

Évaluation des organisations des enseignements de français et de mathématiques en sixième et cinquième

Premiers résultats pour l'année scolaire 2024-2025

Bechichi Nagui, Charpentier Axelle, Cornetet Jules, Dextre Polo Marisol, El Bah Fatimetou (DEPP)

Barone Carlo (Sciences Po), Grenet Julien (CNRS, École d'économie de Paris), Gurgand Marc (CNRS, École d'économie de Paris), Guyon Nina (École Normale Supérieure-PSL, École d'économie de Paris), Huillery Elise (Université Paris Dauphine-PSL, École Normale Supérieure-PSL), Souali Laïla (Université Paris Dauphine-PSL), Voisin Camille (Sciences Po)

Série Études

Document de travail n° 2026-E17
Septembre 2026

Évaluation des organisations des enseignements de français et de mathématiques en sixième et cinquième

Premiers résultats pour l'année
scolaire 2024-2025

Direction de l'évaluation, de la prospective et de la performance

61-65, rue Dutot
75732 Paris Cedex 15

Directrice de la publication

Magda Tomasini

Auteur(s)

Bechichi Nagui, Charpentier Axelle, Cornetet Jules, Dextre Polo
Marisol, El Bah Fatimetou (DEPP)

Barone Carlo (Sciences Po), Grenet Julien (CNRS, École
d'économie de Paris), Gurgand Marc (CNRS, École d'économie
de Paris), Guyon Nina (École Normale Supérieure-PSL, École
d'économie de Paris), Huillery Elise (Université Paris Dauphine-
PSL, École Normale Supérieure-PSL), Souali Laïla (Université Paris
Dauphine-PSL), Voisin Camille (Sciences Po)

e-ISSN 2779-3532

SOMMAIRE



➤ Remerciements.....	9
➤ Synthèse (en français)	10
➤ Synthèse (en anglais)	12
➤ Introduction.....	14
Principaux résultats	15
Contribution de l'étude aux connaissances sur les pratiques de regroupement des élèves	16
➤ Revue de littérature	17
Des effets théoriques ambigus.....	17
Des effets observés globalement nuls sur les résultats scolaires dans le second degré	18
Des effets possiblement différenciés selon le niveau initial des élèves et selon l'allocation des ressources	21
Le rôle des effets de pairs pour les élèves les plus en difficulté ou les plus doués scolairement	22
Des politiques de regroupement difficiles à qualifier.....	24
Des résultats encourageants pour les formes les plus flexibles en cas de moyens conséquents investis	25
La question de la mixité sociale et scolaire à l'école	26
Les effets liés à une réduction de la taille du groupe.....	26
Une politique efficiente ?.....	27
Conclusion.....	27
➤ Cadre institutionnel.....	29
La différenciation pédagogique et la gestion de l'hétérogénéité en France.....	29
Le « Choc des savoirs »	29
Mission d'évaluation.....	31

➤ Données et taux de réponse	33
Matériel d'enquête	33
Données administratives	34
Données extraites des logiciels de vie scolaire.....	34
➤ Les organisations pédagogiques mises en œuvre en 2024-2025	35
Premiers éléments descriptifs	35
De nouvelles organisations pédagogiques dans la quasi-totalité des établissements à la rentrée 2024.....	35
La mise en place des organisations pédagogiques a nécessité des moyens supplémentaires	38
Des regroupements moins flexibles que dans les consignes de l'arrêté du 15 mars 2024.....	39
Une perception de la réforme initialement très défavorable chez les enseignants	41
Une amélioration nette de la perception de la réforme chez les enseignants à la fin de la première année	42
Un impact perçu comme négatif sur le rôle de professeur principal et le suivi des élèves.....	44
Une collaboration entre les enseignants perçue comme renforcée, en particulier en français.....	45
Après la réforme, la difficulté scolaire perçue comme mieux prise en charge en mathématiques.....	46
Plus de 70 % des établissements ont mis en place un seul type de regroupement	46
Présentation de la typologie.....	51
Des choix organisationnels qui varient selon la taille des établissements et leur indice de position sociale.....	51
Détail des modalités organisationnelles	53
Constitution des regroupements	54
Réduction des effectifs	54
Retour en classe de référence et flexibilité des regroupements.....	55
Analyse de l'hétérogénéité des réponses aux questionnaires « chefs d'établissement » selon la nature des regroupements mis en place	57
Une amélioration de la perception de la réforme quel que soit le type de regroupement, davantage marquée en cas de groupes hétérogènes	57
L'avis des chefs d'établissement sur les effets de la réforme dépend de la nature des groupes et du niveau des élèves	59
Les chefs d'établissement perçoivent que la collaboration entre élèves est favorisée par la mise en place de groupes hétérogènes	61
Des effets perçus comme défavorables sur la vie de l'établissement, sans différences marquées selon les dimensions considérées	62

➤ Les effets sur les apprentissages des élèves 63

Analyses corrélationnelles : évolution des résultats aux évaluations nationales63

Des résultats moyens globalement stables entre 2024 et 2025 dans les deux disciplines pour l'ensemble des établissements 63

Une évolution globale des scores aux évaluations nationales de 4^e qui ne semble pas dépendre du type de regroupement mis en place..... 66

Conclusion relative aux premières analyses corrélationnelles 68

Étude corrélationnelle en fonction du niveau initial de performance des élèves 68

Évolution des scores moyens en 4^e selon le niveau initial de performance en 6^e et le type de regroupement 68

Analyses causales..... 71

Méthodologie : Différences de différences71

Cadre général.....71

Première approche : comparaison des résultats des élèves à l'échelle des établissements, avant et après la mise en œuvre de la réforme..... 72

Seconde approche : comparaison des progressions scolaires d'une même cohorte d'élèves dans le temps 75

Résultats de la première approche : comparaison des élèves de la première cohorte exposée aux nouvelles organisations pédagogiques avec ceux de la cohorte précédente..... 77

Validité de l'hypothèse des tendances parallèles..... 77

Une absence d'effet des nouvelles organisations pédagogiques sur la réussite des élèves aux évaluations nationales de 4^e 82

Résultats de la seconde approche : comparaison de la progression des élèves selon leur scolarisation dans un établissement ayant introduit au non de nouvelles organisations pédagogiques en 2024-2025 85

Validité de l'hypothèse des tendances parallèles..... 85

Analyse événementielle..... 87

L'absence d'effet moyen des nouvelles organisations pédagogiques sur la réussite des élèves confirmée par la seconde approche 93

Discussion des limites associées aux résultats 95

➤ Conclusion 98

➤ Bibliographie100

➤ Annexes105

Annexe 1 : Protocole d'enquête 105

Cadre général 105

Définition de la base de sondage et de l'échantillon principal 105

Annexe 2 : Pondération initiale, non-réponse et redressement 107

Pondération initiale des établissements	107
Repondération liée au non-renvoi des formulaires de contact.....	107
Repondération liée à la non-réponse aux questionnaires adressés aux chefs d'établissement	108
Calage sur marges	108
Pondération initiale des coordonnateurs	110
Repondération des questionnaires adressés au coordonnateurs.....	110
Annexe 3 : Construction de la typologie des établissements	113
Annexe 4 : Évolution de la répartition des élèves dans les groupes de performance aux évaluations nationales entre 2024 et 2025	116
Annexe 5 : Évolution de la répartition des élèves dans les groupes de performance aux évaluations nationales entre la 6^e et la 4^e, selon le type de regroupement mis en œuvre en 2024-2025	119
Annexe 6 : Données des évaluations nationales et scores standardisés.....	121
Annexe 7 : Hétérogénéité des résultats selon le niveau des élèves à l'entrée en 6^e et leur sexe	124
Première approche	124
Seconde approche.....	130
Annexe 8 : Résultats par domaine évalué en français et mathématiques	142

➤ Remerciements

Les auteurs expriment leur profonde reconnaissance à l'ensemble des chefs d'établissement, adjoints et enseignants, qui ont participé aux enquêtes par questionnaire et ont transmis les données nécessaires au déploiement du protocole d'évaluation des organisations pédagogiques mises en œuvre à la rentrée scolaire 2024 pour l'enseignement du français et des mathématiques en 6^e et en 5^e.

Ils remercient chaleureusement les membres de la mission IGÉSR, pilotée par Bruno Claval et Christine Darnault, pour leur contribution précieuse à la conception du matériel d'enquête et à l'analyse des premiers résultats relatifs à la description des organisations retenues dans les collèges.

La qualité des travaux d'enquête et d'analyse tient à l'engagement fort de plusieurs stagiaires et vacataires qui, à différentes étapes du projet, ont participé au déploiement du protocole de recueil de données, au nettoyage des données et aux analyses : Sofiane Akli, Laura Berchoux, Théodore Bogaty, Mattéo Gauthier, Léo Gonzales, Hugo Heyninck, Patrick Mielle, Anouck Mille, Lou-Ann Toujas-Tort.

Enfin, l'ERC Starting Grant EMANCIPATE d'Elise Huillery (n°101040822, référence : ERC-2021-STG) a permis de financer pour partie le recrutement de vacataires pour la collecte de données organisée par la DEPP au printemps 2025.

