'éducation semblent peu concluantes (Richardson, 1999). Il s'agit là de politiques visant à accroître l'efficacité de l'enseignement. Mais une politique d'allongement de la scolarité peut aussi être envisagée, elle consiste aussi à dépenser plus pour améliorer le niveau final de qualification, avec des résultats qui, en France, ont été positifs du point de vue de l'efficacité comme de l'équité (Thélot et Vallet, 2000 ; Duru-Bellat et Kieffer, 1999).

Enfin, certaines politiques semblent bien être moins efficaces que la RTC. Parmi elles, on trouverait sûrement les redoublements, la seule politique en direction des élèves faibles financée à guichets ouverts (Crahay, 1996). Leroy Audoin et Mingat (1995) trouvent des effets très faibles pour la RTC, mais ils trouvent des effets encore plus faibles du niveau de formation générale des

²³ La fréquence de l'évaluation des élèves, un facteur d'efficacité selon de nombreuses recherches sur l'efficacité des établissements, l'est aussi selon Verdon et Thaurel (1997).

enseignants du primaire²⁴, et un effet négatif (non significatif) de leur degré de formation professionnelle. Des résultats semblables (il est vrai datant d'avant les IUFM) ont été obtenus dans d'autres études de l'IREDU, dont Duru-Bellat et Leroy-Audoïn (1990), Bressoux (1993) et Mingat (1991) citées ici. De même, les évaluations de l'aménagement du temps scolaire n'ont en général pas établi l'existence d'un bénéfice pour les apprentissages.

²⁴ En revanche, Schmidt-Rolland et Tharel (1997), Verdon et Tharel (1997) trouvent un lien positif entre l'efficacité des enseignants de sixième et le nombre d'années universitaires suivies dans leur discipline.

Conclusion

Les recherches sur la RTC laissent avec bien des incertitudes, notamment :

- sur ses effets non académiques
- sur les processus qui expliquent ses effets académiques, et donc sur les moyens d'en accroître l'efficacité comme sur les raisons pour lesquelles ces effets perdurent.

Elles proposent cependant quelques conclusions relativement sûres sur les effets en situation expérimentale, si l'on se fonde en particulier sur les réanalyses des données de STAR et sur les méta-analyses les plus récentes :

- Pour être efficace, la réduction doit être importante et amener les classes nettement en dessous de 20 élèves.
- Une réduction importante ne conduit qu'à un effet faible, de l'ordre de 0.2 écart-type, soit, pour un élève médian, un gain de 8 rangs sur 100 ou de 2 rangs dans une classe de 25.
- Pour les élèves défavorisés, cet effet devient important. Il peut atteindre 0.4 écart-type, soit un gain de 16 rangs sur 100.
- Cet effet perdure longtemps après que les élèves ont rejoint de grandes classes. Si ce caractère durable exige ou non que la fréquentation des petites classes dure plus d'un an est discuté.

Les études corrélationnelles, en France comme aux Etats Unis, donnent des résultats moins probants, dans le primaire comme dans le secondaire. Certaines études concluent en faveur des grandes classes, d'autres en faveur des petites, la plupart à des effets non significatifs. Les modalités précises de la baisse, les mesures d'accompagnement, importent sans doute fortement.

Les recherches ne justifient donc certainement pas une RTC " au fil de l'eau " qui procède du fait qu'il est difficile de retirer un poste ou de fermer une classe lorsque les effectifs baissent, ni une baisse générale de deux ou trois élèves par classe. C'est le résultat le plus clair des études menées en France. Elles ne justifient pas non plus une attitude fondée sur l'idée que la RTC est forcément la politique la moins efficiente qui soit. Elles peuvent effectivement, semble-t-il, servir d'argument à une politique visant les populations défavorisées pendant les premières années du primaire, pourvu que la baisse soit importante, que des mesures de formation adéquates soient prises, et aussi que l'on puisse en mesurer les effets. En effet, s'il est vrai, comme le soutient Hanushek, que cette politique est efficace dans certaines conditions sans que nous connaissions exactement lesquelles, il n'est pas aberrant d'imaginer qu'on en limite le bénéfice aux écoles dans lesquelles elle se révèle efficace.

