

CONFÉRENCE DE C O N S E N S U S

**DIFFÉRENCIATION PÉDAGOGIQUE :
COMMENT ADAPTER L'ENSEIGNEMENT
POUR LA RÉUSSITE DE TOUS LES ÉLÈVES ?**

La différenciation dans les systèmes éducatifs :
pourquoi, comment, avec quels effets ?

Dominique LAFONTAINE

Mars 2017

La différenciation dans les systèmes éducatifs : pourquoi, comment, avec quels effets ?

Dominique Lafontaine^o

^o Université de Liège
Département des Sciences de l'Éducation

Mars 2017

Contribution dans le cadre de la conférence de consensus sur la différenciation pédagogique du Cnesco et de l'Ifé/ENS de Lyon.



Ce rapport s'inscrit dans une série de contributions publiées en 2017 par le Conseil national d'évaluation du système scolaire (Cnesco) dans le cadre de la conférence de consensus du Cnesco et de l'Ifé/ENS de Lyon sur la différenciation pédagogique les 7 et 8 mars 2017.

Bien que ce rapport ait fait l'objet d'une relecture par le Cnesco, les opinions et arguments exprimés n'engagent que l'auteur.

Pour citer ce document, merci d'utiliser la référence suivante :

Lafontaine, D. (2017). *La différenciation dans les systèmes éducatifs : pourquoi, comment, avec quels effets ?*
Rapport commandé par le Cnesco. <http://www.cnesco.fr/fr/differenciation-pedagogique/>

Publié en Mars 2017
Conseil national d'évaluation du système scolaire
Carré Suffren - 31-35 rue de la Fédération
75015 Paris

Table des matières

Introduction	6
I La différenciation verticale : le redoublement	8
1 L'importance du taux de retard dans les pays de l'OCDE	8
2 L'efficacité pédagogique du redoublement	11
3 Le retard scolaire amplifie-t-il les inégalités liées à l'origine socioéconomique et culturelle?	14
II La différenciation horizontale : effets du regroupement des élèves par aptitude sur les acquis scolaires	16
1 Les études quasi expérimentales	17
2 Les études en milieu naturel	19
III L'impact des mécanismes de différenciation sur les performances et les inégalités scolaires et sociales : l'éclairage des enquêtes internationales	21
1 Regroupement des élèves par niveau, performances, inégalités scolaires et sociales	21
2 Quand les politiques éducatives accentuent les différences dans les contextes d'apprentissage	29
IV L'impact du regroupement des élèves par aptitude sur le plan socio-affectif	36
V Quelles modalités d'organisation/différenciation pour une école plus efficace et moins inégalitaire?	43
VI Différencier, oui, mais comment? Questions, tensions, débats.	50
Bibliographie	52

Table des figures

Figure 1 Regroupement homogène des élèves selon des critères sociaux et déterminisme social.....	25
Figure 2 Regroupement homogène des élèves selon des critères sociaux et déterminisme social.....	31

Figure 3	Décomposition de la variance école (en pourcentage de la variance totale) (Source : base de données PISA 2009) - Enseignement secondaire	33
-----------------	--	-----------

Liste des tableaux

Table 1	Pourcentage d'élèves en retard à l'âge de 15 ans dans les pays de l'OCDE (PISA 2009, PISA 2012)	8
Table 2	Influence (sous la forme d' <i>Odds ratio</i>) de l'origine sociale (SES) sur le redoublement, avant et après contrôle de la performance scolaire (seules les valeurs statistiquement significatives à 0,05 sont reprises dans le tableau)	15
Table 3	Pourcentage de variation de la performance entre écoles ou entre classes (<i>Source</i> : bases de données PIRLS 2006, TIMSS 1995 et PISA 2009) et de l'origine sociale des élèves (PISA 2009).	26
Table 4	Corrélation entre pourcentages de variation entre écoles et performances, par discipline et par cycle (Pisa 2003, Pisa 2006 et Pisa 2009).....	27
Table 5	Corrélation entre pourcentages de variation entre écoles et dispersion de la performance, par discipline et par cycle (Pisa 2003, Pisa 2006 et Pisa 2009).	27
Table 6	Corrélation entre ségrégation académique et sociale et déterminisme social.	28
Table 7	Corrélation entre ségrégation académique et sociale et déterminisme social des aspirations académiques et socioprofessionnelles	29

Introduction

Chaque élève, chaque apprenant, est différent : différent dans ses aptitudes et ses attitudes, sa capacité et sa motivation à apprendre, son rythme d'apprentissage de matières, savoirs ou habiletés qu'il a choisi d'apprendre, ou qui lui sont imposées dans un programme d'études. Toute forme d'éducation ou d'enseignement est confrontée à ce défi central - faire apprendre les mêmes contenus, faire acquérir des compétences similaires - à des élèves qui sont, par nature ou compte tenu d'apprentissages antérieurs, bien différents. Sans doute est-ce pour cela que la forme d'enseignement la plus efficace est le préceptorat où un enseignant s'occupe d'un seul élève à la fois, et peut adapter ses méthodes et définir pour lui la progression idéale dans les apprentissages. Quand on porte le questionnement au niveau des systèmes éducatifs, le défi est qu'il s'agit d'enseigner à des individus singuliers, mais qui sont nichés dans des groupes, dans des classes, dans des écoles dont le degré d'hétérogénéité peut être d'ampleur variable, mais n'est jamais nul.

Pour entrer en matière, considérons d'abord les élèves qui sont d'emblée reconnus comme "différents" - frappés d'un handicap, présentant un trouble reconnu ou, à l'opposé, des dons particuliers, ce que l'on appelle aujourd'hui enfants à "haut potentiel". Comment éduquer ces enfants, ces jeunes, dans quelle(s) structure(s), avec quelle adaptation de l'enseignement à leurs besoins spécifiques ?

Au départ de ces cas, on comprend aisément que l'alternative se situe entre une **éducation "à part", spécialisée**, conçue pour répondre à leurs caractéristiques particulières, mais qui les isole au moins partiellement des autres enfants, ou une **éducation inclusive**, qui les **intègre** dans des classes dites normales avec une adaptation voire des moyens matériels ou humains supplémentaires. Cette tension entre éduquer ensemble le plus et le plus loin possible des enfants différents, ou les séparer et les répartir dans des groupes distincts avec des enfants qui présentent des caractéristiques plus semblables aux leurs, concerne tous les élèves, et traverse tous les choix éducatifs que doivent poser tant les familles que les équipes pédagogiques ou les décideurs en matière de politiques éducatives. Faut-il, pour mieux faire apprendre les uns et les autres, séparer les élèves, ou les intégrer, les éduquer ensemble ? **Différencier l'enseignement** est inévitable, mais comme **Legrand (1995)** le disait si justement : "différencier l'enseignement peut avoir deux sens complémentaires. Il s'agit, dans tous les cas, de prendre en compte la réalité individuelle de l'élève. Mais cette prise en compte peut se faire en considération de deux objectifs différents : ou bien il s'agit d'adapter l'enseignement à la destination sociale et professionnelle des élèves ; ou bien, un objectif commun étant défini et affiché, il convient de prendre en compte la diversité individuelle pour y conduire" (**Legrand, 1995**, p. 6).

Tout système éducatif peut être caractérisé par sa propension à séparer ou intégrer les élèves et de multiples combinaisons existent entre les modèles d'organisation constituant les deux extrémités du continuum de la différenciation. En une première approche, listons simplement les mécanismes les plus répandus consistant à séparer certains élèves de leur cohorte d'âge ou groupe d'appartenance hétérogène pour les regrouper, de manière permanente, avec des élèves qui leur ressemblent davantage, créant de la sorte des groupes moins hétérogènes dans leurs aptitudes et leur degré de mixité sociale voire de genre :

- enseignement ordinaire *versus* spécialisé ;
- redoublement (les élèves moins performants sont regroupés avec des élèves plus jeunes, moins avancés) ;

- classes de niveau constituées sur la base des résultats antérieurs ;
- écoles de niveau ou de réputations différentes (surtout dans les systèmes éducatifs où le choix d'école est libre, sans sectorisation) ;
- filières liées à une orientation (académique, technique, artistique, sportive, professionnelle...);
- options dont la difficulté relative ou la rareté engendre *de facto* des hiérarchies et donc des regroupements par aptitude (par exemple : enseignement immersif, choix des langues anciennes ou du chinois).

Parmi ces mécanismes de différenciation, il est d'usage de distinguer la **différenciation verticale (redoublement)** de la **différenciation horizontale** (tous les autres mécanismes listés ci-avant). À côté de ces mécanismes créant des regroupements par niveau relativement permanents, d'autres mécanismes de différenciation de l'enseignement opèrent aussi des regroupements par niveau, mais transitoires. Il en va ainsi des groupes flexibles, de besoin, qui se créent dans une discipline, de manière temporaire, voire de la création de cours de niveaux différents dans les principales disciplines (cours de base ou cours avancés en langues, en mathématiques, en sciences) comme cela se pratique couramment dans les pays anglo-saxons¹. Nous les envisagerons aussi dans cette revue, mais en les distinguant des premiers, car le caractère permanent ou temporaire de ces regroupements par niveau est précisément ce qui fait toute la différence... Enfin, il existe des pratiques de différenciation pédagogiques internes à la classe, qui ne seront pas prises en compte dans ce rapport, mais qui ont été analysées dans le rapport rédigée par Alexia Forget (2017) pour le Cnesco.

L'objectif de cette synthèse est double. D'une part, il s'agira de **décrire la diversité des modes d'organisation** qui existent de par le monde pour prendre en compte et gérer l'hétérogénéité des publics scolaires en s'appuyant sur des travaux d'éducation comparée et sur les enquêtes internationales. D'autre part, il s'agit d'apporter une réponse fondée scientifiquement à la question des **effets de ces modes d'organisation**, en synthétisant les apports des nombreuses recherches expérimentales ou en milieu naturel existantes. Au-delà des synthèses déjà dressées dans les méta-analyses - regroupant elles-mêmes des centaines de recherches, le présent rapport privilégiera les recherches utilisant des données internationales ou les recherches menées dans des systèmes éducatifs contrastés.

Lorsque l'on s'interroge sur les effets de ces modes de gestion de la diversité, il convient de "balayer large". Il s'agit d'abord d'envisager, à côté des **effets cognitifs** (acquis scolaires, performances à des tests standardisés), les **effets socio-affectifs** (motivation, concept et estime de soi, sentiment d'efficacité, aspirations d'études ou professionnelles). Ensuite, il convient de prendre en compte, à côté des indicateurs de performances ou d'**efficacité**, des indicateurs de dispersion ou d'**équité** tels que l'écart entre les élèves les plus forts et les plus faibles et les écarts de performances liés à l'origine socio-économico-culturelle des élèves ou encore les différences de performances entre écoles. Enfin, il convient de se demander si les effets bénéfiques ou délétères concernent tous les élèves ou affectent plus particulièrement certaines catégories d'entre eux - les plus faibles, les plus défavorisés socialement, les filles ou les garçons par exemple (**effets d'interaction**).

1. En anglais, on oppose ainsi le *tracking* ou *streaming* (*between-class ability grouping*) au *setting* (*within-class ability grouping*) (Gamoran, 2002).

Le présent rapport² est organisé de la façon suivante :

- dans la première partie, les effets spécifiques des deux principaux mécanismes de différenciation verticale et horizontale, le redoublement et le regroupement par aptitude (classes de niveau, filières, groupes de niveau) sont examinés ;
- dans la deuxième partie, sont présentés des travaux menés principalement au départ des bases de données internationales qui ont étudié l'impact de différents modèles d'organisation mis en place par les systèmes éducatifs pour gérer les différences entre élèves ;
- la troisième partie est consacrée à la question de l'impact de la fréquentation d'écoles ou de classes plus ou moins performantes sur la motivation, en particulier sur le concept de soi académique.

Le rapport se conclut sur une série de recommandations et d'interrogations pour la pratique, en soulignant les points de mise en tension, de débat et de choix autour de la question de la différenciation.

I La différenciation verticale : le redoublement

1 L'importance du taux de retard dans les pays de l'OCDE

L'étude PISA³, complémentaire à la mesure des performances en mathématiques, en lecture et en sciences, prélève un vaste ensemble d'informations contextuelles auprès des élèves testés. Les élèves sont ainsi invités à mentionner s'ils ont redoublé une fois ou plus d'une fois une année d'étude durant leur scolarité primaire, secondaire inférieure ou secondaire supérieure. Ces données permettent de déterminer si l'élève présente un retard d'un an ou de plus d'un an et, ainsi, d'estimer pour chacun des pays participants la proportion d'élèves de 15 ans en retard. Le tableau 1 présente, par pays de l'OCDE, le pourcentage d'élèves en retard pour PISA 2009 et PISA 2012⁴.

Tableau 1 – Pourcentage d'élèves en retard à l'âge de 15 ans dans les pays de l'OCDE (PISA 2009, PISA 2012)

			Pourcentage d'élèves de 15 ans qui déclarent avoir redoublé au moins une fois durant leur scolarité⁵		
			Pisa 2009	Pisa 2012	Différence 2009-2012⁶
Belgique	(Fédération Wallonie-Bruxelles)		46,0 %	47,8 %	

2. Plusieurs passages de ce rapport empruntent à des écrits existants rédigés en collaboration avec différents collègues de l'Université de Liège. Nous les remercions ici au passage.

3. Enquête internationale menée à l'initiative de l'OCDE depuis 2000 auprès d'échantillons représentatifs d'élèves de 15 ans dans les pays de l'OCDE et dans certains pays dits "partenaires".

4. Pour un panorama plus complet des pratiques et réglementations en matière de redoublement, on consultera le rapport qu'[Eurydice \(2011\)](#) a consacré à la question.

5. Les données 2012 sont manquantes (M) pour le Japon et la Norvège.

6. Seules les différences significatives sont mentionnées.

France	36,9 %	28,4 %	-8,5 %
Luxembourg	36,5 %	34,5 %	-2,0 %
Espagne	35,3 %	32,9 %	-2,4 %
Portugal	35,0 %	34,3 %	
Belgique (Communauté germano- phone)	34,1 %	31,7 %	
Pays-Bas	26,7 %	27,6 %	
Belgique (Communauté flamande)	26,0 %	27,4 %	
Chili	23,4 %	25,2 %	
Suisse	22,8 %	19,9 %	-2,9 %
Mexique	21,5 %	15,5 %	-6,1 %
Allemagne	21,4 %	20,3 % ⁷	
Italie	16,0 %	17,1 %	
États-Unis	14,2 %	13,3 %	
Turquie	13,0 %	14,2 %	
Autriche	12,6 %	11,9 %	
Irlande	12,0 %	8,6 %	-3,3 %
Hongrie	11,1 %	10,8 %	
Canada	8,4 %	8,0 %	
Australie	8,4 %	7,5 %	-0,9 %
Israël	7,5 %	1,9 %	-5,6 %
Grèce	5,7 %	4,5 %	
Estonie	5,6 %	3,5 %	-2,1 %
Pologne	5,3 %	4,2 %	
Nouvelle-Zélande	5,1 %	5,4 %	
Suède	4,6 %	4,0 %	
Danemark	4,4 %	4,7 %	
République tchèque	4,0 %	4,9 %	
République slovaque	3,8 %	7,6 %	3,8 %
Finlande	2,8 %	3,8 %	1,0 %
Royaume-Uni	2,2 %	2,7 %	

7. Il manque 14 % des données 2012 pour l'Allemagne.

Slovénie	1,5 %	3,4 %	1,9 %
Islande	0,9 %	1,2 %	
Japon	0,0 %	M	
Corée	0,0 %	3,6 %	
MOYENNE	14,71	14,48	

Les pays sont classés par ordre décroissant, en fonction de leur taux de retard en 2009.

Tant en 2009 qu'en 2012, c'est la Fédération Wallonie - Bruxelles (Belgique francophone) qui arrive en tête de classement. En 2009, elle est suivie par la France. La moyenne des taux de retard des pays repris dans le tableau 1 est de 14,7 % en 2009 et 14,5 % en 2012. Dans 17 pays sur 34, le taux d'élèves en retard à 15 ans est inférieur à 10 %. Autrement dit, **dans une majorité de systèmes éducatifs, on redouble très peu, voire pas du tout**. La Fédération Wallonie - Bruxelles, la France, le Luxembourg, l'Espagne et le Portugal, en recourant massivement au redoublement, font figure d'exception.

Par ailleurs, ces résultats permettent déjà de battre en brèche l'idée selon laquelle une pratique intensive du redoublement irait de pair avec un haut niveau de performances : les pays qui apparaissent dans la partie supérieure du classement en fonction du taux de redoublement affichent des scores très moyens aux épreuves PISA. La corrélation entre taux de retard et performances en mathématiques dans PISA 2012 n'est d'ailleurs pas statistiquement significative (0,06, $p=,74$). Les pays dans lesquels beaucoup d'élèves ont redoublé ne sont ni plus ni moins performants dans PISA que ceux dans lesquels le redoublement est peu ou pas du tout pratiqué.

En comparant les données 2009 et 2012, on peut constater que le redoublement est en recul dans neuf systèmes éducatifs et en légère augmentation dans trois pays seulement. Le recul est sensible en France et en Allemagne.

Comme le souligne l'étude Eurydice de 2011 consacrée au redoublement, il n'existe pas de relation univoque entre les réglementations en matière de redoublement et leur incidence statistique. Ainsi, parmi des systèmes éducatifs qui connaissent des réglementations restrictives assez similaires, des variations importantes dans les taux de redoublement sont observées. Par ailleurs, des systèmes éducatifs où la réglementation est quasi inexistante affichent des taux de redoublement faibles. Il n'y est en somme pas nécessaire d'interdire le redoublement, car celui-ci ne fait pas partie des pratiques pédagogiques en usage et personne ne songerait à l'utiliser pour remédier à des difficultés d'apprentissage. "C'est une culture du redoublement qui expliquerait son recours fréquent. Dans les pays où cette pratique est récurrente, la conception selon laquelle répéter une année est bénéfique pour les apprentissages de l'élève reste très présente. Cette vision est partagée par le corps enseignant, la communauté scolaire et les parents eux-mêmes. En Europe, c'est surtout en Belgique, en Espagne, en France, au Luxembourg, aux Pays-Bas et au Portugal que cette conception persiste dans la pratique. L'adaptation des réglementations en matière de redoublement n'est pas suffisante pour

changer cette conviction. Celle-ci doit être remplacée par une autre approche de la gestion des difficultés d'apprentissage des élèves. En définitive, le défi majeur réside plus dans la remise en question de certaines convictions et croyances⁸ qu'en des changements de réglementations" (Eurydice, 2011, p. 61).

2 L'efficacité pédagogique du redoublement

L'intensité de la pratique du redoublement et le coût qu'il engendre amènent inmanquablement à poser la question de son efficacité. Très souvent, en faisant répéter son année à un enfant scolarisé dans l'enseignement maternel ou primaire⁹, un enseignant ou une équipe éducative postule que cette répétition va lui permettre de s'améliorer. Mais est-ce vraiment le cas ? Est-il efficace de faire répéter une année ?

S'interroger sur l'efficacité du redoublement impose de se demander si les élèves en difficulté qui redoublent progressent dans leurs apprentissages, si leurs résultats scolaires s'améliorent. Pour répondre à cette question, on peut se baser sur les constats de terrain, et c'est ce que font généralement les enseignants : lorsque dans leur classe, ils comparent les résultats d'un élève qui a redoublé avec les résultats du même élève l'année précédente, ils observent une amélioration.

Parallèlement aux constats de terrain, évaluer l'efficacité de la pratique du redoublement a été et reste la préoccupation de nombreux chercheurs. Les résultats des travaux les plus récents vont dans le même sens que le constat des enseignants : un élève qui redouble s'améliore durant l'année qu'il répète. Toutefois, les recherches menées sur la problématique s'efforcent de compléter l'analyse en essayant de se poser la question de l'efficacité du redoublement dans les termes suivants : que serait devenu le même élève s'il n'avait pas redoublé ? Aurait-il moins, autant ou davantage progressé ?

Pour fournir des éléments de réponse probants à cette question, les chercheurs appartiennent, au sein d'un important échantillon d'élèves, ceux qui ont des niveaux de compétence similaires (niveaux établis sur la base d'évaluations externes). Les différences d'exigences entre enseignants et entre écoles font qu'à niveau de compétence égal (ou plus précisément à niveau de difficultés égal), certains élèves redoublent et d'autres passent de classe. Les chercheurs examinent alors comment des élèves d'aptitudes comparables avant la décision de redoublement ou de passage de classe évoluent dans les mois et les années qui suivent cette décision. Leurs conclusions sont instructives : certes **l'élève qui redouble s'améliore, mais celui qui passe de classe et qui avait les mêmes difficultés s'améliore davantage.**

De très nombreuses recherches quasi expérimentales ont été consacrées au redoublement et plusieurs méta-analyses synthétisant ces recherches ont été consacrées à la question (Holmes, 1990 ; Jimerson, 2001). La synthèse des méta-analyses réalisée par Hattie (Hattie, 2009) donne une taille de l'effet moyenne de - 0,16¹⁰ : les élèves promus réussissent un peu mieux que les élèves qui ont connu le redoublement. La méta-analyse de Jimerson (2001) met aussi en évidence un effet négatif sur les performances scolaires (taille de l'effet de - 0,39, basée sur 169 effets) et sur les variables socio-émotionnelles (-0,22, basée sur 246 effets) et surtout sur le risque de décrochage (-0,65).

Plus récemment, certains chercheurs ont dénoncé les biais méthodologiques affectant la quasi-totalité des

8. Pourcentage d'élèves en retard à l'âge de 15 ans dans les pays de l'OCDE (PISA 2009, PISA 2012)

9. Dans l'enseignement secondaire, le redoublement est davantage vu comme un moyen de pousser les élèves à s'engager dans le travail.