▸ Synthèse (en français)

La rentrée scolaire 2024 marque la première année de mise en œuvre des nouvelles organisations pédagogiques en français et en mathématiques dans les classes de 6^e et de 5^e. L'étude sur l'organisation des enseignements au collège (EEC), menée par la DEPP auprès d'un échantillon représentatif de 1 500 collèges publics, a permis de collecter, grâce à un fort taux de participation, des données standardisées de qualité sur le déploiement concret du dispositif, les perceptions des chefs d'établissements et coordonnateurs des organisations pédagogiques en français et en mathématiques. L'étude réalisée en collaboration étroite avec une équipe de recherche constituée d'économistes et de sociologues spécialistes de l'évaluation des politiques éducatives mobilise également les résultats aux évaluations nationales de plusieurs cohortes d'élèves concernées ou non par les regroupements afin de mener une première évaluation causale des effets du dispositif sur les apprentissages grâce à deux approches : d'une part, en comparant à l'échelle des établissements les scores obtenus aux évaluations nationales de 4^e en français et en mathématiques pour la première cohorte concernée par le dispositif à ceux des cohortes l'ayant précédée ; d'autre part, en comparant la progression d'une année sur l'autre des élèves d'une même cohorte selon le type de regroupement auquel ils ont été exposés.

Au niveau descriptif, les résultats montrent qu'une grande majorité des établissements a mis en place en 2024-2025 une nouvelle organisation des enseignements (94 %), mais selon des modalités très diverses. Les types de regroupements retenus comportent d'une part des groupes hétérogènes sans lien avec le profil scolaire des élèves (21 % des collèges) et d'autre part des groupes formés en fonction du profil scolaire des élèves (50 % des collèges). Cette deuxième catégorie comprend plusieurs modalités : groupes homogènes (19 %), regroupements des élèves en difficulté (15 % des collèges), groupes constitués selon le niveau scolaire des élèves dans les deux disciplines confondues (8 %), groupes formés selon une compétence spécifique (7 %) et, marginalement, regroupements des élèves performants (1 % des collèges). Enfin, 23 % des collèges combinent plusieurs modalités de regroupement parmi celles mentionnées. Les chefs d'établissement déclarent une réduction des effectifs dans les groupes d'enseignement de français et de mathématiques dans la quasi-totalité des établissements du fait des nouvelles organisations pédagogiques ainsi que d'importants ajustements organisationnels, obligeant les établissements à mobiliser des moyens supplémentaires à ceux alloués pour constituer les groupes. Par ailleurs, les organisations pédagogiques effectivement mises en œuvre apparaissent souvent moins flexibles que celles prévues par les textes réglementaires. Les retours temporaires en classe de référence ainsi que les changements de groupe en cours d'année demeurent relativement rares et, lorsqu'ils existent, concernent généralement un nombre limité d'élèves. Malgré une perception très défavorable du dispositif au moment de son annonce, l'étude révèle une amélioration du jugement porté par les enseignants au cours de cette première année de mise en œuvre, qui pourrait en partie s'expliquer par la réduction des effectifs dans les groupes induite par la mise en place du dispositif. Les perceptions restent cependant contrastées et plutôt négatives concernant certains aspects organisationnels intrinsèques à l'établissement tels que le rôle du professeur principal ou le suivi individualisé des élèves.

Les premières analyses causales des effets sur les apprentissages des élèves scolarisés en classe de 5^e en 2024-2025 montrent que les nouvelles organisations n'ont pas eu d'effet significatif sur les scores en français et en mathématiques aux évaluations nationales de début de quatrième. En effet, les variations estimées entre les élèves de 5^e concernés par la réforme et ceux de l'année précédente sont faibles et d'une ampleur comparable aux variations moyennes observées entre cohortes avant la mise en œuvre de la réforme. Premièrement, les élèves concernés par les regroupements en 5^e en 2024-2025 ne réussissent ni mieux ni moins bien que les élèves de la cohorte précédente aux évaluations nationales de début de 4^e en français et mathématiques, et ce, quel que soit le type de regroupement adopté par les établissements qu'ils fréquentent en 5^e. Deuxièmement, lorsque l'on se concentre sur l'évolution des trajectoires des élèves de la première cohorte exposée au dispositif en 5^e, aucune des organisations pédagogiques retenues (groupes hétérogènes ou regroupements

fondés sur le profil scolaire des élèves) n'a eu d'effet statistiquement significatif sur les performances moyennes à l'entrée en 4^e, en comparaison d'établissements qui n'ont pas réalisé de regroupements. Les analyses complémentaires, réalisées en comparant les élèves selon leur niveau initial ou selon leur sexe aboutissent à des conclusions similaires, c'est-à-dire à une absence d'effet du dispositif sur les scores moyens.

Ces premiers résultats rejoignent les conclusions de la littérature scientifique internationale, qui tend à montrer une absence d'effet moyen des regroupements d'élèves sur les performances scolaires. Cependant, ils doivent être interprétés avec prudence en raison de l'hétérogénéité importante dans la mise en place de la réforme, i.e. dans la nature du dispositif évalué, défini ici par les types de regroupements adoptés, qui pourrait mener à des effets différenciés qui ne seraient pas détectés par la méthode utilisée. D'autres approches réalisées à l'échelle des élèves et de leur groupe d'appartenance seront menées dans de prochains travaux d'évaluation en mobilisant des données issues des logiciels de vie scolaire afin de proposer une estimation plus fine des effets des regroupements.

Le prolongement de ces premières analyses grâce aux données recueillies pour la deuxième année de mise en œuvre du dispositif (2025-2026) permettra d'affiner et de consolider l'estimation de ses éventuels effets.

➤ Synthèse (en anglais)

The start of the 2024 school year marks the first year of implementation of new instructional arrangements for French and mathematics for students in grades 6 (approximately 11–12 years of age) and 7. The Study on the Organization of Instruction in Middle School, conducted by the DEPP (the Department for Evaluation, Prospective and Performance, at the French Ministry of Education) among a representative sample of 1,500 public middle schools, yielded high-quality, standardized data thanks to the high participation rate. The data pertains to the practical implementation of the program and the perceptions of school principals and coordinating teachers regarding classroom organization in mathematics and French. In close collaboration with a research team of economists and sociologists specializing in the evaluation of educational policies, the study also draws on the results of national assessments of several student cohorts—both those affected and those unaffected by the grouping policy—to conduct an initial causal assessment of the program's effects on learning. This assessment uses two approaches. First, it compares the scores on the eighth-grade national assessments in mathematics and French for the first cohort subject to the program with those of the preceding cohort at the school level. Second, it compares the year-over-year progress of students within the same cohort based on the type of grouping to which they were exposed.

At a descriptive level, the results show that the vast majority of middle schools (94%) implemented a new teaching organization in 2024-2025, albeit with very different modalities. The adopted grouping types comprise, on the one hand, mixed-ability groups not based on students' academic profiles (21% of middle schools) and, on the other hand, groups formed according to students' academic profiles (50% of middle schools). This second category comprises several approaches: homogeneous groups (19%), groups of students with learning difficulties (15%), groups formed according to students' academic level across both subjects (8%), groups formed according to a specific skill (7%) and, to a lesser extent, groups of high-achieving students (1%). Finally, 23% of middle schools use a combination of the aforementioned grouping methods. School principals report a reduction in the size of teaching groups for French and mathematics in almost all schools due to the new pedagogical organizations. They also report significant organizational adjustments requiring schools to mobilize additional resources beyond those allocated for group formation. Furthermore, the instructional arrangements that are actually implemented often appear less flexible than those outlined in regulatory texts. Temporary gatherings with the "reference class" and changes in groups during the school year remain relatively rare. When they occur, they generally involve a limited number of students. Despite a highly unfavorable perception of the reform when it was first announced, the study reveals that the teachers' opinions have improved during this first year of implementation, which could be partly explained by the reduction in class sizes resulting from the program's rollout. However, perceptions remain mixed and predominantly negative regarding certain organizational aspects specific to the school, such as the role of the homeroom teacher or individualized student support.

The first causal analyses of the effects on learning for seventh-grade students in the 2024–2025 school year do not reveal a significant impact of the new organizational structures on scores in French and mathematics scores on national assessments. The estimated effects are weak and comparable in magnitude to the average variations observed between cohorts prior to the implementation of the reform. First, students affected by grouping in seventh grade in the 2024–2025 school year perform neither better nor worse than students in the previous cohort on the national assessments in French and mathematics administered at the beginning of eighth grade, regardless of their seventh-grade school's grouping method. Second, when focusing on the progression of students from the first cohort exposed to the program in seventh grade, none of the selected instructional arrangements (heterogeneous groups or groupings based on students' academic profiles) had a statistically significant effect on average performance upon entering eighth grade, compared to schools that did not implement such groupings. Further analyses, conducted by comparing students based on their initial level or gender, yield similar conclusions, namely the absence of effect on average scores.

These preliminary results are consistent with the findings of the international scientific literature, which tends to show an absence of average effects of student grouping on academic performance. However, these results must be interpreted with caution due to the significant heterogeneity in the implementation of the reform—that is, in the nature of the “treatment” being evaluated, defined here by the types of grouping adopted, which could lead to differentiated effects that are not detected by the method used. Further analyses conducted at the student and group levels, will be pursued in future evaluations using school management software data to provide a more detailed estimates of the effects of grouping.

Extending these initial analyses with data from the second year of the program’s implementation (2025–2026) will allow for a more refined and robust estimation of the program’s potential effects.

➤ Introduction

La question du regroupement des élèves selon leur niveau scolaire ou leurs besoins spécifiques occupe une place importante dans la recherche en éducation et dans les débats relatifs aux politiques éducatives. Les arguments en faveur de cette pratique reposent sur l'idée qu'elle permettrait d'améliorer les performances scolaires grâce à la différenciation des apprentissages *via* une meilleure gestion de l'hétérogénéité scolaire des élèves et une adaptation des pratiques pédagogiques aux besoins des élèves. Cependant, la littérature sur les regroupements d'élèves ne permet pas de dégager un consensus clair quant à leurs effets. Si certaines études expérimentales ont mis en avant des effets positifs des pratiques de regroupements selon le niveau dans des systèmes éducatifs de pays en développement, plusieurs revues de littérature soulignent une absence d'effet des regroupements fixes sur la réussite des élèves, ainsi que les risques qu'ils peuvent comporter en matière d'accroissement des inégalités, de stigmatisation ou de détérioration du sentiment de compétence. Néanmoins, la recherche sur le sujet se caractérise majoritairement par des études corrélationnelles, de qualité variable, et qui semblent indiquer des effets potentiellement différenciés selon les profils d'élèves. Les effets paraissent également dépendre fortement des modalités concrètes de mise en œuvre de la politique qui peuvent être très variées : est-ce que les élèves changent de groupe, à quelle fréquence et selon quels critères, comment les professeurs sont affectés aux groupes, comment ils sont formés pour adapter leur méthode d'enseignement à la taille et au type de regroupement, comment ils choisissent d'adapter les contenus des enseignements au type de regroupement, comment les différents regroupements sont présentés aux élèves et aux parents, ou encore comment les élèves sont réintégrés en classe de référence à la suite des regroupements.