Toutefois, d'autres politiques, orientées vers un accroissement du temps d'enseignement, semblent pouvoir être plus efficaces et peut être aussi équitables²⁵. Une comparaison de l'efficacité de diverses politiques, plus approfondie que celle qui a été esquissée ici, est sûrement opportune.

Dans le secondaire, les recherches concluent de façons diverses, et aucune recherche française au collège ne laisse penser qu'une politique de RTC aurait des effets positifs. Toutefois, elles portent sur la sixième, tandis que les résultats de certaines études américaines, et certaines théories qui les expliquent, ouvrent la possibilité d'un effet positif de réductions fortes dans les niveaux où les élèves sont les plus difficiles (5^{ème} et 4^{ème}, semble-t-il), effet dont il peut être envisagé de tester la présence dans une académie.

Parmi les incertitudes qui demeurent, figure la question des dédoublements, autrement dit de l'effet de diminutions sporadiques de la taille du groupe. Nous n'avons trouvé aucune recherche à ce sujet. Or, si certaines des théories qui ont été avancées pour expliquer l'effet de la RTC, semblent pouvoir aussi expliquer un effet favorable des dédoublements (l'effet de congestion, par exemple), d'autres (l'effet de socialisation scolaire dans les petites classes) ne permettent pas de penser que les dédoublements aient un effet positif. *Il s'agit là sans doute d'un thème de recherche opportun.*

²⁵ Plusieurs études, française et américaines, montrent que les performances des élèves défavorisés baissent pendant les grandes vacances, tandis que celles des élèves favorisés ne baissent pas. Cela conduit certains chercheurs à préconiser des "écoles d'été" pour les premiers (par exemple, Krueger, in Goldwater Institute Journal (2000), Proceedings of the Charter School Research Conference in Scottsdale, Arizona, 1(2)). Cela dit, la durée horaire annuelle de l'enseignement est déjà élevée en France : Pour les élèves de 7 ans, seuls quatre systèmes éducatifs sur les seize systèmes de l'U.E. pour lesquels ces chiffres sont disponibles ont une durée plus élevée que la nôtre ; pour les élèves de 10 ans, six (Les chiffres clés de l'éducation dans l'Union Européenne, 1997).

Bibliographie

- Akerhielm, K., 1995, "Does class size matter ?" *Economics of Education Review*, 14 (3), 229-241.
- Annevelink, E. , Doolaard, S. et Bosker, R., 2001, *Adaptive Instruction, class size and pupil teacher ratio*, Communication au 14ème congrès School Effectiveness and School Improvement, Toronto.
- Blatchford, P., et Goldstein, H., 2000, *New evidence on class size*, Summary of a symposium at BERA annual conference, Cardiff, 2p.
- Bloom, B., 1984, *Le défi des deux sigmas*, in L'art et la science de l'enseignement, Crahay et Lafontaine, éd., Labor, Liège.
- Bressoux, P., 1993, *Les performances des écoles et des classes*, MEN-DEP, Dossiers Education et Formations, n°30.
- Card, D, et Krueger, A., 1990, *Does school quality matters ? Returns to education and the characteristics of public schools in the USA*, Working paper series, NBER, Cambridge, Massachussets.
- Card, D, et Krueger, A., 1996, "School resources and Student outcomes : An overview of the literature and New Evidence from North and South Carolina", *Journal of Economic Perspectives*, 10 (4), 31-50.
- Crahay, M., 2000, *L'école peut elle être juste et efficace?* de Boeck, Bruxelles, 440 p.
- Crahay, M., 1996, *peut-on lutter contre l'échec scolaire?* de Boeck, Bruxelles, 332 p.
- De Peretti, A. 1987, *Sur la taille des classes*, pp 99-126, in Pour une école plurielle, essais en liberté, Larousse, Paris.
- Duru-Bellat, M. et Leroy-Audoine, C., 1990, Les pratiques pédagogiques au CP, *Revue Française de Pédagogie*, N°93, 5-15.
- Duru-Bellat, M. et Kieffer, A., 1999, La démocratisation de l'enseignement revisitée, *Cahiers de l'IREDU*, N°60, 312p.
- Figlio, D.N., 1999, "Functional form and the estimated effect of school resources", *Economics of Education Review*, 18, 241-252.
- Fraser, B.J., 1987, "Synthesis of educational productivity research ", *International Journal of Educational Research* ,11(2).
- Génelot, S. 1997, " L'enseignement de l'anglais en sixième ", *Revue Française de Pédagogie*, N°118
- Glass G.V., Smith M.L., 1979, "Metaanalysis of research on class size and Achievement" , *Educational Evaluation and Policy Analysis*, 1(1).
- Glass G.V., Cahen L.S., Smith M.L., Filby, N.N., 1982, *School class size : research and policy*, Sage, Beverly Hills, California.
- Goldstein, H, Yang, M. et al., 1999, *Meta analysis using multilevel models with an application to the study of class size, final report*, 32p., Institute of Education, University of London.
- Goldstein, H. and Blatchford, P., 1998, "Class size and Educational Achievement : a review of methodology with particular reference to a study design", *British Educational Research Journal*, 24 (3), 255-268.
- Grisay, A., 1993, *Le fonctionnement des collèges et ses effets sur les élèves de sixième et de cinquième*, MEN-DEP, Dossiers Education et Formation, n°32, 208p.