10. Il est d'usage de considérer qu'une taille de l'effet doit dépasser 0,15 ou 0,20 pour être considérée comme digne d'intérêt. Un effet entre 0,20 et 0,40 est considéré comme léger ou modéré ; au-delà de 0,40, l'effet est considéré comme important sur le plan éducatif.

études sur lesquelles s'étaient fondées les premières grandes synthèses ou méta-analyses sur le redoublement de Shepard et Smith (1989), de Holmes (1990) et, dans une moindre mesure, de Jimerson (2001). On a ainsi montré que des techniques de contrôle plus sophistiquées étaient indispensables pour éviter que la comparaison entre l'échantillon de redoublants et le groupe témoin de non-redoublants ne penche systématiquement en faveur des non-redoublants, en raison des biais spécifiques liés à l'étude des groupes de sujets se situant aux extrêmes d'une distribution (élèves les plus faibles ou les plus forts). Plusieurs méthodes nouvelles ont été mises en œuvre pour remédier à ces artefacts : contrôle fondé sur de nombreuses variables sociodémographiques et utilisation de techniques d'analyse d'inférences causales dans la remarquable étude de Hong et Raudenbush (2005) ; méthode utilisant la discontinuité de régression dans les cas où le redoublement est décidé sur la seule base d'une note-seuil à un test ou un examen de passage standardisé¹¹ chez Jacob et Lefgren (2004, 2009) ; méthode s'appuyant sur l'utilisation de variables instrumentales¹² (Dong, 2010), ou encore analyse structurale (Grisay, 2001).

Cette nouvelle génération d'études a permis de nuancer les résultats de la première, en montrant que le véritable effet *à court terme* du maintien en maternelle ou du redoublement à l'école primaire est moins négatif que ne le laisse supposer une comparaison prétest/post-test de type classique effectuée sur les mêmes données. La répétition d'une année s'avère même dans quelques cas légèrement positive pour les résultats scolaires durant l'année ou les deux années suivantes (Dong, 2010), ce qui explique un peu mieux pourquoi les enseignants ont toujours cru, croient, et continueront à croire que le redoublement "remet en selle" les élèves faibles. Mais, à *moyen et long termes*, les conséquences négatives sur le plan des apprentissages et de la confiance en soi de l'élève et, surtout, les risques accrus d'abandon scolaire ont été confirmées (Ou et Reynolds, 2010a).

Les travaux de Goos en Belgique (Goos, 2013) s'inscrivent dans cette nouvelle génération de recherches. Goos a étudié les effets du redoublement en 1^{re} année primaire sur les apprentissages, les aspects psychosociaux et la carrière scolaire ultérieure des élèves en Flandre. Elle a ainsi suivi 3 264 élèves, issus de 121 écoles, de leur entrée en primaire jusqu'au début de l'enseignement secondaire. Les élèves ont été testés à intervalles réguliers en lecture et en mathématiques, ainsi que sur divers aspects psycho-sociaux de l'apprentissage : relations agressives avec les pairs, comportements d'hyperactivité ou asociaux, intérêt pour l'école et confiance en soi. Goos utilise les modèles statistiques les plus récents et adéquats pour tenir compte des biais de sélection et combine systématiquement une approche comparant les élèves promus et redoublants à âge constant et à année d'étude constante.

Elle met ainsi en évidence que pendant l'année qu'ils répètent, les redoublants de 1^{re} année réussissent mieux en mathématiques et en lecture que leurs pairs "à risques" qui ont été promus. Mais cet effet positif ne dure guère, puisqu'il disparaît complètement dès la 2^e année primaire, c'est-à-dire dès le moment où les redoublants doivent apprendre de nouvelles matières. Par ailleurs, les élèves qui ont redoublé leur 1^{re} primaire auraient mieux réussi les épreuves de mathématiques et de lecture tout au long de leur scolarité

11. Cette technique consiste à comparer les groupes d'élèves admis à l'examen de passage mais dont le score était supérieur de très peu à la note seuil au groupe de ceux qui ont dû redoubler alors que leur note était très proche de ce seuil. Ces différences étant inférieures à la marge d'erreur inhérente à l'épreuve, l'appartenance à l'un ou l'autre groupe d'élèves peut être considérée comme effectivement aléatoire, dans le cas où aucun autre facteur que le résultat à l'examen n'est pris en compte dans la décision d'admission ou de refus d'admission.

12. On peut lire Gary-Bobo et Robin (2011) pour une description plus détaillée de plusieurs de ces techniques, et une synthèse des études récentes sur le redoublement faisant appel à des méthodes inspirées des travaux sur l'économie de l'éducation.

primaire s'ils avaient été promus plutôt que maintenus en 1^{re} année primaire. Sur le plan psychosocial, les élèves redoublants montrent des profils plus négatifs sur l'ensemble des aspects évalués; à nouveau, les indicateurs auraient été moins négatifs si les élèves avaient été promus.

Le suivi longitudinal tout au long de la scolarité primaire permet de voir que les élèves qui ont redoublé leur 1^{re} primaire connaissent ensuite moins le redoublement que leurs pairs promus. Toutefois, ils sont plus susceptibles de changer d'école primaire, d'être orientés vers l'enseignement spécialisé ou vers des classes de pré-orientation vers l'enseignement professionnel.

Une méta-analyse menée par [Allen et al. \(2009\)](#), recensant 207 effets issus de 22 études publiées entre 1990 et 2007 confirme le lien existant entre la qualité méthodologique de l'étude et ses résultats. Les études les plus rigoureuses trouvent des effets moins négatifs au redoublement (effet moyen de 0,04 pour les études de meilleure qualité, contre -0,30 pour les études de moins bonne qualité). Toutefois, comme le soulignent les auteurs, "ces résultats ne peuvent constituer un argument solide pour les tenants du redoublement. Étant donné les coûts du redoublement et les dégâts sur le plan émotionnel, une différence non significative du redoublement sur les acquis scolaires met en cause les bénéfices pédagogiques des politiques de redoublement" ([Allen et al., 2009](#), p. 492).

En résumé¹³, les nombreuses recherches menées depuis des décennies sur la question des effets du redoublement, qu'ils soient cognitifs ou socioaffectifs, ne plaident pas en faveur du redoublement comme moyen de remédier aux difficultés d'apprentissage : non seulement les élèves faibles promus progressent mieux que les élèves qui ont redoublé, mais les effets sur l'estime de soi et la persévérance sont systématiquement négatifs. La nouvelle génération d'études apparue dans les années 2000 qui s'appuie sur des avancées d'ordre statistique met en évidence des effets qui peuvent être positifs à court terme, ce qui correspond au ressenti des enseignants et des parents. Toutefois, ces études confirment que les effets à plus long terme, qui sont au final les seuls qui importent vraiment, s'avèrent négatifs.

Pourquoi n'est-il pas efficace de maintenir ou de faire redoubler un élève ? Sans doute parce que **le redoublement est du "prêt-à-porter" là où il faudrait du "sur mesure"** : c'est une réponse tardive, qui ne cible pas spécifiquement les difficultés rencontrées par l'élève et qui le replace dans les conditions qui ont conduit à l'échec. Peterson, cité par [Allen et al. \(2009\)](#) ainsi que [Goos \(2013\)](#) se sont penchés sur le soutien reçu par les élèves durant l'année qu'ils répètent. Leurs résultats, menés dans des contextes différents, vont dans le même sens : dans la plupart des cas, le redoublement ne s'accompagne pas d'un soutien pédagogique accru ou spécifique, "le redoublement semble être considéré comme LA mesure pédagogique" en lui-même ([Allen et al., 2009](#), p. 495).

Non seulement la répétition d'une année entière est une solution inefficace et inadaptée (effet limité sur les apprentissages à court terme), mais en plus, c'est une pratique contre-productive : ses conséquences sur la confiance en soi accroissent à moyen et à long termes les risques de décrochage scolaire. **Un élève qui a connu un échec sanctionné par un redoublement construit une image négative de lui-même. Lorsque la confiance en soi diminue, l'élève se "résigne", l'apprentissage et l'engagement scolaire sont de moins en moins garantis.**

En revanche, et ceci ne surprendra pas, avancer les élèves les plus brillants, les faire "sauter" une classe produit des effets favorables, surtout si on compare ces élèves "avancés" à des élèves du même âge (l'effet

13. Les conclusions de la conférence de consensus sur le redoublement organisée par le CNESEO en 2014 vont largement dans le même sens.

est alors de 0,80). Si on les compare à des élèves qui fréquentent la même année d'étude, ils font aussi bien (effet de - 0,04) (Kulik, 2004). Selon Hattie, si l'accélération n'est pas plus souvent pratiquée, c'est sur la base de l'argument qu'elle serait défavorable sur le plan psychologique. Toutefois, selon une méta-analyse réalisée par Kent en 1992, cité par Hattie (2009), si les bénéfices sur le plan socio-émotionnel sont moindres que sur le plan cognitif, ils n'en sont pas pour autant négatifs comme on le pense souvent ; ils seraient même légèrement positifs.

3 Le retard scolaire amplifie-t-il les inégalités liées à l'origine socioéconomique et culturelle ?

C'est un fait établi depuis longtemps : tous les élèves ne sont pas affectés de la même manière par le redoublement. Ainsi, dans nombre de systèmes éducatifs, les élèves provenant de milieux socioéconomiquement défavorisés sont plus nombreux, comparativement aux autres, à connaître le redoublement. Mais ceci est-il injuste ou inéquitable ? Le redoublement vient en effet sanctionner des faiblesses ou difficultés scolaires dont on sait qu'elles sont liées à l'origine socioéconomique. La base de données PISA comporte des informations relatives au retard scolaire, une mesure précise du statut économique, social et culturel des élèves, ainsi qu'une estimation de leurs compétences en compréhension de l'écrit, en culture mathématique et en culture scientifique. Il est dès lors aisé de déterminer l'influence de l'origine sociale sur le risque de redoublement, à compétences égales. En d'autres termes, cette analyse revient à déterminer si des élèves ayant des performances identiques dans PISA, mais issus de milieux sociaux contrastés, ont des probabilités différentes de présenter un retard scolaire.

Des analyses réalisées sur PISA 2012 (Monseur et Lafontaine, 2012) montrent qu'à performances égales dans PISA, plus le niveau socioéconomique d'un élève est faible, plus sa probabilité d'avoir redoublé est élevée. Autrement dit, **un élève issu d'un milieu socioéconomiquement défavorisé aura plus de risques d'être en retard dans son parcours scolaire qu'un élève issu d'un milieu socioéconomique moyen ou favorisé ayant pourtant le même niveau de performances**. Une telle relation s'observe dans la moitié des pays de l'OCDE (17 pays sur 34).

Une analyse plus précise a été réalisée en prenant uniquement en considération les systèmes éducatifs qui comptent plus de 10 % d'élèves en retard scolaire à 15 ans. Le tableau 2 ci-dessous présente le lien entre l'origine sociale de l'élève et le retard scolaire sous la forme de risque relatif (*odds ratio*¹⁴) d'abord de manière brute (colonne SES) ensuite sous contrôle de la performance. L'indice composite PISA de statut social, économique et culturel¹⁵ a été recodé en une variable dichotomique : au sein de chacun des pays, les 25 % d'élèves les plus défavorisés sont ainsi opposés aux 75 % d'élèves les plus favorisés. L'*odds ratio* permet de décrire l'intensité du lien entre l'origine socioéconomique et le retard. Plus cet indice s'écarte de 1, plus le lien avec l'origine socioéconomique sera élevé. Par ailleurs, il s'interprète sous la forme d'un rapport de chances. Ainsi, en Allemagne, les chances d'un élève issu d'un milieu défavorisé d'être en retard scolaire plutôt qu'à l'heure dans son parcours sont 2,20 fois plus élevées que celles d'un de ses condisciples de classe sociale moyenne ou favorisée.

14. Estimé au départ d'une régression logistique.

15. La variable ESCS de la base de données PISA a été utilisée comme indicateur du statut économique, social et culturel de l'élève.

Tableau 2 – Influence (sous la forme d'*Odds ratio*) de l'origine sociale (SES) sur le redoublement, avant et après contrôle de la performance scolaire (seules les valeurs statistiquement significatives à 0,05 sont reprises dans le tableau)

Pays	SES	SES sous contrôle de la performance scolaire
Allemagne	2,20	1,28
Autriche	1,65	n.s.
Belgique	2,30	1,14
Espagne	2,82	1,73
États-Unis	2,71	1,52
France	3,05	1,51
Hongrie	5,09	2,12
Irlande	1,59	n.s.
Italie	1,82	n.s.
Luxembourg	2,33	n.s.
Pays-Bas	1,40	n.s.
Portugal	2,61	1,40
Suisse	1,76	n.s.

Comme tous les *odds ratio* diffèrent statistiquement de zéro, le lien entre origine sociale et redoublement est donc confirmé dans les systèmes éducatifs où le redoublement est une pratique répandue. Quel que soit le système éducatif, les élèves de milieux modestes redoublent plus souvent. Les valeurs de l'*odds ratio* oscillent néanmoins entre 1,40 aux Pays-Bas et 5,09 en Hongrie ; les inégalités sociales face au redoublement sont donc nettement plus marquées dans certains systèmes éducatifs.

Quand on calcule (colonne 3), le lien entre l'origine sociale et le retard à performances égales (à savoir sous contrôle de la performance scolaire dans PISA)¹⁶, le tableau se modifie sensiblement. Dans certains systèmes éducatifs, le lien entre l'origine sociale et le retard scolaire n'est plus significatif (n.s.) dès lors que l'on tient sous contrôle la performance scolaire. Dans ces systèmes, si les élèves de milieu modeste sont davantage frappés par le redoublement, c'est d'abord parce que leurs performances scolaires sont moins bonnes que celles de leurs pairs de milieux plus aisés. En revanche, en Allemagne, Belgique, Espagne, France, Hongrie, aux États-Unis et au Portugal, la propension des élèves d'origine défavorisée à redoubler davantage tient manifestement à d'autres facteurs que leurs performances scolaires ; on ne peut s'empêcher de considérer cette inégalité "ajoutée" comme une forme d'injustice. Dans ces systèmes éducatifs, le redoublement est un mécanisme amplificateur des inégalités sociales.

Monseur et Lafontaine (2009) ont par ailleurs montré qu'en Communauté française de Belgique¹⁷, 35 % des différences entre écoles - très marquées dans ce système éducatif - s'expliquaient par le retard scolaire.

16. La performance en compréhension de l'écrit a été introduite dans la régression logistique.

17. La Communauté française de Belgique est l'une des trois communautés fédérées de la Belgique. Depuis le 25 mai 2011, sans que sa dénomination constitutionnelle ait été changée, elle se désigne elle-même sous le nom de Fédération Wallonie-Bruxelles. Pour des raisons d'intelligibilité, nous avons conservé l'appellation Communauté française de Belgique. La Communauté française de Belgique est un système éducatif à part entière, géré en toute autonomie.

Alors que seulement 3 % des différences entre écoles s'expliquent directement par la composition socioéconomoculturelle du public de l'école, 16 % des différences entre écoles sont de la variance jointe entre la composition du public et le retard scolaire. En d'autres termes, les écoles qui accueillent une proportion importante d'élèves en retard scolaire accueillent aussi un public plus défavorisé socialement, et inversement. Un récent rapport de l'OCDE consacré aux élèves présentant de faibles performances en mathématiques (OCDE, 2016) montre aussi à quel point les variables de parcours (avoir ou non fréquenté l'enseignement maternel, avoir redoublé, fréquenter une filière professionnelle) contribuent d'une part à accroître les risques de se retrouver parmi les élèves dont les performances en mathématiques sont très faibles, d'autre part à accentuer les inégalités sociales. Des nuances existent bien entendu par pays, mais la Belgique et la France notamment font partie de ces pays où les écarts en fonction de l'origine sociale sont au départ assez faibles : ce sont essentiellement les différences de parcours mentionnées ci-avant qui accentuent les inégalités de départ liées à l'origine sociale.

II La différenciation horizontale : effets du regroupement des élèves par aptitude sur les acquis scolaires¹⁸

Le regroupement des élèves par aptitude obéit à la même logique de séparation que le redoublement. Il s'agit de séparer les élèves de leur cohorte ou d'éclater cette cohorte en groupes distincts pour leur proposer un enseignement en principe mieux adapté à leurs caractéristiques, leurs besoins ou leur rythme d'apprentissage. La réduction de l'hétérogénéité des groupes d'apprentissage qui résulte de cet "écrémage" apparaît souvent, aux yeux des acteurs de terrain, comme une condition nécessaire à la mise en place d'une pédagogie adaptée aux besoins de chacun. Or, selon Oakes *et al.* (1992), la justification explicite au regroupement par aptitude - la diminution de l'hétérogénéité des groupes - est loin d'être atteinte dans les faits : "des études mettent en évidence un recouvrement considérable entre les savoir-faire et les habiletés des élèves des différents groupes. Ainsi, la mesure dans laquelle le regroupement par niveau réduit l'hétérogénéité est peut-être bien moindre que ce que l'on prétend habituellement" (Oakes *et al.*, 1992, p. 594)¹⁹. En dépit de l'existence de classes de niveau ou de filières, l'hétérogénéité des classes reste importante, trop importante sans doute aux yeux de nombreux enseignants pour que ceux-ci mettent en place les formes de différenciation pédagogiques réputées efficaces.

Comme pour le redoublement, nombre d'acteurs du monde éducatif - enseignants et parents au premier chef - sont convaincus des vertus du regroupement des élèves par aptitude : non seulement celui-ci serait garant d'efficacité, mais aussi nécessaire - enseigner à des élèves aux aptitudes trop diverses relèverait de la mission impossible. Ces croyances, moins étudiées que celles relatives au redoublement, trouvent une expression emblématique dans la crainte du "nivellement par le bas" : pour nombre d'acteurs, il est évident que si on éduque tous les élèves ensemble, sans filières, sans classes de niveau, les plus faibles tireront les meilleurs vers le bas. Nombreux sont ceux qui le pensent. On verra que cette crainte est en grande partie non fondée, en examinant différents types d'études menées sur la question.

18. Les effets sur les variables non cognitives telles que le concept de soi ou les aspirations d'études seront envisagés plus loin, dans une section à part.

19. "Studies document considerable overlap of students' skills and abilities among groups. Thus, the degree to which tracking reduces heterogeneity maybe far less than we typically assume", (Oakes *et al.*, 1992, p. 594).

La question des effets du regroupement des élèves par aptitude sera abordée sous différents angles :

- comme pour le redoublement, il existe de très nombreuses **études de type (quasi) expérimental** qui comparent l'évolution d'élèves de niveau comparables, placés dans des classes de niveau ou des classes hétérogènes et qui essaient de contrôler d'autres paramètres qui sont sur le terrain étroitement liés à la constitution de classes de niveau, tels que le rythme auquel on avance dans la matière et le volume de matière couvert ;
- des études et des observations menées en **milieu naturel** qui permettent d'appréhender le phénomène du regroupement par aptitude dans toute son amplitude et sa complexité.

Nous examinerons dans un premier temps les effets des **classes de niveau** : par classes de niveau, on entend le fait de regrouper les élèves d'une manière permanente (au moins pour un an, mais il est rare que la composition des classes soit revue après un an ; une fois mise en place, elle a tendance à perdurer), mais sans distinction de parcours ou de curriculum (dans l'enseignement primaire par exemple). La répartition des élèves en classes de niveau, contrairement aux filières, s'opère en général de manière informelle et peu transparente, sans sélection ou orientation explicite. Les classes de niveau se distinguent des **groupes de niveau**. Ceux-ci ont un caractère temporaire, le groupe-classe de référence de l'élève restant hétérogène, et ne concernent qu'un nombre limité de disciplines. Enfin, les classes de niveau se distinguent des **filières** qui proposent dans le secondaire un curriculum distinct à orientation générale, technique, artistique, professionnelle, à l'issue d'une sélection ou orientation explicite. Les filières, en principe, ne correspondent pas nécessairement à un regroupement par aptitude, elles peuvent au moins partiellement relever d'un choix d'études. Toutefois, dans nombre de systèmes éducatifs, en particulier ceux où l'orientation est précoce, cette orientation se fait bel et bien sur des critères de réussite scolaire, donc d'aptitudes.

1 Les études quasi expérimentales

Nous passerons rapidement sur les résultats de ces études (quasi)-expérimentales, en nous limitant à rapporter les principales conclusions des méta-analyses consacrées à la question. Comme signalé ci-avant, ces études, dans l'optique d'isoler l'impact du mode de regroupement par aptitude *per se*, bloquent les mécanismes d'adaptation/différenciation qui font précisément l'intérêt du regroupement par aptitude aux yeux de ses défenseurs. Ainsi, dans ces études expérimentales, pour en prendre l'exemple le plus éclairant, avec les classes composées d'élèves forts, les enseignants ne peuvent pas avancer plus vite ou couvrir davantage de matière que dans les classes d'élèves moyens ou faibles. Ces recherches permettent donc d'apporter une réponse rigoureuse, mais assez artificielle à la question du regroupement par aptitude. Elles fixent une balise théorique, dont la validité écologique est toutefois limitée. Sur le terrain, le regroupement par aptitude va nécessairement de pair avec des différences dans l'approfondissement ou le rythme auquel est vue la matière, ainsi que dans les attentes implicites ou explicites des enseignants.