La réforme des « groupes de besoins » a été conçue dans le cadre du « Choc des savoirs » comme un levier pour répondre à l'hétérogénéité scolaire dans le but d'élever le niveau des élèves. Annoncée en décembre 2023 pour une mise en place en septembre 2024 prévue dans l'ensemble des collèges pour les enseignements de français et de mathématiques en classes de 6^e et 5^e, la mesure vise à constituer des groupes selon les besoins des élèves identifiés par les professeurs. Le dispositif préconise que les élèves les plus en difficulté bénéficient de groupes à effectifs réduits et suggère un retour partiel en classe de référence afin de permettre une réévaluation des besoins et une mobilité des regroupements au cours de l'année. Cette réforme ne prévoyait pas que les enseignants reçoivent une formation particulière supplémentaire. Une extension des regroupements était initialement prévue pour les classes de 4^e et 3^e en septembre 2025, mais n'a finalement pas eu lieu. Le caractère obligatoire du dispositif prendra fin en septembre 2026 après deux ans d'implémentation.

Cette étude menée par la Direction de l'évaluation, de la prospective et de la performance (DEPP), en étroite collaboration avec l'Inspection générale de l'éducation, du sport et de la recherche (IGÉSR) et une équipe de chercheurs en sociologie et économie (École d'économie de Paris, École normale supérieure-PSL, Université Paris Dauphine-PSL, Sciences Po), vise à documenter précisément le déploiement du dispositif et ses modalités de mise en œuvre ainsi qu'à en évaluer les effets causaux sur les apprentissages des élèves. Elle s'appuie sur un échantillon national représentatif d'environ 1 500 collèges publics, dont les chefs d'établissement et coordonnateurs de français et de mathématiques ont été interrogés par questionnaires auto-administrés à deux reprises, au printemps 2025 et au printemps 2026. Grâce à un fort taux de participation, ces questionnaires ont permis de collecter des informations standardisées de qualité sur la mise en œuvre des regroupements, le ressenti des personnels et leurs perceptions des premiers effets. Ces données d'enquête sont complétées par une collecte de données administratives issues des logiciels de vie scolaire et par les données relatives aux résultats des élèves aux évaluations nationales de rentrée scolaire.

Après deux années de mise en œuvre, trois cohortes d'élèves ont été concernées par les nouvelles organisations pédagogiques. Les élèves entrés en classe de 6^e à la rentrée 2023 ont été concernés par la réforme au cours de l'année scolaire 2024-2025, lorsqu'ils étaient en classe de 5^e, c'est-à-dire pendant une seule année. Les élèves entrés en 6^e à la rentrée 2025 ont également été concernés par la réforme pendant une seule année, en classe de 6^e au cours de l'année scolaire 2025-2026. Entre ces deux cohortes, les élèves entrés en 6^e à la rentrée 2024 constituent l'unique cohorte ayant été concernée par les nouvelles organisations pédagogiques pendant deux années consécutives : en 6^e en 2024-2025 puis en 5^e en 2025-2026. Ce document de travail s'appuie sur la première collecte de données réalisée au printemps 2025, portant sur des informations relatives à l'année scolaire 2024-2025. Les analyses d'impact sur les apprentissages portent donc exclusivement sur la cohorte entrée au collège en septembre 2023, seule cohorte concernée par la réforme pour laquelle il est possible d'observer à ce jour les résultats des élèves au niveau national après la mise en œuvre de la réforme, en l'occurrence *via* les évaluations nationales de début de 4^e menées en septembre 2025. **Le présent document de travail analyse donc les effets pour des élèves concernés par les nouvelles organisations pédagogiques pendant un an en classe de 5^e.**

L'estimation des effets causaux des nouvelles organisations pédagogiques proposée ici repose sur deux approches causales complémentaires fondées sur la méthode des doubles différences. La première compare les résultats aux évaluations nationales de début de 4^e des élèves concernés par la réforme en 2024-2025 pendant leur 5^e, avec ceux des deux cohortes précédentes non concernées par la réforme, en confrontant l'évolution observée dans les établissements ayant adopté différents types de regroupements à celle observée dans les établissements n'ayant pas modifié leurs organisations pédagogiques. La seconde approche suit quant à elle les trajectoires scolaires individuelles de la cohorte d'élèves de 5^e en 2024-2025 depuis le début de leur scolarité jusqu'à leur entrée en 4^e, afin de comparer les progressions observées selon le type d'organisation pédagogique mis en œuvre dans le collège fréquenté. Cette approche mobilise les résultats aux évaluations nationales de CP, mi-CP, CE1, 6^e et 4^e de cette cohorte afin d'analyser les trajectoires scolaires individuelles avant et après l'exposition à la réforme. Ces deux approches s'appuient également sur les résultats aux évaluations nationales de 6^e afin d'analyser de potentiels effets différenciés selon le niveau initial des élèves en début de 6^e.

Principaux résultats

Les résultats descriptifs de l'évaluation de la première année de mise en œuvre des regroupements montrent que le dispositif a été déployé de manière très hétérogène, selon des formes organisationnelles diverses. Si une très grande majorité des établissements (94 %) a modifié ses organisations pédagogiques en 2024-2025 pour mettre en place des regroupements, les établissements se différencient fortement dans les modalités adoptées. Ainsi, on distingue des établissements ayant mis en place des groupes hétérogènes (élèves de tous niveaux au sein d'un même regroupement), qui représentent environ 20 % des collèges, et pour le reste, des groupes homogènes (élèves d'un même niveau au sein d'un regroupement), des groupes constitués selon le niveau des élèves dans les deux matières confondues, des groupes fondés sur les compétences des élèves, des groupes constitués uniquement pour les élèves les plus en difficultés ou uniquement pour les plus performants, ainsi que des établissements ayant combiné des modalités multiples de regroupements selon les niveaux et disciplines. On note par ailleurs que les chefs d'établissement déclarent une réduction des effectifs dans la quasi-totalité des établissements du fait des nouvelles organisations pédagogiques, une flexibilité très faible des regroupements mis en place par les établissements ainsi qu'un recours limité au retour en classe de référence, pourtant initialement préconisé par le dispositif.

Malgré une perception très défavorable du dispositif au moment de son annonce, l'étude révèle une amélioration du jugement porté par les enseignants au cours de cette première année d'implémentation, qui pourrait en partie s'expliquer par la réduction des effectifs dans les groupes induite par la mise en œuvre du dispositif. Les perceptions restent cependant contrastées et plutôt

négligentes concernant certains aspects organisationnels intrinsèques à l'établissement tels que le rôle du professeur principal ou le suivi individualisé des élèves.

Les analyses causales menées sur l'évolution des apprentissages des élèves scolarisés en 5^e en 2024-2025 ne font pas apparaître d'effets significatifs des regroupements sur les résultats des élèves exposés aux différentes formes du dispositif, par rapport à ceux scolarisés dans des établissements n'ayant pas mis en place de regroupements. Aucune différence statistiquement significative n'est observée dans les résultats aux évaluations nationales de français et de mathématiques à l'entrée en 4^e des élèves scolarisés en classe de 5^e en 2024-2025, que l'on compare les élèves de cette cohorte entre eux selon qu'ils ont été concernés ou non par les regroupements, ou qu'on les compare avec les cohortes qui ont précédé la mise en œuvre du dispositif.

Contribution de l'étude aux connaissances sur les pratiques de regroupement des élèves

Cette étude apporte une contribution importante aux connaissances sur les effets du regroupement des élèves selon leurs niveaux ou besoins au collège, une pratique dont les effets causaux restent incertains dans la littérature et encore peu documentés dans le contexte français. En mobilisant des informations déclaratives recueillies à l'échelle des établissements et les évaluations nationales exhaustives de français et de mathématiques, elle apporte un éclairage sur la diversité des organisations effectivement mises en place dans les collèges et sur les conditions concrètes de déploiement d'une réforme nationale laissant une marge importante d'autonomie aux établissements.

L'exploitation à venir des données issues des logiciels de vie scolaire permettra d'affiner les premiers résultats présentés dans ce document, en permettant notamment de connaître précisément la composition des différents groupes dans un établissement et de mesurer ainsi le degré d'hétérogénéité sociale et scolaire en leur sein. Les effets des regroupements sur la cohorte d'élèves entrés en 6^e en 2024-2025 et concernés par les regroupements pendant deux années consécutives pourront également être documentés, ainsi que les effets à plus long terme, en particulier sur les résultats aux épreuves écrites du diplôme national du brevet.

La première partie de ce document propose une revue de la littérature, qui souligne les effets incertains et différenciés du regroupement des élèves selon leur niveau ou leurs besoins. Le cadre institutionnel dans lequel s'inscrit la réforme est ensuite décrit avant de présenter en détail la méthodologie associée à l'enquête et les taux de réponse. La quatrième partie du document se concentre sur la description détaillée de la mise en œuvre des organisations pédagogiques dans les établissements et de sa réception. La cinquième et dernière partie présente les résultats des analyses économétriques portant sur l'évaluation causale des effets des regroupements sur les résultats des élèves aux évaluations nationales. Enfin, la conclusion synthétise la portée et les limites des premiers résultats présentés et propose des pistes pour approfondir l'évaluation du dispositif et de ses effets.