- Hanushek, E.A., 1998, *The Evidence on Class Size*, University of Rochester : Institute of political Economy, 40p.
- Hoxby, C.M., 1998, *The effects of class size and composition on students achievement : New evidence from natural population variation*, Working paper 6869, NBER, Cambridge, Massachussets, 59p .
- Koretz, D.M. et Barron, S.L., 1998, *The validity of gain score on the Kentucky Instructional results Informational system (KIRIS)*, Rand, 130 p.
- Krueger A.B., 1997, *Experimental estimates of education production functions*, Working paper 6051, NBER, Cambridge, Massachussets. 31p.
- Lazear, E.P., 1999, *Educational Production*, Working paper 7349, NBER, Cambridge, Massachussets. 45p.
- Lee, V.E. et Loeb, S., 2000, "School size in Chicago elementary schools : Effects on teachers attitudes and student's achievement", *American educational research Journal*, 37(3), 3-32.
- Leroy Audoin, C. et Mingat, A. , 1995, *L'école primaire rurale en France*, Rapport pour la Direction de la Prévision du Ministère des Finances, 51p.
- Meuret, D. et Marivain, T., 1997, *Inégalités de Bien-être au collège*, MEN-DEP, Dossiers Education et Formations, n°89.
- Meuret, D., Broccholichi, S. et Duru-Bellat, M., 2000, *Le choix et l'autonomie des établissements scolaires : quels effets ?*, Cahiers de l'IREDU, N°61, Dijon, 252p.
- Mingat, A., 1991, "Expliquer les acquisitions au Cours Préparatoire : les rôles de l'enfant, de la famille et de l'école", *Revue Française de Pédagogie*, n° 95, pp. 47-63.
- Mac Beath, J., Myers, K., Kirwan, T., 2001, *Studying the development of study support : What's the evidence of increased student achievement ?*, Communication au 14ème congrès School Effectiveness and School Improvement, Toronto. 23p.
- Muijs, D. et Reynolds, D., 2000, Some findings from the evaluation of the mathematics enhancement program , *School effectiveness and school improvement*, 11(3) 273-303.
- Nye, B. Hedges L.V. and Konstantopoulos, S., 2000, "The effects of small classes on Academic achievement : the results of the Tennessee Class Size Experiment", *American Educational Research Journal*, 37(3), 123-152.
- Pong, S. et Pallas, A., 1999, *The Impact of class size on eighth-grade math achievement*, Population Research Institute, Working Paper 99-11, 27p.
- Pritchard, I., 1999, *Reducing class size, what do we know ?* , US Department of Education, OERI, 25p.
- Richarson, R., 1999, *Performance related pay in schools, An assessment of the green papers*, A report for the National union of teachers, London school of Economics, 32p.
- Schmitt-Rolland et Thauvel Richard, 1997, *pratiques pédagogiques de l'enseignement du français en sixième et progrès des élèves*, MEN-DPD, Dossier Education et Formations, n° 87, 192p.
- Slavin, R., 1990, *Class Size and Student Achievement : Is smaller better ?* *Contemporary Education*, 62(1), 6-12.
- Suchaut, B., 1996, *Le temps scolaire : Allocation et effets sur les acquisitions des élèves en Grande section et au CP*, Thèse de doctorat, Université de Bourgogne, 294p.
- Suchaut, B., 1998, *Evaluation d'un dispositif de lecture au CP*, Rapport technique, IREDU-Université de Bourgogne.
- Suchaut, B. et Le Bastard, S., 2000, *Lecture-écriture au cycle II, Evaluation d'une démarche innovante*, Cahiers de l'IREDU, Dijon, N°61.
- Tomlinson, T., 1988, *Class size and Public policy : Politics and Panacea*, US Deptment of Education, OERI.