Les principales méta-analyses portant sur la question sont celles qu'y a consacrées **Slavin (1987a)** dans l'enseignement primaire et secondaire (**Slavin, 1990**)²⁰. La méta-analyse des études menées dans le primaire (**Slavin, 1987a**) conclut à l'absence d'effet spécifique du regroupement par aptitude (moyenne de l'ampleur de l'effet = 0). Les résultats des études originales reprises dans la méta-analyse vont en sens

20. Pour plus de détails sur ces études, voir **Crahay et Wanlin (2012)**.

divers - effets positifs, négatifs ou nuls selon les études, mais dans tous les cas les effets sont limités, bien en-deçà du seuil de 0,20 considéré comme devant être atteint pour conclure à un effet significatif. Les résultats dans le secondaire (Slavin, 1990) vont dans le même sens : l'ampleur de l'effet moyenne est de -0,02. Contrairement à ce que l'on aurait pu attendre, le regroupement par aptitude ne s'avère ni plus positif ni plus négatif pour les élèves en fonction de leur aptitude de départ, pas plus dans le primaire que dans le secondaire. En d'autres termes, le regroupement par aptitude n'est pas davantage bénéfique pour les élèves forts et davantage négatif pour les élèves faibles. Dans certaines études, c'est même l'inverse qui est observé. À nouveau, les effets vont en sens divers selon les études et sont d'ampleur limitée. La conclusion s'impose : **l'effet spécifique du regroupement des élèves en classes de niveau est nul. Si la qualité et la quantité de l'enseignement dispensé sont maintenues constantes, le mode de regroupement en soi n'affecte pas la qualité des apprentissages** ; les élèves forts ne pâtissent pas de la compagnie d'élèves moyens ou faibles.

Les méta-analyses que synthétise Hattie dans son ouvrage (Hattie, 2009) recensent au total plus de 300 études portant sur le regroupement par classes de niveau menées dans des contextes et cultures divers, à différents niveaux scolaires, et dans différentes matières. L'ampleur de l'effet moyenne sur les acquis scolaires est de 0,11 (effet considéré comme insignifiant). L'effet est de 0,14 pour les classes d'élèves forts, de -0,03 pour les classes d'élèves moyens et de 0,09 pour les classes d'élèves plus faibles. "Personne n'en profite" en conclut Hattie. Les effets ne varient pas non plus en fonction des disciplines.

Une étude expérimentale menée dans l'enseignement primaire au Kenya par des économistes de l'éducation (Duflo *et al.*, 2011) mérite que l'on s'y attarde. Cette étude a fait pas mal de bruit, notamment parce qu'elle débouchait sur des résultats peu concordants avec la littérature de recherche ; elle a aussi le mérite d'avoir été menée dans un contexte pour lequel il existe moins de résultats de recherche. Cette étude expérimentale montre qu'après 18 mois, les scores sont plus élevés de 0,14 écart type dans les écoles avec classes de niveau que dans les écoles sans regroupement par niveau. Pour cette étude, dans 60 écoles choisies aléatoirement, les élèves avaient été assignés à une classe de niveau (classe regroupant des élèves forts vs des élèves plus faibles) en fonction de leur niveau initial (groupe expérimental) tandis que dans les 60 écoles du groupe contrôle, les élèves étaient assignés au hasard à l'une des deux classes. Il est à noter que les auteurs n'ont pas cherché à contrôler les caractéristiques de l'enseignement, et "même s'ils ne disposent pas d'observations sur le niveau d'enseignement et la manière dont il varie selon les écoles et les sections dans l'expérience, nous disposons d'indices probants selon lesquels le comportement des enseignants est affecté par l'existence de classes de niveau" (Duflo *et al.*, 2011, p. 1741). Ainsi, les enseignants étaient moins souvent absents dans les écoles avec classes de niveau, en particulier dans les classes d'élèves forts. Quand il existe des classes de niveau, les élèves les plus faibles améliorent essentiellement leurs habiletés de base, tandis que les plus forts enregistrent des gains pour les habiletés plus avancées, ce qui "est cohérent avec l'hypothèse selon laquelle les enseignants ajustent leur enseignement en fonction de la composition de la classe" (Duflo *et al.*, 2011, p. 1741). En outre, les enseignants reçoivent des incitants sur la base des scores enregistrés par leurs élèves à l'examen national de fin d'enseignement primaire (grade 8), ce qui pousse tous les enseignants à faire progresser leurs élèves autant que possible.

À certains égards, en ne contrôlant pas la qualité et la quantité de l'enseignement dispensé, cette étude expérimentale se rapproche des études menées en milieu naturel, et il n'est donc en rien surprenant que ses résultats se rapprochent de ceux de ces dernières. La présence d'incitants pour les enseignants liés à un test

final à enjeux élevés donne aussi un autre horizon au regroupement en classes de niveau, empêchant ou limitant la tendance à fixer des objectifs moins ambitieux pour les classes composées d'élèves plus faibles. Somme toute, cette étude, qui débouche sur des résultats plutôt favorables aux classes de niveau, porte sur une combinatoire de classes de niveau incitantes à faire réussir une épreuve à enjeux élevés. Le titre de l'article ne laisse d'ailleurs planer aucune ambiguïté à ce propos, pas plus que les conclusions prudentes de ses auteurs.

Avant d'en venir à la présentation des études en milieu naturel, envisageons les résultats des études expérimentales portant sur l'effet des groupes de niveau à l'intérieur des classes (*within-class grouping*).

Slavin (1987b) a aussi consacré une méta-analyse à cette question, dans l'enseignement primaire, en lecture et en mathématiques. Les modalités de groupement à l'intérieur des classes sont les suivantes : le temps passé en groupes homogènes doit être nettement inférieur au temps passé en classes hétérogènes, ce qui a pour conséquence que le groupe auquel les enfants s'identifient le plus est la classe hétérogène. La constitution des groupes homogènes repose sur l'évaluation d'une compétence spécifique et non pas sur celle d'une aptitude générale ; enfin, les groupes sont flexibles : en fonction des progrès des élèves, les groupes peuvent être réorganisés.

L'amplitude de l'effet dans le domaine de la lecture est de +0,45, et de +0,35 dans le domaine des mathématiques, ce qui correspond à un effet relativement important. Aucune étude n'affiche de résultats négatifs. Dans trois des cinq études qui ont vérifié quels élèves bénéficiaient le plus de cette forme de groupement, ce sont les élèves faibles qui en profitent le plus ; toutefois, dans l'ensemble, tous les types d'élèves y gagnent.

L'organisation flexible au sein de classes hétérogènes de groupes homogènes constitués en fonction du niveau de maîtrise d'une compétence spécifique produit donc d'indéniables effets positifs, quel que soit le niveau d'aptitude des élèves, et en particulier pour les plus faibles.

Au vu de la littérature expérimentale, pour répondre aux besoins des élèves et à leurs différences de rythme, **il apparaît donc nettement préférable de constituer des groupes de niveau ou de besoin temporaires et flexibles que de séparer en classes distinctes des élèves réputés forts, moyens ou faibles.** Non seulement cette modalité est de loin plus efficace, mais vu son caractère temporaire, elle n'a aucun impact sur le parcours scolaire ultérieur de l'élève, à la différence des classes de niveau.

2 Les études en milieu naturel

Les études menées en milieu naturel débouchent sur des résultats sensiblement différents de ceux des études expérimentales et quasi-expérimentales. Ces études comparent la progression d'élèves de même niveau dont les uns se retrouvent dans des classes fortes à celles d'élèves placés dans des classes plus faibles. Dans ce cas, l'enseignement (qualité et quantité de matière) peut évidemment varier et varie effectivement d'une classe à l'autre. Sans surprise, la progression est en général nettement meilleure pour un élève quand il est placé dans une classe plus forte (Gamoran, 2002 ; Kerckhoff, 1986).

Ceci s'explique aisément. Un ensemble de paramètres autres que la composition homogène/hétérogène de la classe varient avec cette composition. Ainsi, dans les classes plus fortes, le curriculum réalisé est plus poussé, les attentes des enseignants ainsi que le temps consacré aux apprentissages sont plus élevés, les professeurs sont même parfois plus qualifiés ; une classe d'élèves plus forts et plus motivés produit aussi un effet d'émulation. Dans les classes plus faibles, les professeurs sont plus fatalistes, les exercices de

répétition mécanique sont plus nombreux ; obtenir la participation, motiver les élèves, prévenir le chahut ou les débordements prend parfois le pas sur les objectifs d'apprentissage. Bref, on retrouve beaucoup plus souvent des facteurs garants d'efficacité et de progrès pédagogique dans les classes les plus fortes. La composition de la classe et les effets de pairs (dynamique plus ou moins favorable aux apprentissages) qui y sont liés agissent donc de façon directe et indirecte sur les acquis scolaires, via la manière dont l'enseignant ajuste son enseignement en fonction de la composition de la classe. Comme le soulignent *Oakes et al. (1992)*, "les chercheurs et les décideurs sont de plus en plus perturbés par les effets du regroupement par aptitude sur l'accès des élèves aux occasions d'apprendre : le choix des cours, l'exposition au contenu du curriculum, les stratégies d'enseignement, les environnements de classe et les relations avec les pairs sont influencés par ce regroupement - en sorte que les programmes les plus pauvres sur le plan éducatif sont offerts aux étudiants les moins capables, en particulier dans l'enseignement secondaire" (p. 596)²¹. Les rares exemples où le regroupement des élèves en classes de niveau ne se révèle pas négatif pour les élèves fréquentant les classes les plus faibles sont des classes où l'on arrive à maintenir ce qui fait l'efficacité des classes les plus fortes : les mêmes enseignants enseignent aux différentes classes de niveau, ils maintiennent des exigences élevées pour tous, se refusent à abaisser les exigences académiques et mettent en avant les discussions orales plutôt que les exercices écrits répétitifs²² (*Gamoran, 2002*).

En bref, en milieu naturel, les politiques de regroupement par aptitude, si elles peuvent se révéler efficaces pour les apprentissages des élèves placés dans des classes fortes, ont des conséquences négatives pour les élèves placés dans des groupes moins forts. Si certains acteurs peuvent trouver un intérêt aux classes de niveau (élèves, parents, enseignants des classes fortes), les décideurs en matière de politiques éducatives soucieux du bien commun doivent peser les avantages de telles politiques. À l'échelle d'un système éducatif, les faibles bénéfices enregistrés par les élèves forts ne suffisent en général pas à compenser les désavantages nettement plus marqués observés pour les élèves les plus faibles.

Par ailleurs, à l'aune de principes de justice et de critères d'équité, les politiques de regroupement en classes de niveau affichent un bilan désastreux. En offrant des environnements d'apprentissage de qualités inégales aux élèves de niveaux différents, celles-ci conduisent à creuser inéluctablement les écarts entre les élèves les plus et les moins performants, et à accroître substantiellement les inégalités liées à l'origine socioéconomique et culturelle des élèves. Les conséquences de ceci débordent largement la sphère des acquis scolaires, on l'aura compris.

Nous allons maintenant examiner les inégalités liées au regroupement par aptitude à la lumière des résultats des enquêtes internationales. Dans ces études, le regroupement par aptitude inclut non seulement les classes de niveau, mais aussi la répartition en filières distinctes.

21. "Moreover, scholars and policymakers have become increasingly troubled by the effects of tracking on students' access to learning opportunities : course-taking patterns, exposure to curriculum content, instructional strategies, classroom environments, and peer relationships all appear to be influenced by tracking - so that the most educationally impoverished programs are offered to the least able students, especially at the secondary level."

22. "the same teachers taught both high- and low level classes ; teachers held high expectations for low achieving students, manifested in a refusal to relinquish or dilute the academic curriculum ; and teachers made extra efforts to foster oral discourse with low-achieving students (*Gamoran, 2002*, p. 3).

III L'impact des mécanismes de différenciation sur les performances et les inégalités scolaires et sociales : l'éclairage des enquêtes internationales

Les enquêtes internationales menées à l'initiative de l'IEA (TIMSS, PIRLS) et de l'OCDE (PISA) constituent une source particulièrement utile pour étudier le lien entre les structures des systèmes éducatifs, leur efficacité (performances) et leur équité (écarts entre les élèves les plus et les moins performants, proportion d'élèves aux performances faibles, écarts liés à l'origine sociale et culturelle). À la différence des études qui ont été présentées jusqu'ici, les recherches menées au départ des bases de données internationales permettent d'envisager des modalités d'organisation ou de différenciation qui font système et pour lesquels différentes formes de typologie ont été proposées (Mons, 2009 ; Trautwein *et al.*, 2006). Cette section se penchera en particulier sur la question de l'impact du degré de différenciation sur les inégalités sociales.

1 Regroupement des élèves par niveau, performances, inégalités scolaires et sociales²³

A Importance du regroupement par niveau

Les pratiques de regroupement homogène des élèves constituent un mécanisme potentiel d'amplification des inégalités sociales. Cette question ne date pas d'hier, comme en témoignent les analyses menées par Husén (1967) sur les comparaisons entre systèmes sélectifs et systèmes compréhensifs au départ des données de la première étude internationale de l'IEA en mathématiques ou les analyses menées plus récemment par Hanushek et Woessmann (2010) ou par l'OCDE (2011a,b) pour tenter d'identifier les caractéristiques des systèmes éducatifs associées à une plus grande efficacité et à une plus grande équité.

Il existe en effet différentes stratégies pour réduire les différences de niveau ou l'hétérogénéité entre élèves au sein des classes et des écoles ; le redoublement ainsi que les filières d'enseignement relèvent sans conteste de ces stratégies. Parmi les autres stratégies, on peut citer l'enseignement spécial, le regroupement homogène des élèves dans des écoles distinctes ainsi que l'organisation de classes de niveau dans les écoles. Ces différentes stratégies ne sont pas mutuellement exclusives et peuvent bien entendu coexister.

Il n'est pas aisé de trouver un seul indice statistique susceptible de capturer au mieux ces différentes stratégies de regroupement homogène des élèves. Le pourcentage de la variation des performances des élèves qui se situe entre classes constitue, à nos yeux, l'indice qui y parvient le mieux. Cependant, cet indice nécessite que des classes soient sélectionnées au sein des écoles, comme dans les études TIMSS et PIRLS de l'IEA, et non des élèves de 15 ans à travers classes et à travers années d'étude, comme dans l'enquête PISA. Dans le cadre des analyses présentées ici, qui se basent presque exclusivement sur les études PISA, nous avons donc choisi, à défaut, comme indicateurs de l'intensité des pratiques de regroupement homogène des élèves les deux indices suivants :

1. Le premier indice correspond au pourcentage de la variation de la performance des élèves qui se situe entre écoles. Plus cet indice tend vers zéro, plus les performances moyennes des élèves sont semblables d'un établissement à l'autre. Par contre, plus cet indice s'approche de 1 (ou de 100 selon qu'on l'exprime ou non en %), plus les performances moyennes des élèves fluctuent d'une école à l'autre. Cet indice ne peut malheureusement pas détecter les pratiques de regroupement homogène

23. Cette section reprend de larges extraits d'un chapitre d'ouvrage rédigé par Monsieur et Lafontaine (2012).

en classes de niveau au sein des écoles. De même, il ne pourra pas refléter le regroupement homogène des élèves qu'opèrent les filières, si différentes filières d'enseignement sont organisées dans chaque établissement.

2. Le second indice correspond au pourcentage de la variation de l'origine sociale des élèves qui se situe entre écoles²⁴. Il permet d'estimer dans quelle mesure un système éducatif tend à concentrer les élèves en fonction de leur origine sociale dans certains établissements, autrement dit, cet indice est une mesure de la ségrégation sociale des établissements ou encore de la ghettoïsation - quels qu'en soient les motifs.

Dans un premier temps, nous décrivons pour les différents pays retenus l'importance des pratiques de regroupement homogène des élèves au départ des bases de données de l'IEA et de PISA.

Ensuite, nous examinerons le lien entre le regroupement homogène des élèves selon des critères académiques et/ou sociaux et :

1. les performances moyennes ;
2. la différence de performances entre les élèves faibles et les élèves forts ;
3. la corrélation entre l'origine sociale de l'élève et sa performance scolaire ;
4. la corrélation entre l'origine sociale de l'élève et ses aspirations académiques et professionnelles.

Ces hypothèses ont été mises à l'épreuve au départ des bases de données PISA 2003, PISA 2006 et PISA 2009.

Le pourcentage de variation de la performance entre écoles a été calculé au départ des bases de données PIRLS 2006, TIMSS 1995 et de PISA. Les valeurs de cet indice sont reprises dans le tableau 3, par pays et par étude.

L'indice calculé au départ des études PIRLS reflète l'importance des différences de performance en compréhension de l'écrit entre classes de 4^e année primaire²⁵. À une ou deux exceptions près, le pourcentage de variance qui se situe entre classes reste sensiblement inférieur à 30 %, témoignant ainsi d'un recours modéré aux pratiques de regroupement homogène des élèves dans l'enseignement primaire. Soulignons toutefois l'ampleur des variations entre l'Islande (9 %), la Suède (14 %) ou la Norvège (15 %) d'une part, et l'Allemagne (30 %), la Hongrie (29 %), la Nouvelle-Zélande (39 %) et la Slovaquie (36 %) d'autre part. Les indices calculés à partir des données PISA montrent une plus grande variabilité. Ainsi, les pays du nord de l'Europe présentent une variance entre écoles inférieure à 20 %. Ce pourcentage, dans les six pays de culture anglo-saxonne, oscille entre 19 % (Canada) et 25 % (Irlande). À l'aune de PISA, il apparaît que les pays scandinaves et, dans une moindre mesure, les pays anglo-saxons ne recourent pas abondamment aux pratiques de regroupement homogène entre écoles. Enfin, neuf pays ont un pourcentage de variance entre écoles supérieur à 50 % : la République tchèque, les deux communautés belges, l'Italie, l'Autriche, l'Allemagne, la France, la Hongrie et les Pays-Bas. Dans l'interprétation des pourcentages calculés au départ des données PISA, il convient toutefois de se rappeler que cette étude évalue des élèves de 15 ans, quelle que

24. La variable HISEI (i.e. l'indice socioéconomique le plus élevé des deux parents) dans les bases de données a été utilisée comme indice socio-économique.

25. Dans PIRLS, selon les pays, une ou deux classes ont été sélectionnées. L'indice calculé au départ de PIRLS confond donc la variance entre écoles et la variance entre classes à l'intérieur des écoles. Il n'est dès lors pas possible de distinguer regroupement homogène des élèves entre écoles ou entre classes au sein des écoles.

soit l'année d'étude (*grade*) qu'ils fréquentent. Ainsi, en France, les élèves de 15 ans se répartissent entre les collèges (élèves en retard) et les lycées (élèves à l'heure). La distinction entre enseignement secondaire inférieur et enseignement secondaire supérieur contribue donc en France à la variance entre écoles, alors qu'elle contribue à la variance à l'intérieur des écoles dans d'autres pays comme la Belgique.

L'étude TIMSS 1995²⁶ apporte quelques nuances aux constats établis au départ de PISA. En effet, le plan d'échantillonnage adopté dans cette étude permet de distinguer la variance entre écoles et la variance entre classes au sein des écoles : les pratiques de regroupement homogène entre classes ne sont pas dans ce cas confondues avec les pratiques de regroupement homogène entre écoles. En début de scolarité secondaire, quatre pays se caractérisent par un pourcentage de variance entre écoles supérieur à 40 % : l'Allemagne, l'Autriche, la Suisse et les Pays-Bas. Pour tous les autres pays, le pourcentage de variance entre écoles est inférieur ou proche de 30 %. Par ailleurs, l'Irlande, l'Australie, les États-Unis, les deux communautés belges et la Suède présentent un pourcentage de variance entre classes à l'intérieur des écoles supérieur ou égal à 20 %, témoignant ainsi d'un recours au regroupement homogène des élèves sous la forme de classes de niveau.

La dernière colonne du tableau 3 présente les pourcentages de variation entre écoles de l'origine sociale calculés au départ de PISA 2009. Alors que les pourcentages de la variation des performances entre écoles et entre classes constituent des indicateurs de la ségrégation académique, le pourcentage de la variation de l'origine sociale présenté dans la dernière colonne constitue un indicateur du degré de ségrégation sociale d'un système éducatif, en d'autres mots de sa propension à la ghettoïsation des établissements. À cet égard, on peut se demander s'il existe un lien entre la tendance d'un système éducatif à regrouper les élèves selon leurs aptitudes académiques et la tendance à les regrouper selon leur origine sociale. La réponse à cette question est sans ambiguïté. Une corrélation très élevée est observée entre les deux phénomènes (0,72 dans PISA 2009). **Plus un système éducatif regroupe les élèves selon des critères académiques, plus il les regroupe selon des critères sociaux et inversement.** On notera toutefois que **le regroupement en fonction de l'origine sociale est moins marqué que le regroupement en fonction de critères académiques, ceci d'une manière systématique dans les pays qui pratiquent abondamment le regroupement par aptitudes académiques.**

Soulignons que lorsque nous évoquons des regroupements en fonction de critères académiques ou sociaux, il s'agit de constats saisis au travers d'indicateurs objectivés. Ceci ne signifie pas que les systèmes éducatifs ont mis en place délibérément des politiques de regroupement par niveau et encore moins des politiques de ségrégation sociale qui empêcheraient l'accès de certaines catégories d'élèves à certains établissements. Ces indicateurs permettent juste de dresser un état du degré de ségrégation académique et sociale ; celui-ci résulte d'une constellation de facteurs qui peuvent être externes au système éducatif (tels que la ségrégation résidentielle en milieu urbain) ou internes (par exemple : les filières s'organisent ou non dans des établissements distincts, le choix d'école est ou non réglementé, des politiques de quota existent, etc.).

26. Cette étude a en effet évalué la performance des élèves de deux années d'étude consécutives, dénommées *grade* inférieur et *grade* supérieur. O'Dwyer (2002) propose d'augmenter la performance de toutes les classes du *grade* inférieur par le gain moyen observé entre les deux années d'études et de considérer ces classes du *grade* inférieur comme de pseudo-classes du *grade* supérieur. Cette procédure, validée par O'Dwyer (2002), permet ainsi d'estimer le pourcentage de variance qui se situe entre écoles et le pourcentage de variance qui se situe entre classes à l'intérieur des écoles.