➤ Revue de littérature

Des effets théoriques ambigus

La théorie prédit que le choix de regrouper les élèves selon leur niveau scolaire ou leurs besoins spécifiques peut avoir des effets positifs, neutres ou négatifs sur les élèves selon leur position dans la classe et selon l'ampleur relative de différents effets pouvant jouer dans des sens contraires, tels que l'adaptation pédagogique à leurs besoins, les effets de pairs, ou encore le renforcement ou la détérioration du sentiment de compétence et de la motivation scolaire.

Les arguments en faveur du regroupement des élèves selon leur niveau reposent sur une conception de l'hétérogénéité scolaire au sein des classes qui serait préjudiciable à l'apprentissage des élèves : les élèves progresseraient davantage au sein de classes homogènes, c'est-à-dire constituées exclusivement d'élèves de profil scolaire similaire, en bénéficiant d'un enseignement adapté à leurs besoins spécifiques et d'enseignants plus disponibles car moins préoccupés par la gestion de leur classe. En particulier, les élèves les plus performants profiteraient d'enseignements plus exigeants leur permettant de progresser davantage, tandis que les élèves plus fragiles profiteraient d'enseignements plus simples adaptés à leur niveau et se sentiraient plus en confiance sans la concurrence avec des élèves ayant plus de facilités. Cette conception de l'hétérogénéité scolaire a inspiré de nombreuses politiques éducatives visant à homogénéiser le niveau des élèves au sein des classes, et en premier lieu les politiques de redoublement, de saut de classe ou de répartition des élèves en groupes de niveau. La présente revue de littérature porte plus spécifiquement sur cette dernière catégorie, en lien avec les objectifs qui sous-tendaient la mise en place de nouvelles organisations pédagogiques en groupes pour l'enseignement du français et des mathématiques en classes de 6^e et 5^e mises en place à la rentrée 2024 dans l'ensemble des collèges¹.

En revanche, d'autres arguments viennent questionner le regroupement des élèves selon leur niveau ou leurs besoins en raison de potentiels effets négatifs sur les apprentissages des élèves, notamment les plus fragiles. Un premier argument tient à la différenciation des apprentissages : les élèves les plus fragiles seraient ainsi confrontés à des contenus pédagogiques moins avancés et à des exigences moindres, ce qui pourrait avoir des répercussions négatives sur la suite de leur scolarité. Un second argument concerne l'allocation des ressources entre les groupes : le regroupement des élèves selon leur niveau peut aussi s'accompagner d'une évolution du profil d'enseignants affectés aux groupes faibles par rapport aux groupes forts, en termes de niveau d'expérience par exemple, avec un potentiel renforcement des écarts d'apprentissage. Un troisième argument a trait aux effets de pairs : les élèves pourraient bénéficier d'avoir des camarades de classe plus performants si cela induisait un environnement plus propice au travail. Dans ce cas, regrouper les élèves par niveaux ou besoins affecterait négativement les élèves les plus en difficulté ou moyens. Inversement, cela pourrait induire, pour les mêmes raisons, des effets positifs pour les élèves les plus à l'aise. Une dernière série d'arguments concerne l'importance du sentiment de compétence sur la motivation et l'engagement dans le travail scolaire : appartenir à un groupe faible pourrait réduire le sentiment de compétence et provoquer une stigmatisation conduisant à un désengagement du travail scolaire. D'anciennes études qualitatives et ethnographiques ont souvent mis en avant un potentiel effet négatif des groupes de niveau sur l'estime de soi des élèves appartenant aux groupes les plus faibles (Rosenbaum, 1980), et indiquaient que cet effet pouvait être particulièrement marqué pour les formes de regroupements les plus rigides et durables, au sein desquelles l'appartenance permanente à l'un des groupes finirait par influencer la manière dont l'élève perçoit ses capacités sur le plan scolaire. Archer et al. (2018) constatent en outre une possible internalisation d'un statut inférieur chez les élèves assignés aux groupes les moins performants qui expriment leur position comme méritée et légitime, les figeant dans une forme d'état d'esprit fataliste qui pourrait aussi affecter leur engagement et leur

¹ Ministère de l'Éducation nationale et de la Jeunesse, « Mettre en place les groupes de besoins. Faire réussir tous les élèves au collège », Vademecum, mai 2024.

motivation. De tels effets négatifs pourraient également se retrouver au sein des groupes constitués des élèves les plus performants en raison de l'augmentation des exigences et de l'intensification de la compétition entre pairs.

Globalement, les effets attendus des regroupements selon le niveau ou les besoins des élèves sont donc potentiellement complexes et variables selon le niveau initial des élèves et leur position relative, ce qui les rend particulièrement difficiles à anticiper. Selon l'ampleur relative des effets d'adaptation pédagogique, d'allocation des moyens, de pairs, et de motivation, les effets d'une telle politique peuvent ainsi s'avérer positifs, neutres ou négatifs en moyenne sur l'ensemble des élèves, et inégalement répartis selon le profil scolaire des élèves, avec des conséquences potentiellement positives sur les inégalités scolaires.

Certaines études expérimentales confirment que la réorganisation du système éducatif en regroupant les élèves selon leur niveau permettrait au professeur d'adapter ses enseignements aux besoins et capacités spécifiques des élèves, ce qui favoriserait l'apprentissage et la progression de l'ensemble des élèves (Duflo et al., 2011 ; Banerjee et al., 2016). Néanmoins, les résultats allant dans ce sens sont observés au sein de pays en développement, dans des systèmes éducatifs marqués par une très grande hétérogénéité des élèves, en particulier en termes d'âge des élèves, une même classe comprenant souvent des élèves ayant plusieurs années d'écart, et des tailles de classe beaucoup plus importantes que dans le contexte des pays européens. Aussi, les réformes étudiées sont mises en œuvre dans un contexte expérimental très spécifique, comportant notamment une allocation aléatoire des enseignants aux différents groupes et un système d'incitation financière original à destination des professeurs, contexte qui diverge donc nettement de la réforme étudiée dans ce présent rapport.

Des effets observés globalement nuls sur les résultats scolaires dans le second degré

Dans les contextes américains et britanniques, où les regroupements d'élèves selon leurs résultats scolaires (*ability grouping*, voir **Encadré 1**) sont mis en place sous différentes formes pour les élèves du premier et du second degré, une ancienne revue de la littérature basée sur des études corrélationnelles ou/et sur petits échantillons indique que les différentes modalités de regroupement ou de non-regroupement n'ont pas d'effet sur les résultats de la majorité des élèves (Sukhnandan et Lee, 1998). En d'autres termes, du point de vue des résultats scolaires de l'élève, cela n'aurait pas d'incidence d'être dans une classe regroupant des élèves de niveaux scolaires différents plutôt que dans une classe regroupant des élèves du même niveau, que les groupes soient constitués pour chaque discipline indépendamment des autres, ou qu'ils soient définis une seule fois pour toutes les matières.

ENCADRE 1 : Qu'entend-on par groupes de niveau ?

Dans la littérature scientifique internationale en éducation, le terme générique pour parler des groupes de niveau est « *ability grouping* », qui renvoie ainsi plus aux compétences des élèves qu'au seul niveau scolaire observé en début d'année scolaire.

Derrière ce terme se cache une grande diversité des pratiques qui peuvent être regroupées en deux familles : les regroupements par classe de niveau (*between class ability grouping*), qui consistent à affecter les élèves à des classes distinctes selon leurs compétences, pour une ou plusieurs disciplines, et les groupes de niveau au sein des classes (*within class ability grouping*), qui désignent des organisations plus flexibles et temporaires où l'enseignant divise ponctuellement sa classe en groupes pour une activité ou une séquence particulière.

D'après les résultats de l'enquête PISA 2022, les regroupements d'élèves selon leurs compétences sont une pratique relativement courante dans les pays de l'OCDE. Bien que ces pratiques soient en recul depuis 2018, le regroupement dans des classes de niveau pour une ou plusieurs disciplines (*between class*) concerne 40 % des élèves de l'OCDE en 2022, et est particulièrement répandu en Irlande, en Israël, à Malte ou au Royaume-Uni (OCDE, 2023a). Le regroupement d'élèves au sein de leur classe (*within class*) concerne la moitié des élèves des pays de l'OCDE, et s'observe notamment aux Pays-Bas ou en Hongrie. Au sein des pays de l'OCDE, les regroupements d'élèves en classes de niveau ne semblent pas dépendre des caractéristiques de l'établissement scolaire, en revanche les regroupements d'élèves au sein de leur classe sont plus fréquents dans les établissements les plus défavorisés. Les résultats descriptifs de l'enquête PISA montrent que les résultats en mathématiques sont en moyenne très légèrement supérieurs dans les pays qui ne mettent pas en place de regroupements d'élèves, mais il est difficile d'attribuer ces écarts avec certitude à cette modalité d'organisation, les différences de politiques éducatives entre pays étant extrêmement nombreuses.

Le **Tableau 1** ci-après présente différentes modalités de regroupement ainsi qu'un bref exposé des effets identifiés dans la littérature.