Verdon, R. et Thauvel Richard, M. 1997, Pratiques pédagogiques de l'enseignement des mathématiques en sixième et progrès des élèves, MEN-DPD, Dossier Education et Formations, n° 842, 28p.

Wenglensky, H., 1997, *When money matters : How educational expenditures improve student performance and how they don't*, Policy information Perspective, Princeton, New Jersey.

Cahiers pédagogiques, 2000, Dossier "Travailler aussi en grand groupe", n° 385, juin.

California CSR Research consortium (CCSRRC), 2000, *Class size reduction in California : the 1998-99 evaluation findings*, 18p.

Educational Evaluation and Policy Analysis, 1999, 21(2), Special issue , *Class size : Issues and New findings* , Washington, DC.

Educational Research Service, 1978, *Class size, a summary of research*, Arlington (Virginie), 75p.

Ministère de l'Education nationale, Direction de la Programmation et du développement, 2000, L'Etat de l'école, 76p.

Texas Education Agency, 2000, *School size and class size in Texas Public Schools*, 32p., www.tea.state.tx.us/research/abs2.htm

US Department of Education, USDE, 2000, *The class size reduction program : Boosting Student Achievement in Schools across the Nation*, First year report, 16p.

**Présentation synthétique des principales recherches
sur l'effet d'une réduction de la taille des classes.**

Méta-analyses

Auteur	Date	Période	Nombre de recherches	Dont expérimentales	Résultats
Blake (Cité in ERS, 1978)	1954	Avant 1950	22		16 pour les petites classes 3 pour les grandes 3 n'observent pas de différences
Glass et Smith	1979	Jusqu'en 1977	77	5	Petites classes plus efficaces dans 60 % des comparaisons. Gains d'efficacité quasi nuls entre 40 et 20 élèves par classe.
Slavin	1990		7		Ecart de 0.17 écart-type entre les performances des élèves dans des classes de 15-16 vs 25-30 élèves, en faveur des petites classes.
Goldstein et Yang,	1999	1970 à 1991	9	9	Toutes les études portent sur la lecture. En moyenne, l'effet est d'une baisse du score de -0.02 écart type pour chaque élève en plus entre quinze et trente élèves.

Quelques recherches étrangères sur les acquisitions de connaissances

Auteur(s) (pays)	Date de l'observation	Méthode	Tailles des classes comparées	Niveau et taille de la population étudiée	Objet	Effets
Marklund (Suède) (cité in ERS, 1978)	1962	Corrélationnelle	16 –25 vs 26-35	Collège (élèves de 13 ans)	Epreuves scolaires dans plusieurs disciplines	Sur 281 comparaisons, 37 en faveur des grandes classes, 22 en faveur des petites, 222 non significatives.
Coleman (USA)	1966	Id.			Id.	Pas d'effet de la taille des classes
Wilsberg et Castiglione (USA, New York)*	1968	Expérimentale	15 vs ≥ 25	1127 élèves de CP et 513 de CE1	lecture	Effet standard ²⁶ en faveur des petites classes : +0.3
Balow * (USA)	1969	Expérimentale 2 classes	15 vs 30	CP, CE2	lecture	Effet standard en faveur des petites classes : + 0.17 au CP + 0.02 au CE2
Johnson et al. (South Carolina)	1977	Corrélationnelle	Moyenne de 20 vs 27	CP	maths, lecture	En faveur des petites classes pour la lecture. Pas d'effet significatif en Maths
Egelund (Dk.)	1979	Corrélationnelle	11 à 29	7 à 16 ans	danois, maths, anglais	Pas de corrélation
Shapson and al.* (Canada)	1980	Expérimentale	4 classes de 16,23,30, 37 élèves	CM1	Lecture, Maths	Effet standard en faveur des petites classes : .2 en Maths non significatif en lecture
Mazareas* (USA, Mass.)	1981	Expérimentale	<20 vs >30	CP, 368 in <20 vs 646 in >30	“ 5 scores dont la lecture ”	Effet standard en faveur des grandes classes : + 0.13
Wagner* (USA, Ohio)	1981	Expérimentale	15 vs ≥ 25	CE1, deux écoles		Effet standard en faveur des petites classes : + 0.4
Doss et Holey* (USA, Texas)	1982	Expérimentale	15 vs 30	Cohorte du CE1 jusqu'en sixième	Lecture, Langage, Mathématiques.	Effet standard en faveur des petites classes : +0.3 au CE1 et au CE2 +0.2 au CM1 +0.4 au CM2 et en sixième
Finn et Achilles ²⁷ * (USA, Tennessee, projet STAR)	Depuis 1985	Expérimentale	13-17 vs 22-25	Cohorte de la grande section jusqu'au CE2. 6000 élèves.	Maths, lecture	Effet standard en faveur des petites classes : GS : .18 en lec., .15 en maths CP : .24 en lec., .27 en maths CE1 : .23 en lec., .20 en maths CE2 : .26 en lec., .23 en maths