B Regroupement homogène et performances

Les pratiques de regroupement homogène des élèves sont-elles associées à de meilleures performances ? Comme le montre sans aucune ambiguïté le tableau ci-dessous, il n'en est rien. C'est tout le contraire qui est observé : des corrélations négatives sont observées dans tous les cycles et les domaines entre le pourcentage de variation de la performance entre écoles et le niveau de performances. Selon les cas, ces corrélations sont faibles à modérées. Autrement dit, **les systèmes éducatifs qui regroupent les élèves dans des écoles de niveaux différents sont en moyenne moins performants que ceux évitant ce type de regroupement**. Les corrélations entre les performances et le degré de regroupement en fonction de l'origine sociale sont elles aussi négatives, et plus marquées.

C Regroupement homogène des élèves et écarts de performances

Les pratiques de regroupement homogène des élèves sont-elles associées à une plus grande dispersion des résultats scolaires ? En d'autres termes, les pays qui regroupent les élèves selon des critères académiques ou sociaux se caractérisent-ils par de plus grandes différences de performances entre les élèves les plus faibles et les élèves les plus forts ? Le tableau 5 présente, pour les trois enquêtes PISA retenues, la corrélation entre les pourcentages de variation entre écoles (indicateurs de ségrégation académique et sociale) et la mesure de la dispersion des performances (écart-type du pays pour les trois domaines évalués).

En mathématiques, la corrélation entre regroupement homogène des élèves selon des critères académiques et dispersion des performances oscille, pour les trois cycles d'études, entre 0,55 et 0,67. **Plus un pays regroupe les élèves selon des critères académiques, plus la différence de performance entre élèves faibles et élèves forts est élevée**. En compréhension de l'écrit, les corrélations entre regroupement homogène des élèves selon des critères académiques et dispersion oscillent entre 0,27 et 0,56 ; pour les sciences, ils vont de -0,14 à 0,29.

Peut-être faut-il voir dans ces variations le reflet de l'importance relative de ces trois domaines dans le cursus des élèves. Discipline scolaire par excellence, les mathématiques, dans les systèmes sélectifs, constituent un pilier de la réussite et de l'excellence scolaire. Un échec en mathématiques se traduit souvent par un redoublement et/ou une réorientation dans une filière académiquement moins exigeante dans les systèmes éducatifs qui ont adopté ces pratiques de gestion du parcours scolaire des élèves. Cet échec peut par ailleurs s'accompagner d'un changement d'école ([Communauté française, 2009](#)). Dans les établissements performants, l'enseignement des mathématiques peut s'accélérer alors qu'il se voit ralenti dans les écoles qui accueillent les élèves en difficulté. Par contre, comme en témoignent les grilles horaires ([Eurydice, 2011](#)), l'enseignement des sciences joue plutôt un rôle de second plan, ce qui pourrait éventuellement justifier la quasi-absence de lien entre regroupement homogène et variabilité de la performance pour cette discipline ; certains programmes peu exigeants sur le plan académique (filères techniques) comportent par ailleurs un nombre important d'heures de sciences, alors que ce n'est jamais le cas pour les mathématiques.

D Regroupement homogène des élèves et déterminisme social

La figure 1 présente la relation entre regroupement homogène des élèves selon des critères sociaux et le déterminisme social estimé par la corrélation entre la performance de l'élève en compréhension de l'écrit

et son origine sociale au départ de la base de données PISA 2009.

Figure 1 – Regroupement homogène des élèves selon des critères sociaux et déterminisme social

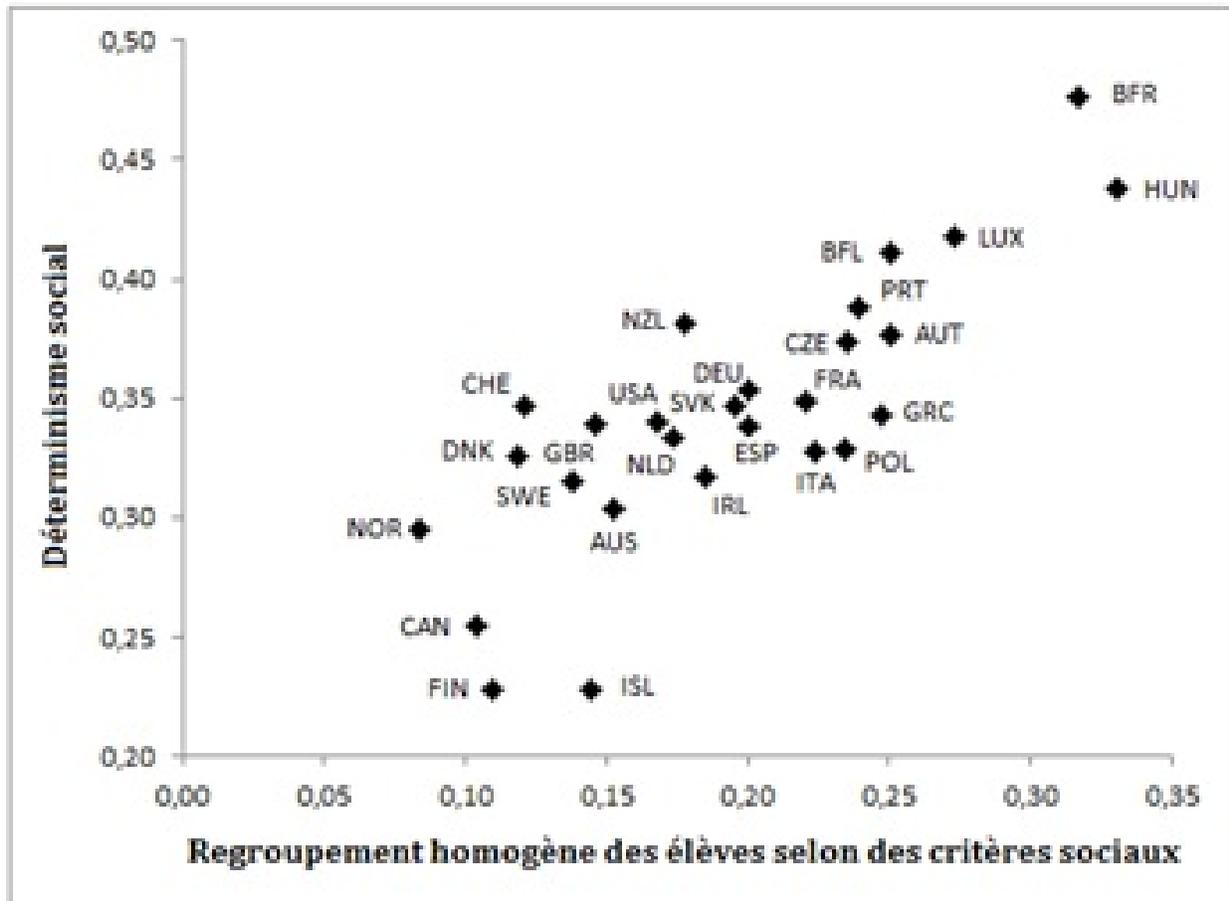


Tableau 3 – Pourcentage de variation de la performance entre écoles ou entre classes (Source : bases de données PIRLS 2006, TIMSS 1995 et PISA 2009) et de l'origine sociale des élèves (PISA 2009).

Pays	PIRLS 2006 Classes et écoles	TIMSS 1995 Ecoles	TIMSS 1995 Classes	PISA 2009 Lecture Écoles	PISA 2009 SES ^a Écoles
Allemagne	30	59	2	61	20
Australie		24	22	22	15
Autriche	18	42	3	57	25
Canada	17	19	7	19	10
Communauté flamande	14	30	22	55	25
Communauté française	26	24	20	57	32
Danemark	12	8	3	16	12
Espagne	20	21	3	19	20
États-Unis	24	31	23	24	17
Finlande				7	11
France	23	23	10	58	22
Grande-Bretagne	19	23	3	20	15
Grèce		15	3	42	25
Hongrie	29	22	4	69	33
Irlande		23	24	25	18
Islande	9	6	8	14	14
Italie	24	26	6	55	22
Luxembourg	17			34	27
Norvège	15			11	8
Nouvelle-Zélande	39			21	18
Pays-Bas	17	46	18	62	17
Pologne	13			16	23
Portugal		17	8	31	24
République Slovaque	36	14	10	45	20
République tchèque		21	11	54	24
Suède	14	1	28	16	14
Suisse		45	3	32	12

a. Dans cette dernière colonne, il s'agit du pourcentage de variation de l'origine sociale des élèves entre écoles.

Tableau 4 – Corrélation entre pourcentages de variation entre écoles et performances, par discipline et par cycle (Pisa 2003, Pisa 2006 et Pisa 2009).

		Compréhension de l'écrit	Mathématiques	Sciences
Pourcentage de variation de la performance entre écoles^a	PISA 2003	- 0,43	- 0,22	- 0,32
	PISA 2006	- 0,35	- 0,03	- 0,10
	PISA 2009	- 0,32	- 0,10	- 0,10
Pourcentage de variation de l'origine sociale entre écoles	PISA 2003	-0,60	-0,57	-0,50
	PISA 2006	-0,56	-0,43	-0,31
	PISA 2009	-0,61	-0,57	-0,51

a. Les indices de ségrégation académique sont basés sur le domaine majeur, à savoir les mathématiques en 2003, les sciences en 2006 et la lecture en 2009.

Tableau 5 – Corrélation entre pourcentages de variation entre écoles et dispersion de la performance, par discipline et par cycle (Pisa 2003, Pisa 2006 et Pisa 2009).

		Compréhension de l'écrit	Mathématiques	Sciences
Pourcentage de variation de la performance entre écoles	PISA 2003	0,30	0,55	0,29
	PISA 2006	0,56	0,67	0,06
	PISA 2009	0,27	0,59	0,28
Pourcentage de variation de l'origine sociale entre écoles	PISA 2003	0,32	0,28	0,06
	PISA 2006	0,45	0,55	-0,14
	PISA 2009	0,35	0,52	0,16

Comme le montre la figure 1, plus un système recourt à la ségrégation sociale, plus le poids de l'origine sociale de l'élève sur sa performance sera élevé. La Communauté française de Belgique affiche à cet égard la corrélation la plus élevée. Avec des pourcentages de variation entre écoles distincts, la Suisse et la France présentent un déterminisme social identique. Le tableau 6 reprend les différentes valeurs de la corrélation entre regroupement homogène (académique et social) et déterminisme social.

Tableau 6 – Corrélation entre ségrégation académique et sociale et déterminisme social.

	Pourcentage de la variation des performances entre écoles	Pourcentage de la variation de l'origine sociale entre écoles
PISA 2003	0,68	0,56
PISA 2006	0,67	0,69
PISA 2009	0,65	0,81

Les pratiques de regroupement homogène en fonction de l'origine sociale renforcent donc le poids de l'origine sociale, comme l'illustre la figure 1. Notons que Dupriez et Dumay (2006) ont de leur côté démontré que l'ampleur des inégalités sociales à l'école n'était pas corrélée aux inégalités socio-économiques dans les différents pays mesurées par les inégalités de revenus, montrant ainsi le rôle déterminant des structures scolaires sur les inégalités sociales.

E Regroupement homogène des élèves, aspirations d'études et professionnelles

Ce renforcement des inégalités sociales par le regroupement homogène produit des effets à long terme, au-delà de la scolarité obligatoire des élèves. Dans le cadre de l'enquête PISA 2003, les élèves ont été invités à indiquer le niveau d'études qu'ils voudraient atteindre et à décrire la profession qu'ils souhaiteraient exercer à l'âge de 30 ans. Comme l'indique le tableau 7, plus un pays regroupe les élèves selon leurs performances académiques et selon leur origine sociale, plus le poids de l'origine sociale de l'élève sur ses aspirations socioprofessionnelles sera élevé (corrélation de 0,68 et 0,70 respectivement). Ainsi, la corrélation entre le statut socio-économique des parents et aspirations professionnelles des élèves est modeste aux États-Unis (0,13) et en Australie (0,19), pays où le regroupement en fonction de l'origine sociale est modéré (voir dernière colonne du tableau 3 et figure 1). En revanche, la corrélation entre le statut socio-économique des parents et aspirations professionnelles des élèves s'élève à 0,39 en Belgique néerlandophone, à 0,34 en Belgique francophone et à 0,37 en Hongrie, systèmes éducatifs où le regroupement en fonction de l'origine sociale est particulièrement accusé.

De même, plus un pays regroupe les élèves selon des critères sociaux, plus cette origine sociale déterminera les projets d'études universitaires ou supérieures non universitaires des élèves. À titre d'exemple, la corrélation entre origine sociale et aspirations universitaires s'élève à 0,20 aux États-Unis et en Islande, et à 0,41 en Hongrie.

Tableau 7 – Corrélation entre ségrégation académique et sociale et déterminisme social des aspirations académiques et socioprofessionnelles

	Pourcentage de la variation des performances entre écoles	Pourcentage de la variation de l'origine sociale entre écoles
Corrélation SES et projet de faire des études supérieures	0,42	0,67
Corrélation SES et projet de faire des études universitaires	0,56	0,83
Corrélation SES et aspirations socioprofessionnelles	0,68	0,70

En résumé, les pratiques de regroupement homogène des élèves, qu'elles soient d'ordre académique ou social, amplifient les différences de performances et renforcent le poids de l'origine sociale sur la performance scolaire et les aspirations académiques et professionnelles des élèves.

Par quels mécanismes les pratiques de regroupement homogène des élèves accentuent-elles les inégalités scolaires et sociales ? Cette question a déjà été abordée en partie dans la section consacrée aux études en milieu naturel. Que peuvent apporter les études internationales à cette question ? Les systèmes éducatifs qui recourent aux pratiques de regroupement homogène des élèves se caractérisent-ils davantage par des phénomènes d'inégalités de traitement ou de discriminations négatives ?

2 Quand les politiques éducatives accentuent les différences dans les contextes d'apprentissage

Différents facteurs peuvent contribuer au regroupement homogène des élèves et aux différences de performances entre établissements qui en résultent. Parmi ceux-ci, on peut citer :

1. la diversité de l'offre d'enseignement, due notamment, mais pas exclusivement, à l'existence de filières d'enseignement ;
2. un recrutement social et/ou académique variable ;
3. des facteurs d'ordre pédagogique, qui peuvent concerner la qualité de l'enseignement dispensé, les caractéristiques des enseignants, une culture, un climat ou un projet pédagogique propres à chaque établissement.

Dans les pages qui suivent, nous allons nous attacher à essayer de déterminer dans quelle mesure, pour chaque système éducatif, les différences de performances entre établissements s'expliquent par des facteurs d'ordre pédagogique tels que la qualité/quantité d'enseignement ou par le recrutement social. Pour ce faire, dans les analyses ci-après, deux groupes de variables ont été retenus. Le premier groupe de variables concerne l'environnement familial. Plusieurs informations ont été intégrées pour décrire au mieux les diverses composantes de cet environnement : le niveau de formation des parents, leur profession, l'importance qu'occupe l'écrit dans l'environnement familial. Le second groupe de variables décrit la qualité du climat scolaire de l'école, en synthétisant des informations recueillies auprès des élèves et auprès des directeurs d'écoles relatives aux problèmes disciplinaires, aux relations entre élèves et professeurs, au bien-être que

l'élève ressent au sein de son école²⁷.

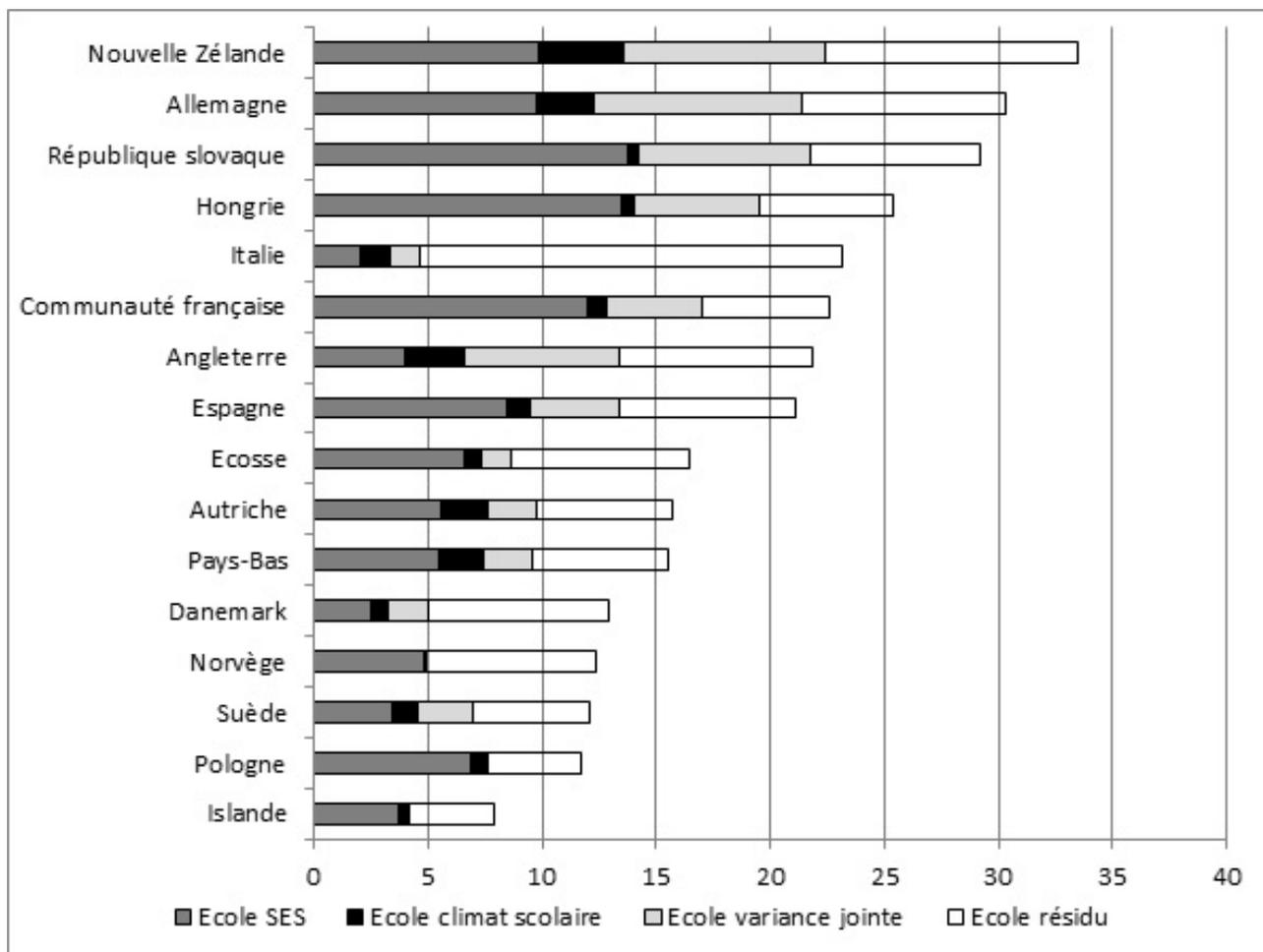
Ces analyses ont été conduites à partir de la base de données internationale PIRLS 2006 pour l'enseignement primaire (4e année) et à partir de la base de données PISA 2009 pour l'enseignement secondaire (les élèves de 15 ans). Dans les deux cas, les analyses portent sur les performances en compréhension de l'écrit.

Plus précisément, il s'agit ici de tenter de cerner par quels facteurs s'expliquent les différences de performances entre établissements. Dans quelle mesure s'expliquent-elles par le recrutement social de l'école, par des variables plus pédagogiques, ou par un effet conjoint ? Pour rappel, il a été montré dans la section consacrée aux études en milieu naturel que les classes fréquentées par les élèves les plus forts et/ou les élèves de milieux plus favorisés présentaient plus souvent des caractéristiques pédagogiques garantes d'efficacité. Il est donc particulièrement important de se pencher sur ces phénomènes de co-variation ou de variance jointe entre l'origine sociale et les environnements d'apprentissage.

La figure 2 présente la décomposition de la variance école, exprimée en pourcentage de la variance totale pour l'enseignement primaire (source : base de données PIRLS 2006).

27. Dans l'idéal, il conviendrait bien entendu d'inclure d'autres variables caractérisant le fonctionnement pédagogique ou les opportunités d'apprentissage au sein des classes. Cependant, ces variables se sont pas nécessairement disponibles à la fois dans les bases de données PIRLS et les bases de données PISA.