TABEAU 1 • Synthèse de la littérature sur les effets des regroupements d'élèves

Type de regroupement		Définition	Effet connu
<i>Within class ability grouping</i>		Les élèves sont affectés temporairement à des groupes au sein de leur classe de référence.	Effet généralement positif mais modeste, à condition que les groupes soient réellement flexibles et qu'il y ait une adaptation des pratiques pédagogiques. [Niveau de preuve très limité basé sur des études corrélationnelles et/ou à échantillons de petite taille]
<i>Between class</i>	<i>Setting</i>	Regroupement en classe de niveau pour une ou plusieurs disciplines au cours d'une année scolaire	Faibles effets positifs si les groupes sont flexibles et réorganisés en cours d'année, effet nul sinon. [Niveau de preuve très limité basé sur des études corrélationnelles et/ou à échantillons de petite taille]
	<i>Tracking / streaming</i>	Regroupement durable : les élèves sont affectés à des classes de niveau qu'ils conservent au cours de leur scolarité	Effet moyen nul, effet potentiellement positif pour les élèves les plus performants, effet stigmatisant pour les élèves les plus fragiles. [Niveau de preuve très limité basé sur des études corrélationnelles et/ou à échantillon de petite taille]
<i>Gifted programs</i>		Classes composées des élèves les plus performants	Résultats incertains, parfois positifs, parfois négatifs. De plus, la littérature n'aboutit pas à un consensus sur les effets indirects sur les autres élèves. [Niveau de preuve limité basé sur un petit nombre d'études causales sur grands échantillons]
<i>Remediation programs</i>		Classes au sein desquelles se retrouvent les élèves les plus en difficulté	Effets positifs si suffisamment de moyens y sont consacrés et si les pratiques pédagogiques sont adaptées aux besoins des élèves. [Niveau de preuve limité basé sur un petit nombre d'études causales sur grands échantillons]
<i>Cross grade grouping</i>		Les élèves de plusieurs niveaux et donc d'âges différents sont regroupés selon les compétences qu'ils maîtrisent ou qu'ils doivent acquérir.	Résultats plutôt positifs, notamment en élémentaire. [Niveau de preuve très limité basé sur des études corrélationnelles et/ou à échantillons de petite taille]

Slavin (1990), en se concentrant sur des travaux de recherche portant sur des élèves américains et britanniques du second degré, explique cette absence d'effet moyen quelle que soit la nature des regroupements par la faible variation des pratiques pédagogiques dans les différents contextes. Selon l'auteur, le regroupement en tant que tel ne peut influencer les apprentissages que si les enseignants tirent profit de cette organisation pour mettre en œuvre des pratiques innovantes, par exemple en faisant intervenir davantage les élèves et en variant les activités pédagogiques. Face à l'absence d'impact observé, Slavin (1990) appelle à l'abandon des regroupements d'élèves, jugés comme une pratique anti-égalitaire, tant que la recherche n'aboutirait pas à une conclusion univoque sur l'effet positif de ces organisations pédagogiques.

Les revues de littérature les plus récentes tendent à confirmer l'absence d'effet moyen sur les performances des élèves (Hattie, 2023). John Jerrim et ses coauteurs, à partir d'analyses corrélationnelles (non causales) sur les contextes britanniques et internationaux, ne trouvent pas

d'association ni entre les différents types de regroupement et les performances scolaires des élèves du primaire et du secondaire, ni entre les différents types de regroupement et la confiance en soi des élèves et le bien-être scolaire (Jerrim, 2021 ; Jerrim et al., 2025). Les auteurs en concluent que c'est aux enseignants de choisir le type d'organisation pédagogique qui leur convient le mieux, en soulignant comme condition nécessaire à l'efficacité du regroupement une affectation des élèves « *transparente, non biaisée, et juste* », et une possibilité réelle offerte aux élèves de changer de groupe en cours d'année. Aussi, les auteurs insistent sur l'importance de ne pas systématiquement assigner les enseignants perçus comme les plus compétents aux groupes d'élèves les plus performants ou bien aux élèves les plus en difficulté.

À ce jour, la dernière méta-analyse sur le sujet des regroupements d'élèves au collège est réalisée par Terrin et Triventi (2023) et fait partie des plus convaincantes car elle prend en compte la qualité des études existantes en termes de mesure d'un impact causal. Elle porte sur l'ensemble des études publiées en langue anglaise entre 2000 et 2021, qui utilisent des méthodes expérimentales ou quasi-expérimentales robustes afin d'isoler l'impact causal des regroupements sur les apprentissages. Leur conclusion est que les regroupements d'élèves au collège n'ont aucun impact sur la réussite des élèves en moyenne, mais contribuent très fortement à accroître le niveau d'inégalités dans le système scolaire. Ils montrent par ailleurs qu'il ne semble pas exister d'effets différenciés des regroupements selon la discipline.

Des effets possiblement différenciés selon le niveau initial des élèves et selon l'allocation des ressources

Si Terrin et Triventi (2023) concluent dans leur méta-analyse que les regroupements d'élèves au collège n'ont aucun impact sur la réussite des élèves en moyenne, ils montrent toutefois qu'ils contribuent très fortement à accroître le niveau d'inégalités dans le système scolaire. Les politiques de de-tracking, qui consistent à abandonner les regroupements d'élèves, pourraient ainsi conduire à réduire les inégalités sans pénaliser le niveau moyen. Les auteurs soulignent ainsi que d'après certaines études, les politiques d'abandon des groupes de niveau peuvent affecter négativement les performances scolaires des élèves les plus performants. Afin de limiter ces effets négatifs, les deux auteurs proposent d'accompagner ces politiques de de-tracking de dispositifs permettant la stimulation de tous les élèves, comme par exemple la collaboration et l'entraide entre élèves de niveaux hétérogènes.

Dans le contexte américain, une étude ancienne et corrélationnelle de Hoffer (1992) observe une corrélation légèrement positive entre regroupements par niveau et progression individuelle pour les élèves ayant les meilleurs résultats scolaires, mais une association négative pour les élèves ayant initialement les résultats les moins bons. Cette observation est cohérente avec la revue de littérature de Dupriez et Draelants (2004), soulignant que l'absence d'effet moyen des regroupements d'élèves pourrait s'expliquer par une compensation d'effets positifs pour les élèves les plus performants par les effets négatifs pour les élèves les plus fragiles. Ces études défendent le principe de classes hétérogènes qui ne nuiraient pas au niveau moyen des apprentissages mais limiteraient les effets inégalitaires et stigmatisants des regroupements d'élèves. Cependant, **ces études corrélationnelles ne permettent pas de conclure avec une grande certitude sur les effets causaux, il faut donc les considérer avec une certaine prudence.**

Ainsi, les regroupements d'élèves selon leurs compétences conduiraient en théorie à une certaine divergence dans les trajectoires scolaires des élèves, marquée par une progression des élèves les plus performants initialement et des conséquences négatives sur les élèves les plus en difficulté. Kerckhoff (1986), dans son étude des groupes de niveau dans les collèges britanniques, propose comme piste d'explication la différenciation des programmes qui conduit les élèves les plus performants à maîtriser des éléments toujours plus avancés tandis que les élèves les plus en difficulté se limitent à l'apprentissage de notions de base. Hermann et al. (2024), dans leur analyse relative à des collégiens

en Hongrie, soulignent également que l'allocation des ressources au sein des établissements se fait en faveur des regroupements composés des élèves les plus performants : ceux-ci bénéficieraient ainsi des enseignants les plus expérimentés et du meilleur matériel. Certaines recherches remettent cependant en cause cette explication en soulignant qu'il est possible que les élèves les plus en difficulté apprennent mieux en face de professeurs peu expérimentés, et en faisant le constat que, dans certains contextes scolaires, les établissements ont déjà tendance à allouer inégalement les ressources entre les élèves, même en l'absence de groupes formalisés (Betts et Shkolnik, 2000a). Plus récemment, Avery et Bacher-Hicks (2026) se sont intéressés à la constitution des classes de la fin de l'école primaire à la fin du collège dans un État du sud-est des États-Unis ; ils ont mis en évidence le fait que, dans le contexte étudié, les chefs d'établissement tendent à attribuer les professeurs les plus performants aux élèves les plus performants.

En étudiant l'effet de classes à destination des élèves les plus en difficulté dans les collèges de Floride, Figlio et Özek (2024) ont montré dans une étude causale sur large échantillon que même lorsque ce sont ces élèves qui bénéficient davantage des ressources de l'établissement, l'effet positif de ces classes se retrouve diminué par le choix des élèves de ce groupe de poursuivre ensuite des enseignements moins avancés et des études moins sélectives. D'après les auteurs, ces élèves intériorisent leur statut d'élèves moins performants, ce qui les inhibe ensuite dans leurs choix d'orientation scolaire. Dans une autre étude causale récente sur le Texas, Antonovics et al. (2026) montrent que les regroupements selon le niveau scolaire induisent une augmentation des inégalités en améliorant légèrement les résultats scolaires des élèves les plus performants sans affecter négativement les résultats des élèves les plus en difficulté, tout en insistant sur l'importance de l'ensemble des pratiques mises en œuvre empiriquement dans les établissements allant de pair avec ces regroupements.

En revanche, les travaux de recherche aboutissent à un consensus relatif autour des bénéfices probables des regroupements pour les élèves à besoins éducatifs particuliers, à condition toujours que des moyens suffisants leur soient alloués, en termes notamment de taille de groupes et de formation des enseignants, et que les organisations pédagogiques permettent la mobilité des élèves en cours d'année (Castle et al., 2005 ; Aisenberg & Lobut, 2023).