²⁶ Pour la définition de l'effet standard, voir note 3, p5.

²⁷ Answers and questions about class size : a state wide experiment, AERJ, 1990, 27,3, pp557-577 et EEPA, 1999.

Auteur(s) (pays)	Date de l'observation	Méthode	Tailles des classes comparées	Niveau et taille de la population étudiée	Objet	Effets
Akerhielm (USA, national)	1988	Corrélationnelle	1 à plus de 30	4 ^{ème} , 24000 élèves	maths, anglais, histoire, science	Effet en faveur des petites classes : Très faible en histoire et science, non significatif en maths et anglais
Figlio (USA, national)	1988 à 1990	Corrélationnelle	Taux d'encadrement de l'école	Une cohorte de 5800 élèves suivie de 4 ^{ème} en seconde.	science	Une hausse de 10% du nombre d'élèves par enseignant se traduit par une baisse de 1% du score en fin de seconde, sous contrôle du score de 4 ^{ème} .
San Juan School district* (USA, Californie.)	1991	Expérimentale	20 vs 30	3 ^{ème}	lecture, compréhension de texte	Effet standard en faveur des petites classes : +0.60
Wenglenski (USA)	1992	Corrélationnelle	<20 vs >20	CM1 ; 4 ^{ème} ;	maths	CM1 : effet direct de petites classes sur niveau de maths : 0.5 années d'avance 4 ^{ème} : effet de petites classes à travers amélioration de la discipline, surtout dans des zones pauvres.
Hoxby (USA, Connecticut)	1986 à 1997	Corrélationnelle	Variations naturelles, entre 15 et 30	CM1 ; 6 ^{ème} ; 4 ^{ème}	maths ; lecture ; expression écrite	Aucun effet significatif à 5%, à aucun des niveaux dans aucune discipline. Quelques effets en faveur des grandes classes, significatifs à 10%. Les résultats ne sont pas différents pour les élèves pauvres ou appartenant à des minorités.
Goldstein et Blatchford (Grande Bretagne)	1996 à 2000	Corrélationnelle (niveau initial)	Deux cohortes suivies les trois premières années du primaire	Entre 10 et 35 élèves par classe	Maths, lecture	Seuls sont disponibles les effets sur la première année : - existence d'effets significatifs en faveur des petites classes en lecture et en maths - effets non linéaires en maths : plus forts entre 15 et 25 qu'entre 25 et 30. - Effets plus forts pour les élèves faibles et pour les élèves éligibles aux repas gratuits.

* études prises en compte dans la méta-analyse de Goldstein et Yang (1999).