Figure 2 – Regroupement homogène des élèves selon des critères sociaux et déterminisme social



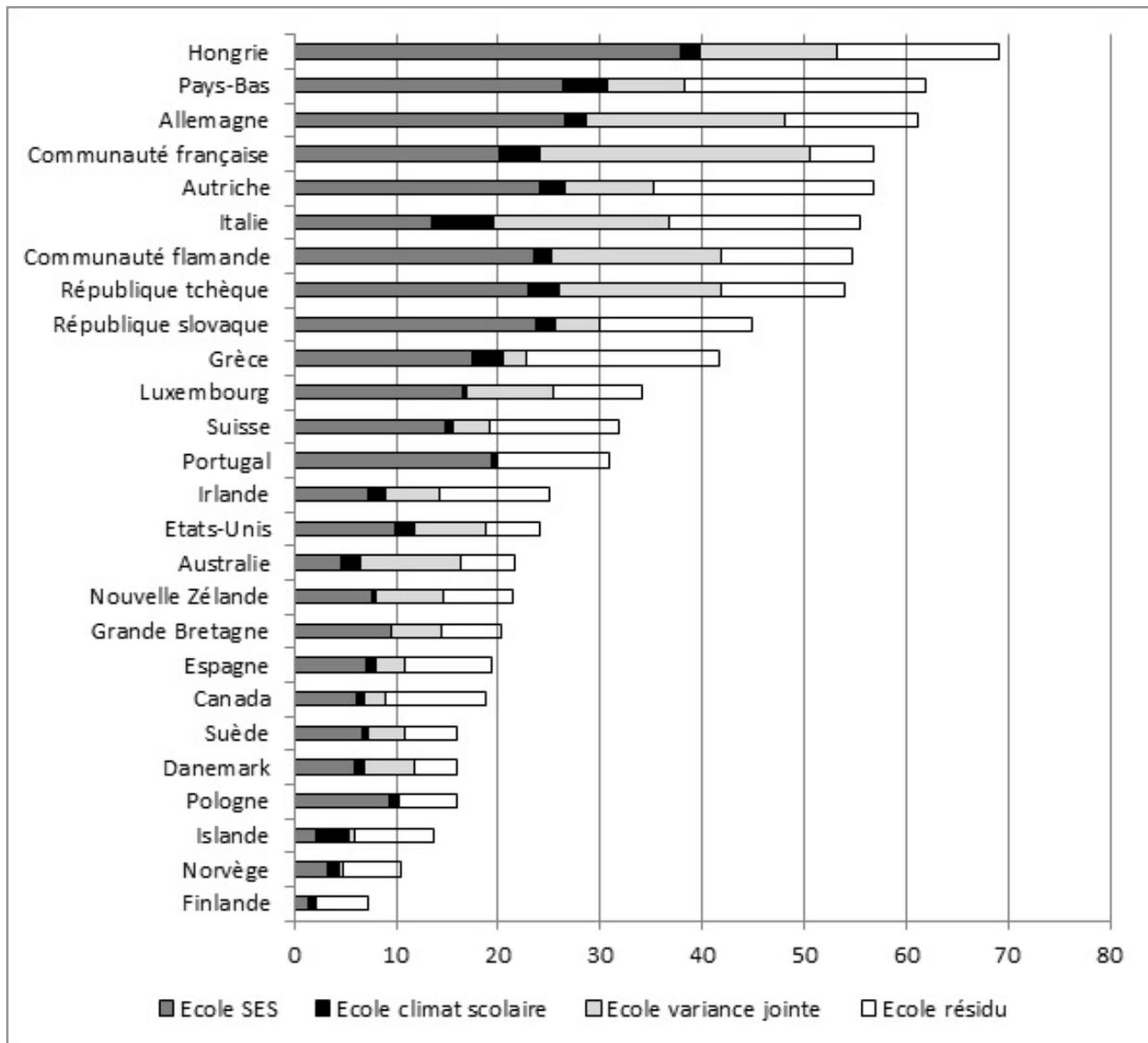
Lire ainsi : en Allemagne, 30 % de la variance des performances se situe entre écoles. Environ un tiers de cette variance école (soit 10 % de la variance totale) s'explique par le recrutement social des écoles. Une faible part de la variance s'explique directement par le climat scolaire. En revanche, presque un tiers de la variance entre écoles constitue de la variance jointe entre recrutement social et climat scolaire. Enfin, une part non négligeable de la variance entre écoles ne s'explique pas par ces variables et reste inexpliquée (résidu).

Les parties ombrées des histogrammes représentent la part de variance expliquée respectivement par

1. le recrutement social de l'école, sous contrôle du climat scolaire ;
2. le climat scolaire, sous contrôle du recrutement social de l'école ;
3. l'effet joint du recrutement social et du climat scolaire.

Comme le montre la figure 2, dans la plupart des pays, ces deux groupes de variables expliquent plus de la moitié des différences de performances entre écoles. En Communauté française de Belgique et en Hongrie, ce pourcentage grimpe même au-delà de 75 %. Par contre, en Italie, il ne dépasse pas les 20 %. Dans beaucoup de systèmes éducatifs, l'effet net du recrutement social (c'est-à-dire sous contrôle des variables de climat scolaire, ce qui revient à rendre les écoles équivalentes quant à ces variables de climat scolaire) à lui seul explique une part importante de la variance entre écoles. C'est notamment le cas en Nouvelle-Zélande, en Allemagne, en Slovaquie, en Hongrie et en Belgique francophone. L'effet net du climat scolaire, dans la plupart des pays, représente un très faible pourcentage de variance entre écoles. En d'autres termes, au-delà du recrutement social des écoles, l'effet du climat scolaire est souvent négligeable. Par contre, le pourcentage de variance jointe, dans bon nombre de pays, est relativement important et nettement plus élevé que l'effet net du climat scolaire, témoignant ainsi d'une étroite intrication entre recrutement social et qualité du climat scolaire. Ainsi, les écoles qui accueillent un public davantage favorisé disposent d'un meilleur climat scolaire que les écoles au recrutement plus défavorisé ; dans les écoles plus favorisées, les problèmes de discipline ont tendance à être moins fréquents, les relations élèves-professeurs plus positives et les élèves expriment davantage qu'ils se sentent bien à l'école. La figure 3 présente la décomposition de la variance école, exprimée sous forme de pourcentage de la variance totale pour l'enseignement secondaire (source : base de données PISA 2009).

Figure 3 – Décomposition de la variance école (en pourcentage de la variance totale) (Source : base de données PISA 2009) - Enseignement secondaire



L'effet observé dans l'enseignement primaire pour les groupes de variables recrutement social et climat scolaire se retrouve dans l'enseignement secondaire, avec quelques nuances. Les variables de recrutement social rendent compte d'une grande part de la variance entre écoles. Les variables de climat scolaire, sous contrôle des variables de recrutement social, n'expliquent qu'une part négligeable de la variance entre écoles. Enfin, la variance jointe reste globalement importante, témoignant à nouveau du lien étroit entre recrutement social et climat scolaire.

En Hongrie, aux Pays-Bas, en Allemagne, en Communauté française de Belgique, en Autriche, en Italie, en Communauté flamande de Belgique, et en République Tchèque, pays qui présentent plus de 50 % de la variance totale entre écoles, le recrutement social de l'école, sous contrôle du climat scolaire, explique environ 40 % de la variabilité entre écoles, ce qui correspond, comme l'indique la figure 3, à plus de 20 % de la variance totale de la performance des élèves en compréhension de l'écrit. Dans ces pays, les écoles se différencient considérablement par la performance moyenne de leurs élèves, mais ces différences sont particulièrement associées au recrutement social. Les établissements performants accueillent en plus large proportion des élèves issus des milieux privilégiés et les établissements peu performants concentrent davantage d'élèves issus des milieux défavorisés. Par contre, en Norvège, en Islande et en Finlande, la part de variance entre écoles expliquée par le recrutement social est très faible ; dans ces systèmes éducatifs - soulignons-le - les différences de performance entre écoles sont d'ailleurs limitées.

Dans tous les pays, le climat scolaire, sous contrôle du recrutement social de l'école, est faiblement associé à la performance moyenne des établissements. Comme dans l'enseignement primaire, l'effet joint est souvent supérieur à l'effet net du climat scolaire. Cet enchevêtrement entre climat scolaire et recrutement social est d'autant plus important que le pays pratique une forte ségrégation académique. Ainsi, l'effet joint représente environ 25 % de la variabilité totale en Communauté française et environ 20 % en Allemagne. Dans ces pays, les écoles performantes recrutent donc des élèves d'origine plutôt favorisée et n'éprouvent pas de difficulté majeure à instaurer un climat serein propice aux apprentissages, alors que les élèves issus de familles défavorisées se concentrent dans des écoles qui peuvent constituer de véritables ghettos où se conjuguent difficultés liées à la précarité sociale, problèmes disciplinaires et climat peu favorable aux apprentissages.

Ces observations rejoignent certaines des conclusions du rapport PISA 2009 que l'OCDE consacre aux inégalités sociales (OCDE, 2011b). Des corrélations ont été calculées entre le milieu socioéconomique moyen des établissements et différentes caractéristiques pédagogiques de ces établissements. On peut ainsi déterminer si les établissements qui recrutent un public socialement plus favorisé présentent des caractéristiques plus favorables en termes de ressources, d'équipement, de qualification du personnel, ce qui contribuerait à accroître les inégalités en ajoutant aux inégalités des chances une inégalité de traitement.

En ce qui concerne l'indice de pénurie d'enseignants qualifiés, la corrélation moyenne pour l'ensemble des pays de l'OCDE est de -0,13. La pénurie est donc plus importante en moyenne dans les établissements où le recrutement est plus défavorisé. Cette situation inégalitaire est particulièrement marquée (corrélation égale ou supérieure à -0,25) en Australie (-0,28), en Communauté française de Belgique (-0,30), aux États-Unis (-0,28), en Islande (-0,37), au Luxembourg (-0,33), au Mexique (-0,26) et en République tchèque (-0,34). Dans 15 des 34 systèmes éducatifs de l'OCDE, la corrélation est non significative et proche de zéro, peut-être tout simplement parce qu'il n'y a pas de pénurie d'enseignants. Dans tous les systèmes éducatifs où il existe un lien significatif entre la pénurie et le recrutement social de l'établissement, celle-ci va dans le

même sens, celui d'une inégalité en faveur des élèves favorisés socialement. Si on considère que l'absence d'enseignants qualifiés constitue un facteur défavorable qui pénalise les établissements, on peut constater que la plupart des systèmes éducatifs ne réussissent pas à lutter efficacement contre la tendance qui consiste à ce que les enseignants les plus qualifiés se retrouvent affectés dans les établissements les plus favorisés. Qu'en est-il de la qualité des ressources pédagogiques ? La corrélation moyenne pour les pays de l'OCDE est de 0,13, indiquant qu'en moyenne la qualité des ressources pédagogiques a tendance à être un peu meilleure quand le recrutement de l'établissement est plus favorisé. Dans 18 des 34 pays, la corrélation est faible, proche de zéro et non significative, traduisant une répartition uniforme des ressources, conforme au principe de l'égalité de traitement. Dans quelques systèmes éducatifs, la corrélation est plus forte (égale ou supérieure à 0,20) et traduit le fait que les ressources sont de meilleure qualité dans les établissements les plus favorisés. C'est particulièrement le cas en Australie (0,31), au Chili (0,35), en Israël (0,25), au Mexique (0,59), au Portugal (0,24), en Suède (0,26) et aux États-Unis (0,22). On ne relève aucun système éducatif où la tendance inverse serait observée - meilleures ressources dans les établissements défavorisés.

Enfin, en ce qui concerne le taux d'encadrement, la corrélation moyenne au niveau des pays de l'OCDE est de 0,15, ce qui signifie qu'en moyenne le taux d'encadrement a tendance à être plus favorable dans les établissements les plus défavorisés. "La moitié environ des pays de l'OCDE tentent d'augmenter le nombre d'enseignants dans les établissements défavorisés, vraisemblablement dans l'intention d'atténuer ce handicap. Cette corrélation est particulièrement marquée en Belgique, en Italie, en Irlande, en Espagne, en Estonie, en Islande, au Portugal, au Japon, aux Pays-Bas et en Corée. Parmi les pays de l'OCDE, les établissements défavorisés ne se caractérisent par un taux d'encadrement inférieur qu'en Turquie, en Slovénie, en Israël et aux États-Unis : les élèves y sont plus nombreux par enseignant." (OCDE, 2011b, p. 42).

À l'aune de ces trois indicateurs de qualité des ressources humaines et pédagogiques, plusieurs tendances se dégagent : celle d'abord, et elle est majoritaire, d'une absence de lien entre les ressources et le recrutement social des établissements - ce qui correspond à une répartition uniforme des ressources, quel que soit le recrutement social de l'établissement (principe d'égalité de traitement). Une autre tendance, assez répandue, consiste à "laisser-faire" les choses - ce qui conduit à ces situations inéquitables d'accroissement des inégalités : les écoles au recrutement le plus favorisé sont aussi celles où les ressources - humaines principalement (davantage d'enseignants qualifiés, meilleures conditions d'encadrement) et pédagogiques, dans une moindre mesure, sont les plus favorables. Les États-Unis, Israël, l'Australie, les pays économiquement les moins avancés (Chili, Mexique, Turquie) se retrouvent à plusieurs reprises dans cette situation susceptible d'accroître les inégalités de départ en y superposant une inégalité de traitement. Enfin, une dernière tendance, plus marginale, témoigne de politiques compensatoires délibérées consistant à allouer davantage de moyens aux écoles les plus défavorisées. C'est notamment ce que l'on peut observer, dans plusieurs systèmes éducatifs, en ce qui concerne les taux d'encadrement. Dans ces systèmes, le ratio élève-enseignant est plus favorable dans les établissements les moins favorisés. Sans doute faut-il voir là la résultante de la combinaison de deux approches. D'une part, une approche directe, qui consiste, dans le cadre de politiques de zones d'éducation prioritaire ou de discrimination positive, à allouer davantage de moyens humains aux établissements dont le recrutement social est particulièrement défavorisé. D'autre part, une approche indirecte, qui consiste à réduire la taille des classes et améliorer l'encadrement dans les établissements qui concentrent des élèves en plus grande difficulté d'apprentissage, généralement regroupés dans des classes ou des filières techniques et/ou professionnelles. Comme les élèves en difficulté d'apprentissage sont plus

fréquemment issus de milieux défavorisés, le taux d'encadrement apparaît comme lié au recrutement de l'établissement. Mais c'est davantage la filière fréquentée ou les difficultés d'apprentissage plutôt que les caractéristiques sociales du public des écoles qui dictent alors l'accroissement de moyens.

IV L'impact du regroupement des élèves par aptitude sur le plan socio-affectif²⁸

Jusqu'à présent, nous nous sommes principalement penchée sur les effets que pouvait avoir la différenciation ou certains de ses leviers principaux (redoublement, classes de niveau) sur les acquis scolaires. Or, d'autres aspects importants de la vie des élèves peuvent être impactés par le mode ou le degré de différenciation du système éducatif, en particulier les différentes composantes de la motivation (attitudes envers l'école, concept de soi...), ou les aspirations des élèves relatives à leur avenir (choix d'études, aspirations socioprofessionnelles).

Une étude longitudinale menée en Suisse par Müller et Hofmann (2014) a par exemple mesuré l'évolution entre le début et la fin d'année au grade 7 de variables socio-affectives (comportement antisocial, gestion de la colère, détresse émotionnelle) d'élèves selon la filière à laquelle ils étaient affectés en fin d'école primaire. Les auteurs ont tenu sous contrôle différentes caractéristiques des élèves : statut socioéconomique, statut relatif à l'immigration, présence d'un retard scolaire, présence de livres à la maison, soutien des parents. En Suisse, dans le canton de Fribourg, les élèves sont répartis en quatre filières distinctes : "avancée", "générale", "de base", "besoins spéciaux". À caractéristiques identiques, les élèves de la filière "de base" voient leurs problèmes d'ajustement social augmenter davantage que ceux des autres filières. Ceci n'est à vrai dire guère surprenant : le fait de se trouver assigné à une filière moins prestigieuse d'une part (effet d'étiquetage) et de fréquenter au quotidien des élèves qui présentent des difficultés d'apprentissage ou d'un autre ordre (effet de pairs) n'est pas de nature à promouvoir des comportements ajustés sur le plan émotionnel et social.

Une littérature abondante a investigué les effets que peut avoir le fait de fréquenter une école ou une classe plus ou moins performante principalement sur le **concept de soi académique des élèves**, et mis en évidence des effets qui sont parfois déroutants et relativement peu connus dans le monde de l'éducation. C'est pourquoi nous avons décidé d'y consacrer une section à part dans ce rapport. Le concept de soi académique (ASC "*academic self-concept*"), à la différence du concept de soi en général, renvoie à la manière dont l'élève se perçoit dans une discipline scolaire donnée. Il est souvent mesuré à l'aide d'items tels que *je suis fort en maths, je suis un des meilleurs de ma classe...* Ceci suppose une comparaison au moins implicite avec d'autres élèves de la classe ou de l'école que l'on fréquente. On comprend immédiatement en quoi le concept de soi peut être influencé par la composition des groupes dans lesquels l'élève s'inscrit. Selon que ce groupe classe ou école est plus ou moins performant, selon que ce groupe est plus ou moins hétérogène, le point de comparaison que prend l'élève sera plus ou moins variable selon l'école ou la classe fréquentée, et le concept de soi en sera inévitablement affecté. D'autres variables comme le sentiment d'efficacité en mathématiques (jugement sur sa capacité à effectuer une liste de tâches mathématiques précises) sont moins sensibles aux mécanismes de comparaison sociale (Monseur et Lafontaine, 2007). À l'échelle du système éducatif, on peut poser l'hypothèse que le mode et le degré de différenciation du système seront

28. Cette section a été rédigée en collaboration avec Virginie Dupont, assistante à l'Université de Liège, que je remercie ici pour sa contribution (Dupont et Lafontaine, umis).

en lien avec la manière dont les élèves y construisent leur concept de soi académique. Le concept de soi académique est une des composantes essentielles de la motivation et est bien corrélé avec les performances dans le domaine. Les élèves les plus performants ont un concept de soi plus positif, ce qui les pousse à s'investir dans les apprentissages du domaine et, en un mouvement vertueux, à devenir plus performants ou à le rester (Marsh *et al.*, 2005 ; McKenna *et al.*, 2012).

Sur la base de la littérature de recherche consacrée aux effets de pairs ou de composition, on tend généralement à considérer que fréquenter une école performante ou socialement favorisée est d'emblée associé à des conséquences positives pour le jeune tant sur le plan des performances que sur le plan motivationnel. Si ceci est largement exact sur le plan des acquis scolaires, il peut en aller autrement sur le plan socio-affectif. Les nombreux travaux de Marsh et ses collaborateurs portant sur le concept de soi scolaire ont mis en évidence un effet appelé *Big-Fish-Little-Pond-Effect* (BFLPE) : le poisson qui nage dans un petit étang se sent plus gros que s'il nageait dans l'océan. La perception de soi, pour les élèves comme pour le poisson, varie en fonction de son environnement immédiat.

En étudiant différents contextes, et en utilisant notamment les données PISA (Marsh et Hau, 2003 ; Seaton *et al.*, 2010) et TIMSS (Marsh *et al.*, 2015), Marsh et ses collaborateurs ont pu montrer qu'au niveau individuel, le concept de soi académique est évidemment corrélé positivement avec les performances : un élève plus performant tend à avoir un meilleur concept de soi en mathématiques. Toutefois, on observe un effet négatif de la moyenne des performances de l'école sur le concept de soi de l'élève. Ainsi, si on compare des élèves qui obtiennent des performances similaires au test PISA par exemple, mais dont certains fréquentent une école plus performante et d'autres une école moins performante, ceux qui fréquentent l'école moins performante ont un concept de soi plus favorable que ceux fréquentant une école plus performante. Selon Marsh, cet effet BFLPE est robuste, généralisé et universel. Il concerne aussi d'autres variables que le concept de soi comme les aspirations ou le choix d'études (Marsh et O'Mara, 2010 ; Nagengast et Marsh, 2012).

Voici, brossé à grands traits, ce qu'est le BFLPE et ce qu'il peut apporter de neuf à la réflexion sur l'impact du regroupement par aptitude. Alors que les élèves fréquentant les écoles, les filières ou les classes les plus performantes en tirent d'incontestables bénéfices sur le plan de leurs apprentissages et acquis scolaires, les avantages seraient nettement moindres, voire se changeraient en désavantages, sur le plan de la motivation et du concept de soi en particulier.

Pour nourrir la réflexion sur les effets de la différenciation sur le plan motivationnel, la vaste littérature sur le sujet sera examinée au travers des questions suivantes :

- Le BFLPE n'est-il pas contrebalancé par d'autres effets, qui seraient tout aussi puissants, comme le prestige/standing perçu de l'établissement, la filière ou la classe fréquentée ?
- Le point de référence/comparaison choisi par les élèves lorsqu'ils s'autoévaluent est-il l'école, le cas échéant la filière, ou encore la classe fréquentée ?
- Certains élèves sont-ils plus sensibles au BFLPE ?
- Le BFLPE (effet négatif de la moyenne des performances de l'école) sur le concept de soi est-il plus marqué dans certains systèmes éducatifs, en particulier selon leur degré ou leur mode de différenciation ?

A BFLPE et prestige perçu : effet de contraste et effet d'assimilation

L'effet de contraste auquel correspond le BFLPE (par contraste avec des pairs très performants, je ne me sens pas très fort...) n'est pas la seule comparaison sociale en milieu scolaire. Les études menées en psychologie sociale portant sur les comparaisons sociales en milieu scolaire (Blanton *et al.*, 1999 ; Huguet *et al.*, 2001) mettent en évidence que se comparer, de façon libre, à des pairs plus performants que soi peut être inspirant plutôt que menaçant. Par ailleurs, comme le soulignent Trautwein *et al.* (2009), les élèves fréquentant une école performante pourraient tirer une certaine fierté du fait d'appartenir à une école ou programme d'étude sélectif qui retentirait sur leur concept de soi scolaire. Dans ces deux derniers cas, on parle d'effet d'assimilation puisque l'individu se compare avec d'autres auxquels il voudrait ressembler, ou au groupe auquel il se sent appartenir et dont il peut tirer fierté.

Deux approches distinctes ont été utilisées afin de mettre en évidence un possible effet d'assimilation lié à la fréquentation d'une école ou d'une filière plus prestigieuse. Soit les chercheurs essaient de capter l'effet d'assimilation en incluant dans l'analyse, à côté de l'école, la filière fréquentée par l'élève (Marsh *et al.*, 2001 ; Lüdtke *et al.*, 2005 ; Trautwein *et al.*, 2006 ; Preckel et Brüll, 2010), soit ils incluent une échelle mesurant la perception par le jeune de la position, en termes de prestige, de l'école ou de la filière fréquentée (Marsh *et al.*, 2000 ; Trautwein *et al.*, 2009). Dans les deux cas, l'objectif est de distinguer l'effet de contraste et l'effet d'assimilation confondus lorsque l'on envisage uniquement l'effet des performances moyennes de l'école sur le concept de soi.