Le rôle des effets de pairs pour les élèves les plus en difficulté ou les plus doués scolairement

Les regroupements d'élèves peuvent influencer indirectement les performances scolaires en modifiant la structure des effets de pairs au sein des groupes. En éducation, la notion d'effets de pairs (*peer effects* en anglais) renvoie à l'impact que les élèves peuvent avoir les uns sur les autres par leurs interactions ou leur simple présence. Ces effets peuvent être de nature très variée, comme par exemple en termes de résultats scolaires mais aussi de comportement, de confiance en soi, d'aspirations ou de motivation. Dès 1997, Duru-Bellat et Mingat soulignent l'importance du « contexte de scolarisation » dans l'analyse des effets des classes de niveau dans les collèges français. Comme il n'y avait pas de « classes de niveau » à proprement parler dans les collèges français au moment de leur travail de recherche, leur définition repose sur l'homogénéité observée au sein des classes. Les deux auteurs soulignent que l'existence de classes de niveau « de fait » est assez répandue, en partie en raison des choix d'options et de langues vivantes. Ce travail qualitatif montre des associations positives entre le niveau général de la classe et la progression individuelle. Ce résultat rejoint ceux de travaux causaux qui s'intéressent à cette question des effets de pairs et concluent également que l'apprentissage des élèves d'une classe ou d'un groupe est d'autant plus favorisé que la proportion d'élèves performants est importante, mais que la taille de ces effets reste très limitée, et sans pour autant que cette relation soit nécessairement linéaire (Brodsky, 2010). D'après ces travaux, isoler les élèves les plus en difficulté d'un côté et les élèves les plus performants de l'autre contribuerait à exacerber les écarts de résultats scolaires, bien que la magnitude de ces associations reste très modérée.

L'amélioration du climat scolaire est une des justifications fréquemment évoquées pour le regroupement des élèves les plus en difficulté. Lavy et al. (2012a), en étudiant les collégiens et lycéens en Israël, montrent par exemple que la proportion d'élèves en difficulté au sein d'un groupe, en particulier les redoublants, influence négativement les performances de l'ensemble du groupe en raison des problèmes de comportement plus fréquents et du temps supplémentaire dédié à la gestion de classe plutôt qu'à l'enseignement.

Pour les regroupements d'élèves les plus performants, les résultats sont incertains et varient selon les études. Kulik et Kulik (1982), dans leur méta-analyse dédiée au second degré, expliquent leurs meilleurs résultats par la plus grande émulation qui s'observerait dans les classes d'élite, favorisant ainsi le progrès des élèves qui s'y trouvent. Ils estiment pour leur part que le regroupement d'élèves les plus en difficulté a un impact positif sur leur scolarité : cela leur permettrait en effet de gagner en motivation, d'améliorer leur perception de l'École ou d'accroître leur intérêt pour les matières enseignées. Cependant, ce travail ancien repose sur des corrélations et ne permet pas d'isoler les effets de pairs de l'effet des caractéristiques initiales propres aux élèves eux-mêmes. Steenbergen-Hu et al. (2016), dans leur revue de cent ans de littérature sur la question des regroupements d'élèves de la maternelle au lycée aux États-Unis, suggèrent qu'il y a des effets positifs et probants pour les classes regroupant les élèves les plus performants, comme pour les classes regroupant des élèves de plusieurs niveaux ou pour les sauts de classe. Là encore, ces études sont corrélationnelles et n'isolent pas les effets de pairs des effets des caractéristiques individuelles initiales des élèves eux-mêmes. Dans un travail publié récemment, Landaud et Maurin (2026) étudient l'effet causal d'être scolarisé dans les meilleures classes préparatoires aux grandes écoles, dites classes « étoiles », plutôt que dans une classe moins prestigieuse dans un même établissement. En comparant des élèves avec des niveaux initiaux similaires, les auteurs observent un effet positif très fort de l'appartenance à ces classes étoiles sur les rangs dans les différents concours et l'admission dans les écoles les plus sélectives, possiblement parce qu'elles permettent aux élèves les plus performants de développer tout leur potentiel, de gagner en confiance et de bénéficier d'un environnement scolaire plus stimulant. Leurs résultats confirment ceux de Wu et al. (2019) portant sur la Chine, qui montrent que les élèves admis dans les meilleures classes des lycées chinois les plus prestigieux obtiennent des meilleurs scores à l'examen d'entrée à l'université que ceux qui sont dans des classes ordinaires. Cependant, Landaud et Maurin (2026) soulignent que ces classes étoiles, qui regroupent finalement les élèves issus des milieux les plus favorisés, contribuent significativement à l'homogénéité sociale observée dans les grandes écoles et à la reproduction des inégalités sociales et scolaires².

D'autres études dans le secondaire ne retrouvent pas ces effets positifs de regrouper les élèves les plus performants et montrent qu'appartenir à une classe très performante n'est pas toujours un avantage. Une étude causale fondée sur l'ensemble des lycées français entre 2010 et 2016 montre un effet négatif d'être dans une classe très performante pour les 25 % des élèves initialement les plus performants, effet qui peut s'expliquer par le fait d'une concurrence accrue ainsi que la perte du bénéfice psychologique et motivationnel d'être situé en tête de classe (Boutchenik et Maillard, 2019). Ainsi, une concentration de très bons élèves entre eux peut pénaliser les bons et très bons élèves en les exposant à une compétition trop importante, source de stress ou de détérioration de l'estime de soi. Cet effet de « rang » est documenté dès 1984 par Marsh et Parker sous l'appellation de *Big Fish Little Pond Effect* (« effet grand poisson, petit étang ») et démontré empiriquement dans plusieurs travaux récents utilisant des méthodes empiriques robustes (Elsner and Isphording, 2017 ; Murphy et Weinghart, 2020 ; Pagani et al., 2021 ; Delaney and Devereux 2022 ; Denning et al. 2023).

Les travaux de recherche ne s'accordent pas sur l'effet, pour les élèves les plus en difficulté, de côtoyer des élèves plus performants en classe. Hoxby et Weingarth (2005), en étudiant les effets sur les scores de fin de collège d'une expérience quasi-naturelle de réallocation d'élèves selon leur statut

² Bonneau et Dousset (2025) ont également mis en évidence un effet différencié selon le genre d'un environnement de préparation compétitif sur les performances académiques.

socio-économique en Caroline du Nord, montrent que les élèves les plus performants ne sont pas pénalisés lorsqu'ils sont en présence d'élèves moins performants et que les élèves plus fragiles progressent davantage lorsqu'ils interagissent avec des élèves plus performants. En revanche, l'étude de deux réformes de suppression des groupes de niveau en Autriche et en Allemagne par Fleischmann et al. (2023) suggère que côtoyer des élèves avec un niveau scolaire plus élevé a un impact négatif sur la progression des élèves les plus en difficulté, en particulier en raison de l'effet de comparaison négatif qui pèse sur leur motivation et sur leur confiance en leur capacité à réussir. Les groupes de niveau permettraient ainsi d'éviter aux élèves les plus en difficulté de subir une concurrence pénalisante. Il convient toutefois de souligner que cette étude repose sur une approche corrélationnelle et non causale. Booij et al. (2017), dans une expérimentation aléatoire menée à l'université d'Amsterdam, mettent également en avant les bénéfices des groupes de niveau sur les étudiants les plus en difficulté, à travers leur implication plus grande en classe et la multiplication des interactions bienveillantes entre les étudiants, notamment dans le travail en groupe et l'aide reçue ou proposée.

De nombreuses dimensions restent à explorer pour comprendre le rôle des effets de pairs sur la réussite scolaire, en particulier le genre des élèves. Lavy et al. (2012b), dans leur étude des effets de pairs dans les collèges britanniques, expliquent qu'une partie de cette ambivalence sur l'impact des élèves les plus performants pourrait s'expliquer par le fait que les filles en tirent un avantage tandis que les garçons s'en trouvent pénalisés. Cette étude confirme ainsi les conclusions de Boaler (1987) obtenues dans ce même contexte 25 ans plus tôt en utilisant des méthodes qualitatives. Cette dimension est peu explorée dans ce document de travail et pourra faire l'objet de travaux futurs, en tirant notamment profit des données au niveau des élèves et de leur groupe d'appartenance plutôt qu'à l'échelle de l'établissement ou du niveau et de la discipline.

Si les travaux de recherche autour des effets de pair dans l'éducation ne convergent pas vers un résultat commun, ils mettent en évidence la faible magnitude des éventuels effets de pairs sur la réussite éducative, ne permettant pas de résoudre la question de la ségrégation sociale et scolaire, ni d'offrir les mêmes opportunités de progression pour tous les élèves (Hoxby et Weingarth, 2005 ; Monso et al., 2019).

Des politiques de regroupement difficiles à qualifier

Une autre difficulté qui complique la mesure d'un effet unique des regroupements d'élèves sur leurs performances provient de la grande diversité des organisations pédagogiques désignées par cette notion, qui complique la compréhension de l'objet étudié et l'identification de ses effets sur les apprentissages des élèves (voir **Encadré 1**). À ce sujet, Rees et al. (2000), dans leur critique de l'article de Betts et Shkolnik (2000a), considèrent qu'il n'est pas possible d'évaluer une telle politique à partir de données déclaratives, en particulier parce que les réalités de terrain peuvent fortement diverger pour des groupes pourtant labellisés de la même manière d'un contexte à l'autre. Selon eux, il faudrait donc plutôt s'appuyer sur les données administratives et recourir à des mesures objectivées et continues de l'homogénéité au sein des groupes.

Même avec des données plus fines au niveau des élèves, il est extrêmement difficile de tenir compte de toutes les spécificités des groupes mis en place, ni des logiques d'assignation qui dépendent des préférences des enseignants et du niveau de collaboration au sein des équipes pédagogiques pour chacune des matières concernées. Lecomte et Maleyrot (2026), dans une étude ethnographique récente sur la mise en place des groupes de besoins dans un collège rural français, montrent ainsi que derrière la qualification de « groupes hétérogènes », constitués selon l'ordre alphabétique des noms des élèves, il y a deux réalités très différentes selon la discipline. En mathématiques, en raison de l'investissement personnel des enseignants, le fonctionnement effectif correspond ainsi davantage à des groupes de besoins flexibles, avec des retours en groupe hétérogène de référence et des réallocations d'élèves en fonction de leurs besoins. En revanche, en français, les élèves restent dans les mêmes groupes hétérogènes tout au long de l'année, à l'exception des élèves les plus fragiles

ou à besoins particuliers qui sont mis dans un groupe à part pour bénéficier de l'enseignement d'un professeur des écoles spécialisé.