Quelques recherches sur d'autres dimensions

Auteur(s) (pays)	Date de l'observation	Méthode	Tailles des classes comparées	Niveau et taille de la population étudiée	Objet	Effets
Pugh	(1965)	Observation d'enseignants	20 ou moins vs 30 ou plus	Premier et second degré	Modes d'enseignement	Davantage d'enseignement individualisé et d'activités en petits groupes dans les petites classes, mais une grande quantité d'enseignement reste adressé à la classe entière. Activités plus variées dans les petites classes. NB : recherche critiquée pour les concepts utilisés pour décrire l'activité des enseignants.
Yeany	(1976)	Observation video de 64 enseignants stagiaires	CM1-CM2		Modes d'enseignement	Les modes d'enseignement ne varient pas avec la taille des classes
Glass et Smith (1982)	Jusqu'en 1979	Synthèse de 60 recherches			Attitude vis à vis de l'école	Sur 172 comparaisons, 85% sont en faveur des petites classes pour des dimensions comme : motivation, attention, intérêt pour l'école, créativité,...
Egelund (Danem.)	1979	Corrél..	7 à 16 ans	De 11 à 29	Climat social de la classe	Meilleur climat dans les grandes classes.
Blatchford et Goldstein, 2000	1997	Corrél.	4 à 7 ans		Climat social de la classe	Les maîtres des petites classes déclarent plus souvent que les élèves sont agressifs et que certains d'entre eux sont rejetés.

Recherches françaises

Auteur(s)	Date de l'observation	Méthode	Tailles des classes comparées	Niveau et taille de la population étudiée	Objet	Effets
Auvinet ²⁸	1958-59			Enseignement élémentaire	Réussite au certificat d'études Taux de retards	“ Un effectif moyen compris entre 30 et 40 correspond à de meilleurs résultats qu'un effectif inférieur à 30 ”
Legrand	1949 vs 1962		25.4 en 49 vs 28.5 en 62	Tout le premier degré	Taux de retards	Le % d'élèves en retard est descendu de 16.2 à 15.4
SIGES	1978			CP et CE1	Taux de retards	Les taux de redoublements baissent quand la taille des classes augmente
Cherkaoui	1979	Corrél.		Lycée	Connaissances	Effet positif de la taille des classes pour des élèves défavorisés en sections techniques.
Mingat, 1991	1989	Corrél.		CP	Lecture et maths	Léger effet négatif de la taille des classes (-0.01 écart type par élève supplémentaire, S ; à 5%). L'effet est non significatif en maths, égal à -0.02 écart-type par élève en plus en français.
Bressoux 1993	1990-91	Corrél.	10 à 33	CE2 au CM2	lecture	Pas d'effet significatif. Effet négatif de la taille dans les classes à cours multiples (S à 10%).
Leroy-Audoin et Mingat 1995	1993-94	Corrél.		CE2	français et maths	Effet non significatif dans les classes à cours simples. Effet <i>positif</i> de la taille dans les classes à cours multiples.
Suchaut 1996	1992	Corrél.	17 à 33	Grande section	Pré-requis pour l'entrée au CP	Effet non significatif .
Suchaut 1996	1992	Corrél.	13 à 30, dont les deux tiers entre 19 et 25	CP	français, maths	En maths, effet positif de la taille des classes (S à 5%). En français, effet non significatif

²⁸ Les trois études premières citées le sont par de Peretti, A. Pour une école plurielle, Larousse, 1987

Auteur(s)	Date de l'observation	Méthode	Tailles des classes comparées	Niveau et taille de la population étudiée	Objet	Effets
Schmitt-Rolland et Thauvel-Richard, 1997	1993-94	Corrél	13 à 31	6ème	français	Effet non significatif .
Verdon et Thauvel-Richard, 1995	1993-94	Corrél.		6ème	maths	Effet non significatif.
Suchaut, 1998	1997	Corrél.	15 à 26	CP	Lecture	Effet non significatif.
Suchaut et Le Bastard, 2000	1998-1999	Corrél.	12 à 30, dont les deux tiers entre 16 et 25	CP	Lecture-écriture	Effet positif de la taille des classes jusqu'à 23 élèves, négatif ensuite (S à 1%).

Rapports établis à la demande du Haut Conseil de l'évaluation de l'école

Directeur de la publication : **Claude THÉLOT**

Secrétariat général : 3/5 boulevard Pasteur 75015 PARIS

Tel : 01 55 55 77 14

Fax : 01 55 55 77 62

Adresse du site du HCéé : <http://cisad.adc.education.fr/hcee>

ISSN en cours

Conception et impression – DPD/BED