Plusieurs études menées principalement en Allemagne - système éducatif relevant du modèle de la séparation, avec filières précoces (dès le grade 5 ou le grade 7, selon les *Länder*) font apparaître l'existence d'un effet positif du prestige perçu de l'école ou de la filière, distinct du BFLPE²⁹. Le BFLPE au sens technique du terme (coefficient de régression négatif des performances moyennes de l'école sur le concept de soi) apparaît comme étant l'effet net ou la résultante d'effets s'exerçant en sens opposés : un fort effet de contraste (effet négatif) lié au groupe qui sert de point de comparaison et un effet d'assimilation (effet positif) plus faible engendré soit par des comparaisons libres avec certains élèves plus performants, soit par l'effet d'appartenance à un groupe valorisé ou prestigieux. Comme l'effet de contraste est généralement plus puissant que l'effet d'assimilation, dans la majorité des cas, un coefficient de régression négatif des performances moyennes de l'école sur le concept de soi (BFLPE) est observé.

B Groupe de comparaison : école ou classe ?

Les recherches consacrées au BFLPE, surtout celles de la première génération, ont pour la plupart testé l'influence des performances moyennes de l'école sur le concept de soi. Quelques-unes ont utilisé les performances moyennes de la classe fréquentée. Dans la plupart des cas, les recherches étudiant l'effet des performances au niveau de l'école sont celles qui utilisent la base de données internationale PISA. Dans ce cas, travailler au niveau école est davantage lié aux contraintes de l'échantillonnage qu'à un véritable choix : dans PISA, ce sont en effet des élèves qui sont sélectionnés à l'intérieur des écoles, et non des classes entières. Peu d'études ont inclus les deux niveaux (classe et école) et testé leur importance respective. Pourtant, selon les études en psychologie sociale (Zell et Alicke, 2009), les individus sont davantage portés à utiliser les informations les plus locales dans les processus de comparaisons sociales, c'est ce qu'on appelle l'effet

29. À l'exception notable de l'étude de Trautwein *et al.* (2006), sur laquelle nous allons revenir en détail plus loin.

de la dominance locale.

Marsh et al. (2014) ont fait l'hypothèse que, dans le cas du BFLPE, la composition académique de la classe (qui est un cadre de référence plus proche de l'élève) serait un meilleur prédicteur du concept de soi académique que la composition académique de l'école. Au départ d'une large base de données composée de 15 356 élèves néerlandais du grade 9 nichés dans 651 classes de 65 écoles, *Marsh et al. (2014)* ont testé ces hypothèses à l'aide d'un modèle à trois niveaux (élèves - classes - écoles) et constatent que le BFLPE au niveau de l'école disparaît complètement sous contrôle des performances moyennes de la classe. L'effet du contexte de l'école généralement observé résulterait de la somme des effets classe, les écoles performantes étant plus souvent constituées de classes elles-mêmes performantes.

Liem et al. (2013) ont également étudié l'effet de dominance locale dans le cas du BFLPE à Singapour en distinguant le niveau classe, filière et école. À Singapour, au sein des écoles, les élèves du secondaire sont répartis en trois filières distinctes (très performante, moyenne, faible) en fonction de leurs résultats à un même test. En accord avec la théorie de la dominance locale, les auteurs confirment la prédominance de l'effet de la classe ou de la filière sur celle de l'école dans la formation du concept de soi scolaire.

C Certains élèves sont-ils plus sensibles au BFLPE ?

L'existence du BFLPE est attestée dans tous les systèmes éducatifs qui ont participé à PISA ou à TIMSS. Son caractère universel semble incontestable. Mais certaines catégories d'élèves ne sont-elles pas davantage sensibles à la comparaison sociale ? Plusieurs recherches ont pris en considération des variables dites "modératrices" du BFLPE. Le BFLPE varie-t-il en fonction du niveau d'aptitude de l'élève, de son sexe ou son statut socioéconomique, les attitudes scolaires ou encore de certains traits de personnalité ? Les résultats qui se dégagent de ces études sont peu concluants. Même s'il existe de légères variations dans l'intensité de l'effet négatif des performances moyennes sur le concept de soi scolaire en fonction de l'aptitude de l'élève, toutes les études confirment la présence d'un BFLPE généralisé à l'ensemble des élèves. Les analyses de données nationales mettent aussi bien en évidence des résultats non significatifs (Allemagne, Pays-Bas, France) qu'un BFLPE plus fort pour les élèves faibles (Allemagne, France) ou l'inverse (France). Les analyses fondées sur PISA mettent quant à elles en évidence un BFLPE plus prononcé pour les élèves les plus forts lorsque les analyses sont menées au travers de tous les pays, et des résultats non significatifs ou un BFLPE plus prononcé pour les élèves les plus forts lorsque les données des pays sont analysées distinctement.

Les rares études qui ont inclus une analyse en fonction du sexe de l'élève font apparaître que les filles sont légèrement plus affectées que les garçons par les performances de leurs camarades de classe. À notre connaissance, *Seaton et al. (2010)* sont les seuls à avoir étudié l'effet modérateur de l'origine socioéconomique sur le BFLPE en utilisant les données PISA 2003. Seul un lien faible avec les possessions culturelles apparaît ; l'effet est à nouveau marginal.

Jonkmann et al. (2012) ont étudié six traits de personnalité comme potentiels modérateurs du BFLPE : le caractère névrosé, extraverti, ouvert, consciencieux, agréable et narcissique de la personne. Deux traits ressortent comme étant des modérateurs du BFLPE : la névrose et le narcissisme. Ainsi, les personnes émotionnellement instables, qui ont un haut niveau d'anxiété et qui se tracassent excessivement ont tendance à être plus sensibles au BFLPE : à compétences égales, les performances moyennes de l'école fréquentée

ont un effet plus négatif sur le concept de soi scolaire de ces personnes. Le deuxième trait de caractère qui modère le BFLPE est le narcissisme. À compétences égales, les personnes qui se sentent supérieures, qui ont des attitudes auto-suffisantes, autoritaires et dominantes sont moins touchées par l'effet négatif des performances moyennes de l'école qu'ils fréquentent.

En résumé, les études peinent à mettre en évidence des caractéristiques d'élèves qui seraient clairement plus sensibles au BFLPE. Seule l'anxiété, le narcissisme et le sexe peuvent moduler de façon marginale l'importance du phénomène. La plupart des élèves sont sensibles au BFLPE.

D BFLPE et formes de différenciation : pour aller plus loin dans la compréhension du phénomène

Comme déjà dit, le BFLPE a été mis en évidence dans des systèmes éducatifs et des contextes culturels variés, et cet effet est en conséquence considéré comme relativement universel. Plusieurs études invitent toutefois à la nuance sans aller jusqu'à remettre en question la généralité du phénomène. Ainsi, [Marsh et al. \(2015\)](#) ont étudié le BFLPE dans 13 pays à partir des données TIMSS pour les grades 4 et 8. Ils mettent en évidence un BFLPE significatif dans les 13 pays tant au grade 4 qu'au grade 8, mais notent toutefois que le BFLPE est de plus grande ampleur au secondaire qu'au primaire, c'est-à-dire au moment où l'on passe, du moins dans certains pays, à un regroupement plus important des élèves par aptitude.

Par ailleurs, quelques études ([Marsh et al., 2001](#) ; [Nagengast et Marsh, 2011](#) ; [van de Gaer et al., 2012](#)) montrent que le BFLPE est d'autant plus important que le système scolaire est différencié. En effet, dans les systèmes scolaires intégrés, c'est-à-dire où les élèves ne sont pas regroupés dans des classes ou des écoles de niveaux différents, les groupes de référence (classe et/ou école selon les études) sur la base desquels les élèves s'autoévaluent sont relativement similaires du point de vue des aptitudes. À l'inverse, dans les systèmes différenciés, où les élèves sont regroupés par niveau, les cadres de référence sur lesquels s'appuie la comparaison sociale sont différents d'une classe ou d'une école à l'autre, ce qui décuple l'effet des comparaisons sociales.

L'étude menée par [Marsh et al. \(2001\)](#) est à cet égard exemplaire. Elle porte sur des données collectées en Allemagne auprès d'élèves du grade 7 en début, milieu et fin d'année scolaire 1991/1992, soit juste après la réunification. Au moment de sa réunification, l'Allemagne constitue un terrain particulièrement propice pour étudier l'impact du caractère différencié du système éducatif. En effet, alors que les élèves de l'Allemagne de l'Ouest ont toujours connu un système différencié comprenant différentes filières, ceux de l'Allemagne de l'Est sont quant à eux habitués à un système non différencié. Au moment de la réunification, c'est le système différencié, déjà en application à l'Ouest, qui est adopté pour l'Allemagne réunifiée. Les auteurs font l'hypothèse que les élèves d'Allemagne de l'Est, qui n'ont donc pas connu de système scolaire différencié jusque-là, seront moins affectés par le BFLPE que ceux d'Allemagne de l'Ouest. Les résultats confirment leur hypothèse : en début d'année, au moment du changement de système, le BFLPE est significativement plus important pour les élèves d'Allemagne de l'Ouest que pour ceux d'Allemagne de l'Est. Cette différence est toujours significative mais de moindre ampleur en milieu d'année ; en fin d'année, elle n'est plus significative. À ce moment, tous les élèves se sont accoutumés au système différencié qui est désormais leur lot commun. Dans un tel système, le concept de soi des élèves subit davantage l'influence des performances moyennes de leur classe.

Enfin, une dernière étude, également menée en Allemagne (Trautwein *et al.*, 2006) apporte des éléments particulièrement intéressants pour comprendre la dynamique des mécanismes en jeu dans différentes modalités de différenciation et leurs retombées sur le plan motivationnel. Ces auteurs commencent par poser un cadre de référence théorique décrivant différents types de modes de groupement des élèves (en anglais *streaming, tracking* ou *ability grouping*) selon trois critères : le niveau institutionnel, le rôle des aptitudes et l'impact du placement de l'élève sur son parcours ultérieur. Au niveau institutionnel, ils distinguent les groupes de niveau à l'intérieur de la classe (*within-class achievement achievement grouping*), surtout utilisés dans le primaire, les cours de différents niveaux (*course-level grouping*), largement répandus dans le secondaire (par exemple : suivre un cours de mathématiques normal ou à un niveau avancé) et les *regroupements au niveau de l'école (school level)* au sein desquels il faut distinguer les regroupements *implicites*, fondés par exemple sur le lieu de résidence, et les regroupements *explicites*, avec une offre d'enseignement différenciée (filières) destinée à des publics d'élèves distincts.

C'est souvent sur la base des performances scolaires antérieures que l'élève est orienté vers l'une ou l'autre filière, mais d'autres facteurs, de l'ordre du choix (*opt-in tracking*) de l'élève ou de ses parents peuvent bien entendu intervenir. Enfin, selon les systèmes éducatifs, le fait d'être placé dans l'un ou l'autre cours ou telle ou telle filière peut restreindre plus ou moins fortement les choix d'études et de carrières. Le caractère plus ou moins explicite des modalités de regroupement rend aussi plus ou moins visible des modalités de différenciation, dont les effets à long terme peuvent échapper aux élèves et aux parents moins complices des attendus et normes de l'institution scolaire.

Comme le font remarquer les auteurs de l'étude, "le système éducatif allemand est sans doute le plus clairement différencié des pays occidentaux industrialisés" (p. 788) et à cet égard, constitue une sorte de laboratoire idéal pour y investiguer l'importance des effets de contraste (BFLPE) et d'assimilation (prestige perçu). La différenciation y est précoce et a des conséquences lourdes pour le parcours de l'élève. Toutefois - et ceci est moins connu - il existe en Allemagne deux modalités de filiarisation (*tracking*) différentes, entre écoles (*between schools*) et à l'intérieur des écoles (*within-school*). Dans la première forme, qui concernait environ 90 % des élèves à l'époque de l'étude, les élèves sont assignés à différents types d'écoles correspondant à des filières³⁰ sur la base de leurs résultats scolaires, en sorte que chaque école regroupe des élèves de niveau relativement semblable (homogènes). Dans la seconde forme (10 % des élèves), les écoles "Gesamtschulen" (*comprehensive schools*) regroupent un ensemble d'élèves plus hétérogène du point de vue de leurs aptitudes, et ceux-ci, à l'intérieur de l'école sont répartis en deux ou trois "filiales" ou classes de niveau en fonction de leurs aptitudes : dans ce cas, les filiales correspondent davantage à des niveaux de cours (*course-level* ou *setting*) plus ou moins poussés en mathématiques, allemand et langues étrangères³¹. On comprend aisément que selon la modalité de filiarisation en vigueur, la manière dont un élève se construit un concept de soi peut varier sensiblement. C'est précisément à cette question que cette étude sophistiquée s'intéresse. Elle inclut également, à côté des résultats de tests standardisés, les notes scolaires obtenues par les élèves, ce qui permet d'investiguer le rôle que jouent ces notes sur le concept de soi.

Dans le cas où les filiales sont organisées entre écoles, Trautwein *et al.* (2006) observent bien un effet de

30. Les trois filiales sont le *Gymnasium* (la plus exigeante), la *Realschule* (intermédiaire) et la *Hauptschule* (moins exigeante sur le plan académique). Dans les deux dernières filiales, les élèves obtiennent leur diplôme après le grade 9 ou 10 et la plupart poursuivent dans le système de formation en alternance (*dual system*). Les élèves fréquentant le *Gymnasium* obtiennent leur diplôme à l'issue du grade 12 ou 13 et doivent obtenir l'*Abitur* (équivalent du bac français) pour pouvoir entrer à l'Université.

31. Cette forme se rapproche d'une modèle d'intégration à la carte, selon la typologie de Mons (2007).

contraste lié aux performances moyennes de l'école fréquentée : à compétences égales, les élèves fréquentant une école plus performante ont un concept de soi moins favorable (BFLPE classique). De manière moins attendue, ils observent également un BFLPE lié à la filière fréquentée, c'est-à-dire que le concept de soi en mathématiques d'élèves aux performances similaires fréquentant la filière *Gymnasium* (la plus prestigieuse) est moins favorable que celui des élèves des autres filières. Autrement dit, même si les filières sont bien visibles, tant pour les élèves que pour les parents, la fréquentation de la filière la plus prestigieuse n'est pas associée à un effet positif sur le concept de soi en mathématiques suffisamment puissant pour annihiler l'effet de contraste. De manière intéressante, il apparaît que les notes scolaires attribuées par leurs enseignants à des élèves qui ont obtenu des scores similaires au test standardisé sont nettement plus sévères au *Gymnasium*, et le sont nettement moins dans la filière la moins prestigieuse (*Hauptschule*) - confirmant la présence d'un effet Posthumus : les enseignants adaptent leur degré d'exigence et leurs notes en fonction du niveau de la classe, ici de la filière, reproduisant ainsi une courbe de Gauss (*grade-on-a-curve*). Les analyses montrent que le concept de soi est davantage influencé par les notes scolaires que par les résultats aux tests standardisés ; l'effet des performances au test sur le concept de soi est d'ailleurs entièrement médié par les notes scolaires. Comme les notes scolaires sont bien plus visibles aux yeux des élèves, les leurs et surtout celles des autres, que les résultats à un test externe *one shot* d'une part, et comme, en raison de l'effet Posthumus, les notes sont finalement à peu près les mêmes dans les différentes filières, les élèves des filières les plus prestigieuses ne se perçoivent pas comme meilleurs et ceux des filières moins prestigieuses comme moins bons, du moins à l'aune des instruments utilisés classiquement pour appréhender le concept de soi en contexte scolaire.

Dans le cas où les filières/classes de niveau sont organisées au sein des écoles, les résultats sont sensiblement différents. Tout d'abord, il apparaît que dans cette configuration, les enseignants ne recréent pas une courbe de Gauss pour chacune des filières/classes de niveau. Les élèves qui fréquentent la "meilleure" filière reçoivent de meilleures notes de leurs enseignants ; en quelque sorte, la courbe de référence se situe au niveau de l'école et non au sein de chaque filière. Les auteurs retrouvent un effet de contraste (BFLPE) lié aux performances moyennes de l'établissement et un puissant effet positif de la fréquentation de la filière la plus valorisée sur le concept de soi. Cet effet positif lié à la filière fréquentée s'explique complètement par les notes que l'enseignant donne aux élèves : sous contrôle des notes scolaires, la filière en soi n'a plus d'effet positif significatif sur le concept de soi scolaire. Mais comme les élèves de la filière la plus exigeante reçoivent des notes qui reflètent leurs habiletés objectivement meilleures, ils se sentent davantage aptes que ceux des filières moins exigeantes.

Que retenir de cet ensemble d'études portant sur l'impact des écoles, filières et classes fréquentées et du degré de différenciation du système sur la construction du concept de soi académique ?

- Tout d'abord, au vu des nombreuses études qui ont confirmé l'existence d'un BFLPE dans de nombreux contextes, l'impact des regroupements par niveau peut être différent sur le plan cognitif et socio-affectif. C'est un élément que les parents en particulier devraient connaître et prendre en considération lorsque la possibilité de choisir un établissement pour leur enfant leur est offerte. Vouloir à tout prix le meilleur établissement, la filière la plus prestigieuse pour son enfant n'est pas toujours le meilleur choix, et peut même se révéler délétère si l'élève est particulièrement sensible aux comparaisons sociales (enfants anxieux, filles dans les matières scientifiques) ;

- Les mécanismes de comparaison sociale - effet de contraste (BFLPE) ou d'assimilation - (prestige perçu) opèrent surtout au niveau le plus local (dominance locale) : pour évaluer ses capacités en mathématiques par exemple, l'élève se compare d'abord aux élèves de sa classe plutôt qu'à ceux de son école ;
- Les élèves s'appuient davantage, pour définir leur concept de soi académique, sur les notes scolaires et ce qu'elles reflètent (attentes, exigences plus ou moins élevées des enseignants) que sur leurs capacités objectivées par la passation d'un test standardisé ;
- À performances individuelles égales de l'élève, l'impact négatif de performances moyennes élevées de l'école sur le concept de soi s'observe davantage dans les systèmes différenciés que dans les systèmes peu différenciés ;
- Le caractère implicite/explicite des regroupements par niveau, l'impact de l'école ou la filière fréquentée pour le parcours ou le devenir de l'élève sont susceptibles de moduler les effets de contraste et d'assimilation ;
- Comme le montre l'étude de [Trautwein et al. \(2006\)](#) comparant en Allemagne les *filiales entre écoles* (modèle de la *séparation*) et les *filiales/classes de niveau à l'intérieur des écoles* (plus proches du modèle de *l'intégration à la carte*), le type de différenciation n'engendre pas les mêmes effets. Dans le cas allemand, l'effet d'assimilation (positif pour les meilleurs élèves et négatif pour les moins bons) que l'on peut retirer de la fréquentation d'une filière plus ou moins exigeante ne percole que quand les filières³² s'organisent à l'intérieur de l'école - ce qui est aisé à comprendre : les élèves peuvent alors se comparer "en local" aux élèves d'autres filières. Ce que les enseignants font d'ailleurs aussi, puisque leurs notes intègrent en quelque sorte une norme plus large que celle de la filière, à l'échelle de l'établissement. L'intérêt de cette dernière étude est qu'elle entre dans les mécanismes fins qui opèrent au sein d'un système éducatif, prenant en considération des procédures de différenciation autres que le seul regroupement par aptitude : après tout, adapter ses exigences, ses attentes ou ses notes en fonction du niveau de la classe, de la filière ou de l'école est une autre façon de prendre en compte les différences entre élèves et l'hétérogénéité des aptitudes, relevant davantage de la différenciation pédagogique que structurelle.

V Quelles modalités d'organisation/différenciation pour une école plus efficace et moins inégalitaire ?

L'école, dans les pays économiquement les plus avancés, ne réussit pas à endiguer les inégalités liées à l'origine sociale. Souvent, elle a tendance à les accentuer plutôt qu'à les réduire. Les études s'appuyant sur les enquêtes internationales menées par l'IEA et par l'OCDE ont permis de consolider une série de constats. S'ils n'apportent pas d'éléments véritablement neufs à la compréhension des mécanismes amplificateurs des inégalités, ces constats montrent à quel point les phénomènes d'inégalités sociales sont largement répandus de par le monde et restent une préoccupation d'une brûlante actualité dans nos sociétés contemporaines. Avant de poursuivre, rappelons ces principaux constats :

32. Dans ce cas, il s'agit plus de *setting* que de véritables filières (cours plus ou moins avancés dans différentes disciplines).