Des résultats encourageants pour les formes les plus flexibles en cas de moyens conséquents investis

Lou et al. (2000) tentent de réconcilier les résultats d'une cinquantaine d'études sur les regroupements d'élèves de l'école primaire à l'enseignement supérieur, présentant toutefois un faible niveau de preuves. Les auteurs arrivent à la conclusion que les regroupements d'élèves selon leurs compétences sont plus efficaces que les classes hétérogènes pour faire progresser les élèves. Cependant, ce résultat n'est valide qu'à plusieurs conditions : les groupes constitués doivent être de taille réduite, les professeurs doivent être suffisamment formés et équipés pour pouvoir adapter leur enseignement aux différents groupes et la répartition des élèves entre les groupes ne doit pas se faire uniquement par rapport aux notes, et doit tenir compte d'autres caractéristiques comme le genre pour assurer la cohésion des groupes. Il faut également noter que cette étude n'emploie pas une méthode qui permet d'estimer véritablement l'effet causal des regroupements et de leurs différentes modalités.

Cette recommandation en faveur de groupes de niveau *sous conditions* est celle qui semble finalement ressortir le plus de la littérature. Steenbergen-Hu et al. (2016) soulignent ainsi que les effets positifs identifiés dans la littérature sont valides pour des regroupements flexibles et temporaires et lorsque les élèves restent tous au sein d'une même salle de classe avec un professeur passant d'un groupe à l'autre. La note IDEE de synthèse de la littérature sur la différenciation des apprentissages met également en avant la flexibilité des groupes et le développement professionnel continu des professeurs comme des conditions nécessaires pour que les groupes de niveau puissent profiter à tous les élèves (Aisenberg & Lobut, 2023). Les modalités de regroupement qui fonctionneraient le mieux seraient ainsi celles qui consistent à regrouper les élèves par groupes au sein des salles de classe, de manière temporaire au cours d'une heure d'enseignement. Ces formats sont plus adaptés aux classes de primaire, ce qui pourrait expliquer que les effets observés sont généralement plus forts et plus significatifs dans le premier degré (EEF, 2021).

À partir d'une étude qualitative dans les collèges britanniques, Ireson et al. (2002) montrent toutefois que la flexibilité des groupes est parfois abandonnée en raison des complexités organisationnelles que cela représente et que l'affectation initiale des élèves dans les groupes peut être imparfaite. Cela peut conduire des élèves à être durablement affectés à un groupe qui ne leur correspond pas et nuire ainsi à leur progression tout au long de leur cursus.

La question de la mixité sociale et scolaire à l'école

Au-delà des effets individuels sur la performance scolaire qui s'avèrent incertains, la mise en place des regroupements d'élèves peut avoir un impact sur le niveau de mixité et de ségrégation sociale au niveau des classes et des établissements. Les effets de la mixité sont documentés au-delà de la simple performance scolaire : une revue de littérature récente souligne ainsi des effets positifs de la mixité en milieu scolaire sur l'accès et la réussite dans l'enseignement supérieur, sur l'intégration sur le marché du travail et sur l'atténuation des comportements déviants pour les élèves issus de milieux sociaux plus défavorisés (Charousset et al., 2023). Pour ce qui concerne la France, une étude récente sur des expérimentations de mixité sociale au collège montre également des effets positifs sur le bien-être d'être scolarisé dans un environnement mixte plutôt qu'un environnement concentrant un nombre important d'élèves de sa propre classe sociale, et ce aussi bien pour les élèves favorisés que les élèves défavorisés (Azmat et al., 2026).

Même en l'absence de regroupements d'élèves, le système éducatif français se caractérise historiquement par une certaine forme de ségrégation sociale entre les classes d'un même établissement (Duru-Bellat et Mingat, 1997 ; Ly et al., 2014 ; Boutchenik et Maillard, 2019). Ainsi, au regard de la corrélation importante en France entre réussite scolaire et origine sociale des élèves (OCDE, 2023b), des regroupements d'élèves par niveau scolaire pourraient augmenter la ségrégation sociale observée au sein des établissements, et ainsi augmenter encore davantage les inégalités de parcours scolaire selon l'origine sociale.

Par ailleurs, une littérature causale récente portant sur les Etats-Unis trouve des effets assez forts sur les convictions politiques des enfants à long terme (Billings et al., 2021 ; Kaplan et al., 2025), des effets qui dépassent donc largement le cadre scolaire et soulignent l'importance des potentiels effets sociétaux d'une politique renforçant la ségrégation sociale entre classes.

Néanmoins, le regroupement d'élèves selon le niveau scolaire pourrait aussi conduire à limiter la ségrégation inter-établissement, en incitant les familles des élèves les plus performants à scolariser leurs enfants dans le collège de secteur si celui-ci met en place de tels regroupements. Dans le contexte américain, Figlio et Page (2002) émettent l'hypothèse que cette rétention des collégiens les plus performants au niveau d'un établissement pourrait expliquer leur résultat selon lequel les élèves les plus fragiles ne sont pas pénalisés lorsque leur collège met en place des groupes de niveau.

Les effets liés à une réduction de la taille du groupe

Les regroupements d'élèves peuvent s'accompagner d'une réduction du nombre d'élèves en face d'un enseignant par rapport aux cours en classe de référence. En France, l'année 2024, date de la mise en place des nouvelles organisations pédagogiques, a été marquée par une baisse significative du nombre d'élèves devant un enseignant pour les enseignements de français et de mathématiques en 6^e et en 5^e, qui ne se retrouve pas pour les autres niveaux ou disciplines (Baradji, 2025)³. Ces réductions d'effectif peuvent être concentrées sur les élèves les plus en difficulté, ou au contraire sur les élèves les plus performants lorsqu'ils sont mis à part : les éventuels effets des regroupements d'élèves par niveau ou compétence s'ajoutent donc dans certains cas à un possible effet de réduction de la taille des classes. Par ailleurs, ces réductions peuvent être contrebalancées par une augmentation de la taille des structures pour les autres niveaux ou disciplines : il s'agit alors d'une réallocation de ressources au sein de l'établissement, qui peut neutraliser l'effet escompté de la réforme (Sims, 2009).

³ La mise en place des groupes entraîne, au niveau national, une baisse de l'indicateur E/S comprise entre 1,5 et 2,4 élèves pour les enseignements de français et de mathématiques en 6^e et en 5^e. L'indicateur E/S correspond au nombre d'élèves dont un professeur a la charge en moyenne pendant une heure de cours.

Bien que la direction et la magnitude de ce dernier effet ne fasse pas consensus dans la littérature, les études expérimentales et quasi-expérimentales les plus robustes semblent indiquer un effet fort et significatif d'une réduction de la taille des classes dans le premier degré, mais les conclusions sont moins claires au niveau du collège (Bouguen et al., 2017 ; Fajeau et al., 2025). En France, pour le premier degré, la DEPP a conduit une évaluation des effets de la réduction de taille des classes de CP et CE1 en éducation prioritaire à partir de la rentrée 2017 (Andreu et al., 2021). Cette étude met en avant un effet positif significatif relativement important sur les résultats scolaires pour les élèves de CP, qui persiste en CE1⁴. Un document de travail récent souligne néanmoins que ces effets positifs s'estompent à l'entrée au collège (El Bah, 2026). Pour le second degré, Piketty et Valdenaire (2006), dans le contexte français, identifient un effet positif et significatif de la réduction de la taille des classes en 3^e sur la réussite au brevet, bien que cet effet soit moindre que ceux identifiés au primaire. De même, Fredriksson et al. (2013) montrent que la réduction de la taille des classes a des effets importants sur les résultats scolaires comme sur les dimensions socio-émotionnelles des élèves au début de leur scolarité au collège, et que ces effets bénéfiques persistent à moyen et long terme. En revanche, en appliquant la même méthode que les deux études précédentes, Leuven et al. (2008) ne trouvent pas d'effet significatif pour des élèves de milieu de collège en Norvège et Angrist et al. (2019) arrivent à la même conclusion pour des élèves de fin d'école primaire en Israël.

À nouveau, le lien positif entre réduction de la taille des classes et progression des élèves pourrait dépendre de l'allocation des ressources entre les établissements scolaires. Jepsen et Rivkin (2009) montrent ainsi que la très forte réduction de la taille des classes dans les écoles primaires californiennes a un effet presque nul sur les résultats scolaires des élèves, probablement car elle s'accompagne d'un recrutement massif de professeurs inexpérimenté, en particulier au sein des écoles les plus défavorisées.

Une politique efficiente ?

Dans le premier degré, les regroupements d'élèves selon leur niveau scolaire semblent efficaces même sans réduction de la taille des classes, notamment lorsqu'ils sont mis en œuvre ponctuellement au sein d'une même salle de classe. De ce point de vue, il s'agirait donc d'une politique efficiente nécessitant peu de moyens supplémentaires. Ce sont les arguments mis en avant par Collins et Gan (2013) et Card et Giuliano (2016) pour justifier la mise en place de groupes de niveaux dans les écoles primaires américaines.

En revanche, s'agissant du second degré, la littérature présentée ci-dessus est moins concluante en ce qui concerne les effets propres des regroupements d'élèves. Ainsi, la mise en place de groupes flexibles de plus petite taille avec des réorganisations en cours d'année et d'éventuelles modifications de pratiques pédagogiques peut impliquer des coûts non négligeables (heures d'enseignement, formation, coordination des équipes pédagogiques, etc.). Cependant, la mise en œuvre de regroupements fixes semble conduire à une absence d'effets sur les performances scolaires, voire des effets négatifs pour les élèves les plus en difficulté. En l'absence d'effets forts d'une telle politique, la question se pose de savoir comment le coût et les bénéfices d'une telle politique se comparent à ceux d'autres politiques éducatives dont l'efficacité peut déjà être mesurée dans la littérature (Grenet et Landais, 2025).