- Quel que soit le pays, quelle que soit la discipline ou le niveau d'études, les élèves issus de milieux favorisés obtiennent, en moyenne, de meilleurs résultats que les élèves issus de milieux défavorisés ;
- Dans les systèmes éducatifs qui régulent l'hétérogénéité des aptitudes scolaires par le redoublement, les élèves de milieux défavorisés sont plus souvent confrontés au redoublement. Plus inquiétant encore, dans certains systèmes éducatifs particulièrement inéquitables, à performances scolaires identiques, les élèves défavorisés redoublent plus souvent ;
- Dans quasi tous les systèmes éducatifs à structure différenciée³³, les élèves d'origine modeste sont surreprésentés dans les filières professionnelles alors que les élèves de milieux privilégiés y sont sous-représentés. À nouveau, cette répartition inégale subsiste sous contrôle de la performance scolaire dans certains systèmes éducatifs (Monseur et Lafontaine, 2012) ;
- Les différentes pratiques de regroupement homogène des élèves, qu'elles se fondent sur les aptitudes ou l'origine sociale, vont de pair avec de moins bonnes performances moyennes des systèmes éducatifs et accentuent, dans les disciplines considérées comme majeures par l'institution scolaire, les différences de performance entre les élèves faibles et les élèves forts ;
- Ces pratiques accentuent le poids du milieu familial sur les performances scolaires ;
- Dans les systèmes fortement différenciés, le BFLPE est plus accentué (effet négatif des performances moyennes de l'école fréquentée sur le concept de soi) ;
- L'influence inégalitaire du regroupement homogène se poursuit au-delà de l'obligation scolaire puisqu'on la retrouve au niveau des aspirations d'études et professionnelles ;
- Les différences de performances entre écoles s'expliquent assez largement, dans nombre de systèmes éducatifs, par le recrutement social de l'établissement. Aux inégalités de départ liées à l'origine sociale vient donc s'ajouter une inégalité de traitement, puisque certains élèves - en l'occurrence ceux issus de milieux plus favorisés - bénéficient de meilleures conditions d'apprentissage (Gamoran, 2002). Dans une étude récente portant sur les occasions d'apprendre en compréhension de l'écrit utilisant les données PISA 2009, Lafontaine *et al.* (2015) ont pu montrer que dans les 34 pays de l'OCDE une proportion importante des différences de performances entre écoles s'expliquait par les occasions d'apprendre et la composition sociale du public de l'établissement (68 % de variance expliquée en moyenne). La part de variance jointe entre ces deux variables était importante (28 %), alors que la variance unique expliquée par les OTL n'était que de 8 % ; ceci montre à quel point les élèves d'origine moins favorisée se voient offrir des occasions d'apprendre moins riches que ceux d'origine plus favorisée. De surcroît, cette inégalité dans les occasions d'apprendre est plus marquée dans les systèmes éducatifs différenciés que dans les systèmes intégrés. Cette variance jointe est par exemple de 44 % en Belgique, 41 % en Allemagne, 39 % en France, 47 % en Hongrie, 53 % au Luxembourg, 43 % aux Pays-Bas... alors qu'elle est inférieure à 10 % dans les pays nordiques ou baltes (Finlande, Estonie, Suède). Les analyses de Schmidt *et al.* (2015) explorant les relations entre les OTL et les performances en mathématiques dans PISA 2012 montrent de même le lien entre OTL et origine sociale ; environ un tiers de la relation entre l'origine socioéconomique et les performances transite par des différences dans les occasions d'apprendre, qui sont donc socialement marquées.

33. Seule la République tchèque fait exception à cet égard.

La lecture de ces constats aurait quelque chose d'accablant si un autre regard n'était possible sur ces résultats, qui, plutôt que de mettre l'accent sur les tendances communément partagées par les systèmes éducatifs, s'attache à repérer les différences d'organisation et à en tirer parti pour voir dans quelle direction infléchir ce qui, à défaut, apparaîtrait comme un processus inéluctable de reproduction.

Dans tous les systèmes éducatifs, les performances autant que les parcours sont liés à l'origine sociale des élèves, mais la situation est plus critique dans certains d'entre eux. Le fil rouge qui traverse ce rapport est que l'organisation structurelle du système éducatif est de nature à tempérer ou à exacerber l'influence de l'origine sociale sur les performances et les parcours. Les pays à taux élevé de ségrégation académique ou sociale - que cette ségrégation s'opère via le redoublement, la sélection précoce ou un quasi-marché scolaire sont presque tous plus inéquitables que la moyenne. Des systèmes éducatifs tels que les deux Communautés belges, l'Allemagne, la France, le Luxembourg, la Hongrie... sont régulièrement épinglés pour leur caractère inéquitable. À l'opposé, les pays scandinaves et certains pays anglo-saxons - principalement le Canada - figurent en bonne place au rang des pays qui réussissent à contenir les inégalités sociales (OCDE, 2011b). À ce type de conclusion, certains ne manqueront pas d'opposer deux types d'arguments : ces pays plus équitables ne paient-ils pas le prix fort pour atteindre plus de justice sociale ? Autrement dit, ces pays ne sont-ils pas moins efficaces ou performants ? Ne sont-ils pas menacés de "nivellement par le bas" ? Par ailleurs, le lien robuste entre différenciation du système et inégalités sociales ne cache-t-il pas d'autres variables agissantes ou d'autres caractéristiques culturelles ou politiques des systèmes éducatifs qui rendent leurs modalités de fonctionnement irréductibles, difficiles à exporter ou à emprunter ? Plusieurs réponses peuvent être apportées à ces légitimes objections.

Tout d'abord, contrairement à ce que beaucoup pensent, l'efficacité (le niveau de performances) et l'équité d'un système éducatif (l'ampleur des inégalités liées à l'origine sociale) vont souvent de pair. La corrélation entre le pourcentage de variance de la performance en compréhension de l'écrit entre écoles et le score moyen dans PISA 2009 en compréhension de l'écrit est de $-0,38$ ³⁴. Cette corrélation est certes modeste, mais significative : plus un système est équitable, plus il a tendance à être efficace, et vice-versa. Il existe à vrai dire, parmi les pays de l'OCDE, assez peu de systèmes vraiment performants qui ne soient pas aussi équitables. La Flandre (Communauté flamande de Belgique) est à cet égard une exception notoire : c'est l'un des rares systèmes qui s'illustre par de très bonnes performances et une équité faible.

L'existence d'une telle corrélation entre efficacité et équité pourrait en laisser plus d'un sceptique, de même peut-être que l'ensemble des constats de cette section, fondés sur des données transversales et corrélationnelles, avec lesquelles il est toujours périlleux de se prononcer sur la causalité des effets. D'autres types de données, plus adéquats pour établir d'une façon plus robuste un lien de causalité entre structure du système éducatif et inégalités sociales, peuvent être convoqués.

Ainsi, Felouzis *et al.* (2011) ont mené une étude longitudinale dans le canton de Genève en suivant pendant trois ans un échantillon d'élèves testés en 9e, à l'âge de 14 ou 15 ans. Les auteurs disposent donc du niveau de performances de 1 654 élèves en 2003 et 1 782 élèves en 2006, ainsi que de leur parcours scolaire de la fin de l'école primaire jusqu'à la sortie du système éducatif. Cette recherche permet de mesurer l'effet respectif de la fréquentation des différentes filières d'enseignement à Genève, qui interviennent dès la fin de l'école primaire, et des performances dans PISA sur la probabilité d'être dans la filière gymnasiale (filière

34. La même corrélation s'élève à $-0,12$ en mathématiques et à $-0,23$ en sciences. Dans ces deux derniers cas, la corrélation n'atteint pas le seuil de significativité (à $p. 05$) sur le plan statistique.

la plus prestigieuse) un an ou trois ans après la 9e. Précisons que Genève offrait au moment de l'étude une situation quasi expérimentale du point de vue des filières dans le secondaire : trois établissements (17 % des élèves genevois) scolarisent les élèves dans un système intégré (tous les types d'élèves sont scolarisés ensemble), les 14 autres établissements proposent un système segmenté à filières : "A" (exigences élevées) et "B/C" (exigences faibles). En moyenne, l'origine socioéconomique des élèves de ces deux systèmes est identique ; il n'existe donc pas de biais compromettant la comparaison entre les élèves des deux systèmes. Les écarts de performances en fonction de l'origine sociale sont systématiquement plus élevés dans le système à filières, pour les deux cohortes (2003 & 2006) et dans les trois domaines évalués (sauf en sciences dans PISA 2006). Un an après la passation du test PISA, pour un niveau de compétence donné en 9e, la probabilité d'être orienté dans une filière gymnasiale est très différente selon que l'on est scolarisé dans une filière "A" ou "B/C". Pour des élèves ayant un même score en mathématiques correspondant à la moyenne, les probabilités d'être orienté dans la filière gymnasiale sont de 10 % s'ils sont dans une filière "B/C", de 50 % s'ils sont dans une filière "H" (hétérogène ou intégrée) et de 60 % s'ils sont dans une filière "A". On relève donc ici une discrimination négative forte puisque, à compétences égales dans PISA, les orientations vers les études longues dépendent de la filière suivie dans le secondaire inférieur. Trois ans après la 9e, les élèves qui étaient scolarisés dans le système unifié (filière "H") en 9e ont autant de chances d'être dans la filière gymnasiale, toutes choses égales par ailleurs, que ceux qui étaient scolarisés dans la filière "A", la plus exigeante.

Cela signifie que les processus d'orientation y sont bien moins inégalitaires, car ce système intègre tous les élèves, sans exclure ceux qui, dans un autre contexte, auraient été scolarisés en B/C. On mesure ainsi l'ampleur des inégalités scolaires que produit la séparation en filières différenciées et hiérarchisées dès la fin de l'école primaire.

Dans un même système éducatif, les inégalités sociales s'avèrent donc moindres lorsque les écoles ont un fonctionnement unifié, sans filières. Ces écoles ne connaissent pas de nivellement par le bas, au contraire : les élèves bénéficiant d'un système unifié sont donc "tirés vers le haut" plutôt que vers le bas.

Des enseignements forts, analogues à ceux issus de cette comparaison dans le canton de Genève, se dégagent aussi de l'examen de l'évolution observée en Pologne dans les cycles successifs de PISA.

La Pologne a connu en 1998-1999 une réforme de son système éducatif d'une ampleur considérable, qui a notamment porté sur la structure du système éducatif (Janowski, 2007 ; OCDE, 2010). Avant la réforme, l'enseignement primaire en Pologne durait 8 ans (jusqu'à l'âge de 14-15 ans) ; il n'existait pas, à proprement parler, d'enseignement secondaire inférieur. Après quoi, les élèves se dirigeaient soit vers le lycée (enseignement général, 4 ans), l'enseignement technique (4 ou 5 ans), l'enseignement professionnel "général" (3 ou 4 ans) ou l'enseignement professionnel "de base" (3 ans). D'après l'OCDE (2010), environ la moitié des étudiants fréquentaient l'enseignement professionnel, un tiers l'enseignement technique et seuls 20 % des étudiants suivaient le lycée (enseignement général). À partir de 1999-2000, la durée de l'enseignement primaire a été ramenée à 6 ans (élèves de 7 à 13 ans) et un nouveau type d'école - le *gimnazjum* - a été créé. Il correspond à l'enseignement secondaire inférieur, est de type intégré (sans filières) et accueille tous les jeunes de 13 à 16 ans.

La réforme a donc consisté non seulement à postposer d'un an le moment de l'orientation en filières, mais aussi à fonder un véritable enseignement secondaire inférieur général, non différencié. Elle s'est accompa-

gnée d'une réflexion importante sur le curriculum de ce nouveau gymnasium. "Celui-ci fixe des attentes élevées afin d'éviter que les enseignants ne continuent à enseigner de la même façon. C'était particulièrement important vu que les attentes étaient faibles pour les élèves supposés rejoindre l'enseignement professionnel de base" ³⁵ (OCDE, 2010, p. 225).

On imagine aisément l'intérêt que présente le cas polonais pour les questions qui nous occupent ici. Quelles retombées observe-t-on sur les acquis des élèves mesurés par PISA à la suite de ce changement de structure et de curriculum ?

Examinons d'abord ce qu'il en est en matière d'inégalités sociales. Compte tenu de tous les résultats présentés dans ce chapitre, on peut s'attendre à ce que les inégalités liées à l'origine sociale se réduisent. C'est effectivement le cas. Le pourcentage de la variation de la performance entre écoles se réduit entre 2000 et 2009 de 79 %. La Pologne qui figurait en 2000, avec 62 % de variance entre écoles, parmi les systèmes où la ségrégation académique était la plus marquée, n'affiche plus en 2009 que 16 % de variance entre écoles, soit une valeur proche de celle des pays nordiques. Par ailleurs, les clivages entre établissements en fonction de l'origine sociale se réduisent, tandis que la mixité sociale à l'intérieur des établissements va croissant (OCDE, 2011b). La réduction est spectaculaire, mais compréhensible : lorsque la Pologne bascule d'un système fortement différencié vers une structure intégrée et bénéficie en outre de l'avantage stratégique de créer de nouvelles écoles *ex nihilo*, vierges de toute histoire, une forte diminution de la variance entre écoles est en effet prévisible.

Qu'en est-il maintenant sur le plan des performances ? Les meilleurs élèves, comme on le craint souvent dans le grand public, ont-ils pâti de côtoyer au quotidien les élèves plus faibles qui dans l'ancien système, étaient écartés de la filière la plus exigeante ? La réponse est sans conteste négative : non seulement la performance moyenne du pays en lecture n'a pas diminué, elle a même sensiblement augmenté (+21 points entre 2000 et 2009, soit un cinquième d'écart-type), tandis que le pourcentage d'élèves peu performants diminuait fortement (moins 8,2 %) et que la proportion d'élèves très performants augmentait légèrement (+1,6 %). La valeur du percentile 5 passe de 304 à 346 (augmentation d'un demi-écart-type) et celle du percentile 95 augmente d'un dixième d'écart type (de 630 à 640). La suppression des filières a bien un effet égalisateur : les élèves très performants progressent, mais moins que les plus faibles, et les écarts se réduisent en conséquence (Lafontaine et Baye, 2012).

Le cas de la Pologne est particulièrement éloquent : **lorsqu'un même système éducatif décide de basculer d'un système différencié avec filières hiérarchisées à un système intégré, les inégalités sociales s'amenuisent, les différences de performances entre écoles s'effondrent, les élèves les plus performants ne le sont pas moins et les élèves à la traîne sont en nette diminution.** Le bénéfice pour le système éducatif est donc positif sur toute la ligne.

La force particulière de l'étude genevoise et l'analyse des évolutions qu'a connues la Pologne - qui les rend particulièrement convaincantes - est que la comparaison synchronique à Genève et diachronique en Pologne des deux modes d'organisation et de leurs effets porte sur un même système éducatif. La démonstration des effets vertueux d'un système intégré ne peut dès lors être écartée au motif qu'il s'agirait de systèmes éducatifs différents, culturellement ou politiquement trop éloignés pour que l'on puisse s'en inspirer ailleurs.

35. "This set high expectations to prevent teachers from teaching the same way they always had. This was especially important given the low expectations for students who were assumed to be going on to the basic vocational school". (OCDE, 2010, p. 225).

Alors que les recherches menées notamment par [Monseur et Lafontaine \(2012\)](#) rapportées ci-avant et les analyses de l'OCDE utilisent un indice continu de différenciation/ségrégation, [Mons \(2007\)](#) a proposé de son côté une typologie qui distingue un modèle de la séparation (filiales précoces) et trois modèles de l'intégration (tronc commun long) s'opposant sur plusieurs points. Les modèles de l'intégration *individualisée* (pays scandinaves et certains pays asiatiques) et de l'intégration *à la carte* (pays anglo-saxons essentiellement) autorisent ou valorisent une différenciation soit au niveau des groupes soit des individus, pour autant que les groupes de niveau ne constituent pas des classes de niveau permanentes, rigides et hiérarchisées. Ces modèles sont sous-tendus par un idéal d'égalité des acquis auquel s'ajoute, dans le modèle *à la carte*, une visée d'excellence pour les meilleurs élèves, qui bénéficient de cours avancés dans certaines matières. Ces deux derniers modèles permettent donc une forme de différenciation à l'intérieur du tronc commun. En revanche, le modèle d'intégration *uniforme* (France et pays du sud de l'Europe) applique un principe strict d'égalité de traitement et fait peu appel à l'individualisation et aux groupes de niveau temporaire dont l'efficacité est mise en évidence par les recherches en éducation ([Crahay et Wanlin \(2012\)](#)), voir p. 12 du présent rapport). Dès lors, les pays qui s'inscrivent dans ce modèle continuent à recourir à la manière classique de gérer les inévitables différences d'aptitude et de rythme des élèves en pratiquant le redoublement, dont le manque d'efficacité et le caractère inégalitaire sont bien connus.

[Mons \(2007\)](#) a utilisé les données de PISA 2000 pour étudier le lien entre les quatre modèles de sa typologie et différents indicateurs d'efficacité et d'équité : le score moyen du pays dans PISA, l'écart-type des performances (inégalités scolaires), le pourcentage d'élèves très peu performants d'une part, très performants d'autre part, la variance entre écoles, et enfin les écarts en fonction de l'origine socio-économico-culturelle (inégalités sociales).

Une série de constats largement concordants avec ceux des études déjà présentées ressortent de ses analyses. Ainsi, "plus la première orientation est repoussée dans le temps, plus le niveau général des élèves est élevé, et le nombre d'élèves en difficulté réduit. Les inégalités interélèves et interétablissements sont faibles et les disparités sociales de réussite scolaire sont peu élevées. Les résultats des différents modèles économétriques montrent que cette caractéristique majeure de l'organisation scolaire est davantage associée aux inégalités scolaires qu'à l'efficacité des systèmes éducatifs" ([Mons, 2007](#), p. 126).

Le modèle de la séparation (sans tronc commun, avec filiales précoces) affiche des performances moyennes à médiocres, ainsi que des inégalités scolaires et sociales fortes. "Alors que la sélection précoce et l'orientation des élèves dans des voies adaptées à leur profil sont censées maximiser les rendements scolaires, ce système est associé à un niveau général faible" ([Mons, 2007](#), p. 126). De surcroît, la proportion d'élèves en difficulté y est alarmante et l'élite peu développée, malgré l'existence de filiales "d'excellence". Ainsi, "le modèle de la séparation se révèle moins efficace que certains modèles d'intégration dans la production d'une élite scolaire, à niveau de développement économique comparable" ([Mons, 2007](#), p. 128).

Toutefois, l'un des modèles d'intégration, le modèle d'intégration *uniforme* (France et pays du sud de l'Europe) présente des performances moyennes plus faibles encore. Il s'illustre par un pourcentage important d'élèves en grande difficulté et ne produit pas une élite scolaire importante. C'est dans ce modèle que le taux de sorties sans diplômes ou qualification est le plus élevé. L'avantage tout relatif du tronc commun

avec modèle *uniforme* est qu'il produit a priori moins d'inégalités sociales que le modèle de la séparation³⁶. "Le système français n'a pas recours aux outils novateurs qui permettent une gestion efficace des disparités entre élèves, il s'appuie sur les instruments traditionnels de gestion de l'hétérogénéité, à savoir le redoublement et la création de classes de niveau dans le secondaire, en dépit de l'interdiction qui en était faite dans la loi Haby" (Mons, 2007, p. 133). Les modèles de l'intégration *individualisée* (pays scandinaves, Japon, Corée) et *à la carte* (pays anglo-saxons) ont en commun plusieurs caractéristiques : ils ne pratiquent pas ou peu le redoublement, l'individualisation y est importante, ne concerne pas les seuls élèves en difficulté et dépasse donc de simples activités de remédiation, le regroupement par classes de niveau n'existe pas dans le primaire. Alors que le modèle de l'intégration *individualisée* table essentiellement sur l'individualisation, le modèle *à la carte* propose en secondaire une offre différenciée de cours en fonction du niveau des élèves, dans chaque discipline. "Cette organisation souple ne conduit pas à l'organisation de classes de niveau permanentes, mais à la construction de groupes de niveau flexibles et évolutifs dans chacune des disciplines" (Mons, 2007, p. 120). Au final, les résultats de ces deux modèles qui évitent tout regroupement par aptitude rigide ou pérenne, sont assez semblables : les taux d'abandon ou les sorties sans diplômes y sont faibles, le niveau de performances élevé. C'est le système *à la carte* qui se révèle le plus performant, en grande partie grâce à la présence d'une élite importante ; en revanche, le système d'intégration *individualisée* est plus égalisateur. Les écarts en fonction de l'origine sociale y sont moindres. Comme le souligne Mons, "chaque modèle d'intégration est caractérisé par des résultats académiques spécifiques voire contradictoires" (Mons, 2007, p. 128).

La typologie de Mons (2007) et les travaux qui en découlent sont d'un grand intérêt pour nourrir la réflexion sur les politiques éducatives. Il ne suffit en effet pas d'en appeler à la mise en place ou à la prolongation d'un tronc commun pour que des résultats positifs en découlent. **L'adoption d'un tronc commun n'évacue pas la question de la gestion de l'hétérogénéité, ni celle de la différenciation et des formes qu'elle peut adopter.** "Selon les modalités de sa mise en œuvre, l'école unique peut être tour à tour performante ou inefficace, égalitaire ou lieu de production de nouvelles disparités sociales. Tout autant que le concept, ce sont donc les politiques éducatives de gestion de l'hétérogénéité mobilisées pour atteindre l'objectif qui importent. Plus que le concept, c'est sa mise en œuvre concrète sur laquelle il faut se focaliser" (Mons, 2007, p. 133).

Les résultats de Mons, faut-il le souligner, sont en parfaite cohérence avec ceux des trois premières sections de ce rapport. Les effets des mécanismes de différenciation que sont le redoublement et les classes de niveau ou filières, sur le plan des acquis scolaires et sur le plan socio-affectif, mais aussi sur les occasions d'apprendre et les environnements d'apprentissage sont particulièrement puissants. Les recherches expérimentales en ont montré l'inefficacité et les recherches en milieu naturel le caractère amplificateur sur les inégalités. Un système éducatif qui utilise ces deux modalités de différenciation "rigides", peu adaptées aux particularités et besoins spécifiques des élèves, comme modalité principale de gestion des inévitables différences d'aptitude et de rythme d'apprentissage (obéissant à un modèle d'intégration *uniforme* ou à un modèle de la *séparation*), renonce *de facto* à mettre au cœur de ses dispositifs pédagogiques des modalités de différenciation plus souples, mieux adaptées, et plus efficaces comme le font les modèles d'intégration

36. Pour rappel, les analyses de Mons (2007) se fondent sur les données PISA 2000. En France, l'ampleur des inégalités sociales a cru fortement entre 2000 et 2012. Des analyses sur des données plus récentes pourraient conduire à de nouvelles conclusions.

individualisée et à la carte.

VI Différencier, oui, mais comment ? Questions, tensions, débats.

Dans l'enseignement, faut-il ou non différencier, la question n'est au final qu'oratoire. Au vu de la diversité des publics d'apprenants, au vu de leurs différences d'aptitudes et de rythmes d'apprentissage, différencier s'impose. La véritable question est celle du comment différencier, étroitement articulée à celle du pour quoi différencier - à quelles fins, pour atteindre quels buts ? Viennent ensuite la question du degré et des modes de différenciation et de leurs effets sur les plans cognitif et non cognitif : la différenciation, est-ce efficace et est-ce juste ?

Cette synthèse montre à l'envi combien la question de la différenciation et de ses effets a été investiguée par les chercheurs en sciences de l'éducation. Il est peu de questions qui aient été aussi étudiées, et pour lesquelles on dispose de résultats de recherche aussi solides, reposant sur des approches diversifiées et menées dans une variété de contextes : recherches quasi expérimentales, recherches en milieu naturel, enquêtes internationales, recherches à caractère plus théorique débouchent sur un ensemble de conclusions remarquablement concordantes.

Sur le plan scientifique, on connaît donc, et depuis relativement longtemps, les effets de la différenciation et de ses principaux mécanismes résumés ici en quelques points :

- le redoublement n'est efficace qu'à court terme ; à long terme, il se révèle inefficace et préjudiciable sur le plan socio-affectif (motivation, engagement scolaire). Il frappe plus souvent les jeunes d'origine défavorisée, d'une manière injuste dans un certain nombre de pays (à compétences égales, les jeunes d'origine défavorisée sont davantage en retard scolaires). Le redoublement, avec les filières précoces, est l'un des principaux vecteurs d'amplification des inégalités scolaires liées à l'origine sociale (Monseur et Lafontaine, 2009 ; OCDE, 2016) ;
- Selon les études expérimentales, regrouper les élèves par aptitude en classes de niveau (*ability-grouping*) ou travailler en classes hétérogènes ne fait aucune différence. De surcroît, le regroupement par niveau n'est pas plus favorable pour certaines catégories d'élèves, à condition que la qualité et la quantité des apprentissages soient contrôlées, condition qui ne se réalise qu'en situation... expérimentale ;
- Sur le terrain, en milieu naturel, le regroupement en classes de niveau, en proposant des groupes ou cours plus avancés ou poussés aux meilleurs élèves, amplifie les écarts entre les classes fortes et les classes plus faibles, et creuse les inégalités liées à l'origine sociale. Les élèves forts ou moyens ont donc intérêt, dans un système différencié, à rejoindre les classes les plus fortes ou les plus exigeantes ;
- L'abondante littérature consacrée aux effets de la composition des écoles, filières ou classes sur le concept de soi (BFLPE) apporte un éclairage un peu différent sur cette question. Pour certains élèves, le fait de fréquenter une école ou une classe performante peut avoir une incidence négative sur leur concept de soi, et les conduire à sous-estimer leurs capacités scolaires. ;
- À l'aune de l'enquête PISA, les systèmes éducatifs les plus différenciés, présentant de grandes variations des performances et de composition sociale du public entre écoles, sont légèrement moins performants que les systèmes compréhensifs ou intégrés. Dans les premiers, les écarts entre les élèves

les plus et les moins performants sont davantage marqués, et les inégalités liées à l'origine sociale nettement plus importantes. La différenciation de structures a un impact léger à modéré sur les performances, mais elle accroît substantiellement les inégalités scolaires et sociales. C'est aussi ce que montre clairement le suivi de cohorte dans le canton de Genève : le système intégré tire les élèves vers le haut, et réduit les inégalités sociales, en donnant davantage de chances de poursuivre des études supérieures aux élèves les moins favorisés. L'évolution qu'a connue la Pologne à la suite de l'instauration d'un tronc commun long va dans le même sens : on y observe une forte réduction des inégalités de performances entre écoles et une augmentation du niveau moyen de performances.

Venons-en maintenant à la question plus précise des formes ou modalités de différenciation. Tant les recherches expérimentales, les études en milieu naturel, que les travaux menés au départ de la typologie de Mons (2007) montrent que les formes de différenciation les plus "rigides" ou "uniformes", qui clivent les groupes d'élèves et différencient leur parcours en fonction de leurs aptitudes, pour le long terme (*via* le redoublement et/ou les filières) ne conduisent ni à élever le niveau de performances moyen d'un système éducatif, ni - de manière plus surprenante - à produire une "élite" plus importante. La forme de différenciation proposée (redoublement ou orientation vers une filière académiquement moins exigeante) réputée efficace aux yeux des équipes éducatives, bloque la mise en place d'autres formes de différenciation plus souples (groupes de niveau temporaires, cours de niveau normal ou avancé, individualisation plus poussée prenant une autre forme que de la simple remédiation pour les élèves en difficulté) qui sont quant à elles bénéfiques non seulement pour les élèves en difficulté, mais aussi pour les meilleurs élèves. Croire aux vertus du "prêt-à-porter" pédagogique empêche en somme d'envisager le "sur-mesure" ou les mesures à *la carte* dont tous les élèves ont besoin pour progresser compte tenu de leurs capacités et leurs acquis antérieurs. Devant de telles évidences scientifiques, on peut se demander pourquoi les systèmes éducatifs hésitent encore à renoncer au redoublement, aux filières précoces et à adopter une organisation intégrée avec des formes souples de différenciation, qui semblent ne présenter que des avantages ? Sans doute ces connaissances scientifiques sont-elles encore insuffisamment connues des décideurs politiques, du terrain et du grand public. Si ce n'était qu'une question de connaissances, ce serait simple : il suffirait d'informer et de former les enseignants et autres acteurs du monde éducatif. Mais ce n'est bien entendu pas qu'une question de connaissances : c'est aussi une question de valeurs, comme toujours en éducation. À cet égard, il serait naïf de penser que les pays scandinaves ou la Pologne ont fait l'économie des débats vifs et des habituelles appréhensions quant à une possible baisse de niveau. Tous les systèmes éducatifs qui se sont lancés dans une unification de structure - qui augmente *de facto* l'hétérogénéité des écoles ou des classes - sont passés par les mêmes débats et les mêmes questionnements. Ils ont dû basculer d'une logique de différenciation/séparation, qui gère les difficultés d'apprentissage en séparant les publics d'élèves et en les orientant sur des voies différentes, vers une logique d'intégration qui implique une gestion plus directement pédagogique des difficultés d'apprentissage et de l'inévitable hétérogénéité des acquis, des rythmes et des aptitudes. Au final, ce ne sont pas seulement les arguments scientifiques qui les ont fait emprunter cette voie, même si ces derniers ont pu y contribuer, mais bien des convictions ou valeurs politiques fortes³⁷ qui ont conduit ces systèmes éducatifs à laisser derrière eux les réticences liées à la peur du mélange pour aller de l'avant, vers une école plus juste et tout aussi efficace.

37. Le meilleur exemple en est sans doute la Pologne qui, aujourd'hui dirigée par un gouvernement très conservateur, envisage de revenir à l'ancien système, en dépit des indubitables avancées observées dans PISA.

Bibliographie

- Allen, C., Q. Chen, V. Willson, et J. Hughes (2009). Quality of research design moderates effects of grade retention on achievement : a meta-analytic multilevel analysis. *Educational Evaluation and Policy Analysis* 4(31), 480–499.
- Blanton, H., B. Buunk, F. Gibbons, et H. Kuyper (1999). When better-than-others compare upward : Choice of comparison and comparative evaluation as independent predictors of academic performance. *Journal of Personality and Social Psychology* (76), 420–430.
- Communauté française (2009). Annuaire statistique de l'enseignement de plein exercice et budget des dépenses d'enseignement (2007-2008). Technical report, Communauté française et Service des statistiques de l'ETNIC, Bruxelles. <http://www.statistiques.cfwb.be>.
- Crahay, M. et P. Wanlin (2012). *Comment gérer l'hétérogénéité des élèves?* (Deuxième édition ed.). Bruxelles : De Boeck Université. In Crahay, M. (sous la direction de), *L'école peut-elle être juste et efficace?*
- Dong, Y. (2010). Kept back to get ahead? kindergarten retention and academic performance. *European Economic Review* (54), 219–236.
- Duflo, E., P. Dupas, et M. Kremer (2011). Peer effects, teacher incentives, and the impact of tracking : Evidence from a randomized evaluation in kenya. *American Economic Review* 101(5), 1739–1774.
- Dupont, V. et D. Lafontaine (Soumis). Fréquenter une école très performante n'a pas que des vertus : impact de l'école ou de la classe fréquentée sur le concept de soi (big-fish little pond effect). *Revue française de pédagogie*.
- Dupriez, V. et X. Dumay (2006). Inequalities in school systems : effect of school structure or of society structure? *Comparative Education* 2(42), 243–260.
- Eurydice (2011). *Le redoublement dans l'enseignement obligatoire en Europe. Réglementations et statistiques*. Bruxelles : Education, audiovisuel et culture, EACEA P9 Eurydice.
- Felouzis, G., S. Charmillot, et B. Fouquet-Chauprade (2011). Les inégalités scolaires en suisse et leurs déclinaisons cantonales : l'apport de l'enquête pisa 2003. *Swiss Journal of Sociology* 1(37), 33–55.
- Forget, A. (2017, avril). La différenciation dans l'enseignement : état des lieux et questionnement. *Publication du Cnesco dans le cadre de la conférence de consensus sur la différenciation pédagogique des 7 et 8 mars 2017*.

- Gamoran, A. (2002). Standards inequality and ability grouping in school. *Center for educational sociology Briefing* (25), 1–4.
- Gary-Bobo, R. et J. Robin (2011). La querelle des redoublements : analyse économique et problèmes statistiques. http://ces.univ-paris1.fr/membre/Gary-Bobo/RGB_Robin4Final.pdf.
- Goos, M. (2013). *Grade retention. The role of the national educational policy and the effects on students' academic achievement, psychosocial functioning, and school career*. Ph. D. thesis, Katholieke Universiteit Leuven.
- Grisay, A. (2001). Évaluer les dispositifs de prise en charge d'élèves faibles (ou forts) : l'utilisation de groupes naturels entraîne des artefacts. *Cahiers du Service de Pédagogie expérimentale - Université de Liège* (7-8).
- Hanushek, E. et L. Woessmann (2010). The economics of international differences in educational achievement. IZA DP 4925, Institute for the Study of Labor, Bonn.
- Hattie, J. (2009). *Visible learning. A synthesis of over 800 meta-analyses relating to achievement*. New York : Routledge.
- Holmes, C. (1990). *Grade Level Retention Effects : A Meta-Analysis of Research Studies*. Bristol : Falmer Press. In L. A. Shepard & M. L. Smith (Eds.), *Flunking Grades. Research and Policies on Retention*.
- Hong, G. et S. Raudenbush (2005). Effects of kindergarten retention policy on children's cognitive growth in reading and mathematics. *Education Evaluation and Policy Analysis* 3(27), 205–224.
- Huguet, P., F. Dumas, J.-M. Monteil, et N. Genestoux (2001). Social comparison choices in the classroom : Further evidence for students' upward comparison tendency and its beneficial impact on performance. *European Journal of Social Psychology* (31), 557–578.
- Husén, T. (1967). *International Study of Achievement in Mathematics*, Volume II. New York et Stockholm : John Wiley et Almqvist & Wiksell.
- Jacob, B. et L. Lefgren (2004). Remedial education and student achievement : A regression-discontinuity analysis. *The Review of Economics and Statistics* 1(86), 226–244.
- Jacob, B. et L. Lefgren (2009). The effect of grade retention on high school completion. *American Economic Journal : Applied Economics* 3(1), 33–58.
- Janowski, A. (2007). Educational restructuring and change : post-communist educational transformation in poland. *Orbis Scholae* 2(1), 80–109.
- Jimerson, S. (2001). Meta-analysis of grade retention research : Implications for practice in the 21st century. *School Psychology Review* 3(30), 420–437.
- Jonkmann, K., M. Becker, H. Marsh, O. Lüdtke, et U. Trautwein (2012). Personality traits moderate the big-fish-little-pond effect of academic self-concept. *Learning and Individual Differences* (22), 736–746.

- Kerckhoff, A. (1986). Effects of ability grouping in british secondary schools. *American Sociological Review* (51), 842–858.
- Kulik, J. A. (2004). *Meta-analytic studies of acceleration*, Volume 2. Iowa City : The University of Iowa. In N. Colangelo, S. G. Assouline & M. U. M. Gross (Eds). A nation deceived. How schools hold back American brightest students.
- Lafontaine, D. et A. Baye (2012). Pisa, instrument ou témoin du changement : évolution des performances en lecture et des politiques éducatives dans cinq systèmes européens. *Éducation Comparée* (7), 59–101. <http://hdl.handle.net/2268/129225>.
- Lafontaine, D., A. Baye, S. Vieluf, et C. Monseur (2015). Equity in opportunity-to-learn and achievement in reading : A secondary analysis of pisa 2009 data. *Studies in Educational Evaluation* (47), 1–11.
- Lüdtke, O., O. Köller, H. Marsh, et U. Trautwein (2005). The frame of reference and the big-fish-little-pond effect. *Contemporary Educational Psychology* (30), 263–285.
- Legrand, L. (1995). *Les différenciations de la pédagogie*. Paris : PUF.
- Liem, G., H. Marsh, A. Martin, D. McInerney, et A. Yeung (2013). The big-fish-little-pond effect and a national policy of within-school ability streaming : Alternative frames of reference. *American Educational Research Journal* 2(50), 326–370.
- Marcoux, G., F. Boraita, et M. Crahay (À paraître). À propos de la structuration, de l'enracinement culturel et de la modifiabilité des croyances des enseignants sur le redoublement : synthèse d'un programme de recherche fns. *Revue Suisse des Sciences de l'Éducation*.
- Marsh, H., A. Abduljabbar, A. Morin, P. Parker, F. Abdelfattah, B. Nagengast, et M. Abu-Hilal (2015). The big-fish-little-pond effect : Generalizability of social comparison processes over two age cohorts from western, asian, and middle eastern islamic countries. *Journal of Educational Psychology* 1(107), 258–271.
- Marsh, H. et K. Hau (2003). Big-fish-little-pond effect on academic self-concept : A cross-cultural (26 countries) test of the negative effect of academically selective schools. *American Psychologist* 5(58), 364–376.
- Marsh, H., H. Kuyper, A. Morin, et P. Parker (2014). Big-fish-little-pond social comparison and local dominance effects : Integrating new statistical models, methodology, design, theory and substantive implications. *Learning and Instruction* (33), 50–66.
- Marsh, H. et A. O'Mara (2010). Long-term total negative effects of school-average ability on diverse educational outcomes. direct and indirect effects of the big-fish-little-pond effect. *Zeitschrift für Pädagogische Psychologie* 1(24), 51–72.
- Marsh, H. W., O. Köller, et J. Baumert (2001). Reunification of east and west german school systems : Longitudinal multilevel modeling study of the big-fish-little-pond effect on academic self-concept. *American Educational Research Journal* (38), 321–350.

- Marsh, H. W., C.-K. Kong, et K.-T. Hau (2000). Longitudinal multilevel models of the big-fish-little-pond effect on academic self-concept : Counterbalancing contrast and reflected-glory effects in hong kong schools. *Journal of Personality & Social Psychology* (78), 337–349.
- Marsh, H. W., H. Kuyper, A. Morin, P. Parker, et M. Seaton (2014). Big-fish-little-pond social comparison and local dominance effects : Integrating new statistical models, methodology, design, theory and substantive implications. *Learning and Instruction* (33), 50–66.
- Marsh, H. W., U. Trautwein, O. Lüdtke, O. Köller, et J. Baumert (2005). Academic self-concept, interest, grades, and standardized test scores : Reciprocal effects models of causal ordering. *Child Development* 2(76), 397–416.
- McKenna, M., K. Conradi, C. Lawrence, B. Jang, et J. Meyer (2012). Reading attitudes of middle school students : Results of a u.s. survey. *Reading Research Quarterly* 3(4), 283–306.
- Müller, C. M. et V. Hofmann (2014). Does being assigned to a low track negatively affect psychological adjustment. a longitudinal study in the first year of secondary school. *School Effectiveness and School improvement* 2(27).
- Mons, N. (2007). *Les nouvelles politiques éducatives : la France fait-elle les bons choix ?* Paris : PUF.
- Mons, N. (2009). Effets théoriques et réels des politiques d'évaluation standardisée. *Revue française de pédagogie* (169), 99–139.
- Monseur, C. et D. Lafontaine (Eds.) (2007). *Why do non-cognitive variables better predict mathematics achievement in some countries than in others ? A methodological study on PISA 2003*. Communication non publiée à la 12th Biennial Conference of the European Association for Research on Learning and Instruction (EARLI).
- Monseur, C. et D. Lafontaine (2009). *L'organisation des systèmes éducatifs : quel impact sur l'efficacité et l'équité ?* Bruxelles : De Boeck. In V. Dupriez & X. Dumay (Eds.), *L'efficacité en éducation, promesses et zones d'ombre*.
- Monseur, C. et D. Lafontaine (2012). *Structure des systèmes éducatifs et équité : un éclairage international*. Bruxelles, Belgique : De Boeck. In M. Crahay (Ed.), *Pour une école juste et efficace*.
- Nagengast, B. et H. Marsh (2011). The negative effect of school-average ability on self-concept in the uk, the uk countries and the world : the big-fish-little-pond-effect for pisa 2006. *Educational Psychology* 5(31), 629–656.
- Nagengast, B. et H. Marsh (2012). Big fish little ponds aspire more : mediation and cross-cultural generalizability of school-average ability effects on self-concept and career aspirations in sciences. *Journal of Educational Psychology* 4(104), 1033–1053.
- Oakes, J., A. Gamoran, et R. N. Page (1992). *Curriculum differentiation : Opportunities, outcomes, and meanings*. New York : Macmillan. In P. W. Jackson (Ed.), *Handbook of research on curriculum : A project of the American Educational Research Association*.

- OCDE (2010). *Strong Performers and Successful Reformers in Education - Lessons from PISA for the United States*. Paris : OCDE.
- OCDE (2011a). *Résultats du Pisa 2009 : savoirs et savoir-faire des élèves*. Paris : OCDE.
- OCDE (2011b). *Résultats du PISA 2009 : Surmonter le milieu social : l'égalité des chances et l'équité du rendement de l'apprentissage*. Paris : OCDE.
- OCDE (2016). *Low-Performing Students : Why they fall behind and how to help them succeed*. Paris : OCDE. <http://dx.doi.org/10.1787/9789264250246-en>.
- O'Dwyer, L. (2002). *Extending the Application of Multilevel Modeling to Data from TIMSS*. London : Kluwer Academic Publishers. In F. Robitaille, & A.E. Beaton (Eds.). *Secondary Analysis of the TIMSS data*.
- Ou, S. et A. Reynolds (2010a). Grade retention, postsecondary education, and public aid receipt. *Educational Evaluation and Policy Analysis* 1(32), 118–139.
- Ou, S. et A. Reynolds (2010b). Grade retention, postsecondary education, and public aid receipt. *Educational Evaluation and Policy Analysis* 1(32), 118–139.
- Preckel, F. et M. Brüll (2010). The benefit of being a big fish in a big pond : Contrast and assimilation effects on academic self-concept. *Learning and Individual Differences* (20), 522–531.
- Schmidt, W. H., N. A. Burroughs, P. Zoido, et R. T. Houang (2015). The role of schooling in perpetuating educational inequality : an international perspective. *Educational Researcher* 7(44), 371–386.
- Seaton, M., H. Marsh, et R. Craven (2010). Big-fish-little-pond effect : Generalizability and moderation-two sides of the same coin. *American Educational Research Journal* 2(47), 390–433.
- Shepard, L. et M. Smith (1989). *Flunking grades : Research and policies on retention*. London : The Falmer Press.
- Slavin, R. (1987a). Ability grouping and student achievement in elementary schools : A best-evidence synthesis. *Review of Educational Research* 3(57), 293–336.
- Slavin, R. (1987b). Grouping for instruction in the elementary school. *Educational Psychologist* (22), 109–127.
- Slavin, R. (1990). Achievement effects of ability grouping in secondary schools : a best evidence synthesis. *Review of Educational Research* 3(60), 471–499.
- Trautwein, U., O. Lüdtke, O. Köller, H. Marsh, et J. Baumert (2006). Tracking, grading, and student motivation : Using group composition and status to predict self-concept and interest in ninth-grade mathematics. *Journal of Educational Psychology* 4(98), 788–806.
- Trautwein, U., O. Lüdtke, H. Marsh, et G. Nagy (2009). Within-school social comparison : How students perceive the standing of their class predicts academic self-concept. *Journal of Educational Psychology* 4(101), 853–866.

van de Gaer, E., A. Grisay, W. Schlutz, et E. Gebhardt (2012). The reference group effect : An explanation of the paradoxical relationship between academic achievement and self-confidence across countries. *Journal of Cross-cultural Psychology* 8(43), 1205–1228.

Zell, E. et M. Alicke (2009). Contextual neglect, self-evaluation, and the frog-pond effect. *Journal of Personality and Social Psychology* 3(97), 467–482.



Cnesco

Carré Suffren

31-35 rue de la Fédération

75 015 Paris

cnesco.communication@education.gouv.fr

École normale supérieure de Lyon
Institut français de l'Éducation

19 allée de Fontenay

69 007 Lyon

conf.consensus.ife@ens-lyon.fr