Conclusion

Si la littérature suggère plutôt des effets nuls des regroupements d'élèves selon leur niveau sur leurs performances scolaires au collège, elle se caractérise par une disparité très importante des résultats de la recherche, aussi bien en ce qui concerne les effets de ces organisations pédagogiques sur les

⁴ En s'appuyant sur une approche descriptive et corrélationnelle, Murat et Rocher (2026) ont montré que les élèves de l'éducation prioritaire, concernés par la diminution de la taille des classes en CP et en CE1, progressent plus rapidement que les autres au cours de ces deux niveaux, notamment dans les domaines moins socialement marqués où l'action de l'école est importante. Ce n'est plus le cas en CM1, lorsqu'il n'y a plus de classes dédoublées.

élèves les plus performants que sur les élèves les plus en difficulté. D'après les différentes études recensées, les formes d'organisation les plus flexibles pourraient être plus efficaces car elles seraient plus adaptées à la diversification des activités pédagogiques et permettraient de garantir une unité de la classe et de limiter ainsi la stigmatisation des élèves selon leur groupe d'appartenance.

Par ailleurs, la diversité des modalités de mise en œuvre des regroupements pédagogiques rend difficile l'évaluation de leurs éventuels bénéfices et invite à la prudence dans la transposition des résultats de la recherche à d'autres contextes scolaires. Plusieurs études soulignent notamment que les écoles et établissements peuvent mobiliser différentes formes de différenciation des ressources, des pratiques ou des enseignements même en l'absence de politiques formelles de regroupements d'élèves. En particulier, de rares études sur données françaises (Ly et Riegert, 2016 ; Souidi, 2023) rapportent une ségrégation scolaire intra-établissement qui existait déjà en France avant la mise en place des nouvelles organisations pédagogiques à la rentrée 2024. Il sera donc important de compléter ces travaux avec une mesure plus fine de l'hétérogénéité scolaire des élèves qui se côtoient dans une même salle de classe afin de quantifier l'effet réel de la réforme sur la mixité sociale et scolaire au collège.

➤ Cadre institutionnel

La différenciation pédagogique et la gestion de l'hétérogénéité en France

Relativement aux autres pays de l'OCDE, le système scolaire français se caractérise par un niveau de différenciation pédagogique relativement faible (Mons et al., 2017)⁵. En France, les enseignants de collège sont moins nombreux à différencier couramment les travaux donnés aux élèves avec le plus de difficultés et aux élèves les plus performants. De même, il y a relativement moins d'élèves en France qui estiment que leurs enseignants proposent une aide individualisée aux élèves les plus en difficulté. D'après les résultats de l'enquête Talis 2024, au collège, 85 % des enseignants affirment adapter fréquemment les méthodes d'enseignement aux besoins des élèves, 74 % indiquent poser fréquemment des questions de niveaux de difficulté différents pour vérifier la compréhension par les élèves de l'objet travaillé et 45 % déclarent orienter fréquemment les élèves vers différents supports d'apprentissages selon leurs besoins. Au niveau de l'Union européenne, ces proportions sont respectivement de 92 %, 86 % et 67 % (Ceesay et al., 2025). Par ailleurs, en France, 70 % des enseignants estiment atteindre l'objectif d'adapter leur enseignement pour répondre aux différents besoins des élèves au collège, contre 81 % en moyenne pour les autres pays européens. Ce moindre recours à la différenciation pourrait s'expliquer par un manque de formation à ces enjeux et pratiques ou encore par une conception de la justice scolaire historiquement fondée sur l'égalité de traitement, qui entre parfois en tension avec l'adaptation de l'enseignement aux besoins des élèves (Dubet et Duru-Bellat, 2004).

Les enquêtes nationales et internationales sur les pratiques enseignantes menées par la DEPP mettent en avant les difficultés que rencontrent les enseignants de collège en France dans la gestion de l'hétérogénéité des élèves, et notamment la prise en charge des élèves en difficulté et des élèves à besoins éducatifs particuliers (Coppé et al., 2021 ; Goube et al., 2024). Le soutien aux élèves à besoins éducatifs particuliers apparaît d'ailleurs comme étant une demande très forte des enseignants, juste derrière l'augmentation des salaires et la réduction de la taille des classes (Charpentier et al., 2020). Bien que les enseignants considèrent les méthodes de prise en charge de cette hétérogénéité comme prioritaires, ils y renoncent souvent en raison des difficultés pratiques que cela représente (Potrel et al., 2024).

Le « Choc des savoirs »

La modification des organisations pédagogiques en 6^e et en 5^e pour les enseignements de français et de mathématiques a été annoncée dans le cadre de la réforme du « Choc des savoirs » présentée en décembre 2023 par le ministère chargé de l'éducation nationale. En s'appuyant sur un double constat d'une baisse du niveau des élèves français au collège dans les études internationales et d'une demande marquée pour une réduction des effectifs et pour une meilleure gestion de l'hétérogénéité des classes chez les enseignants, le ministère a annoncé la création de groupes de niveaux flexibles avec des effectifs réduits pour les élèves les plus en difficulté⁶.

Le terme privilégié pour décrire ces organisations pédagogiques est, par la suite, passé de « groupes de niveaux » à « groupes de besoins » afin d'accentuer le caractère flexible et transitoire de ces groupes. Ce glissement tend vers un rapprochement avec les résultats de la recherche qui semblent identifier des effets plus positifs de ce type de mesure lorsque les regroupements sont organisés de manière temporaire autour de compétences spécifiques (voir section **Revue de littérature**).

L'arrêté du 15 mars 2024 modifiant l'arrêté du 19 mai 2015 relatif à l'organisation des enseignements dans les classes de collège⁷, qui définit les nouvelles organisations pédagogiques retenues pour la

⁵ Ce rapport scientifique du Cnesco s'appuie sur les données de Talis 2013 et de PISA 2015.

⁶ Source : [Choc des savoirs. Élever le niveau de notre école, Dossier de presse du Mardi 5 décembre 2023](#).

⁷ Source : <https://www.legifrance.gouv.fr/jorf/id/JORFTEXT000049286467>.

rentrée de septembre 2024, confirme cette flexibilité dans les regroupements d'élèves prescrits. Il y est indiqué que les regroupements doivent être définis selon les besoins des élèves identifiés par les professeurs et peuvent être modifiés en cours d'année pour s'adapter à l'évolution des élèves et de leurs besoins. Si des retours en classe de référence sont prévus sur de courtes périodes pour une durée maximale de dix semaines, l'organisation retenue est celle des groupes pour l'ensemble des enseignements communs de français et de mathématiques au collège. L'objectif de réduction de la taille des groupes pour les élèves les plus en difficulté est confirmé dans cet arrêté.

Initialement prévues pour l'ensemble des classes et des niveaux du collège, les nouvelles organisations pédagogiques ont été mises en place pour les classes de 6^e et de 5^e à la rentrée de septembre 2024. Dans l'enseignement public, les moyens humains et financiers nécessaires à la mise en œuvre de ces nouvelles organisations ont été évalués à 2 330 équivalents temps plein (ETP) supplémentaires, correspondant à 1 509 000 heures de travail et un accroissement de la dotation horaire globale (DHG)⁸. Cette mise en œuvre a été accompagnée par une communication du ministère encourageant la souplesse et le pragmatisme pour ajuster au mieux les dispositions de l'arrêté du 15 mars 2024 au contexte et aux moyens de chaque établissement.

Au printemps 2025, l'Inspection générale de l'Éducation nationale, du sport et de la jeunesse (IGÉSR) a publié un premier rapport d'évaluation s'appuyant sur des observations qualitatives au sein de trente-neuf collèges (voir **Encadré 2**). Ce rapport indique dans ses conclusions que, malgré un contexte peu porteur (« *forte opposition syndicale* », « *évolutions des discours ministériels* »), les regroupements d'élèves ont été mis en place dans la quasi-totalité des établissements, selon des modalités différentes d'un établissement à l'autre.

ENCADRE 2 : Les principales conclusions du premier rapport de l'IGÉSR sur la mise en place des groupes de besoins lors de l'année scolaire 2024-2025

L'IGÉSR a publié un premier rapport sur la mise en œuvre des groupes de besoins à la rentrée 2024, en s'appuyant sur un travail d'observation auprès de 39 établissements. À travers des visites et des échanges avec les différents acteurs de la communauté éducative, l'IGÉSR a documenté une mise en pratique quasi-généralisée de la mesure mais sous des formes assez différentes, avec une majorité d'établissements adoptant des groupes de niveaux permanents, d'autres préférant conserver des groupes hétérogènes mais en profitant des moyens supplémentaires pour diminuer la taille des groupes et d'autres ne définissant que des groupes que pour une seule catégorie d'élèves, les élèves les plus en difficulté ou les plus compétents.

Les regroupements mis en place apparaissent finalement être beaucoup plus rigides que ce que ne le laissait entendre le décret du 15 mars 2024, et seule une minorité d'établissements ont véritablement suivi les consignes de l'arrêté du 15 mars 2024, avec des retours en classe de référence et des réallocations d'élèves en cours d'année. Les auteurs du rapport soulignent que ces formes plus flexibles demandent un effort de coordination beaucoup plus important pour l'ensemble de l'équipe pédagogique, et que l'ambition initiale de flexibilité a pu être abandonnée en raison des contraintes qu'elle représentait.

Par ailleurs, les regroupements ont conduit à réduire le nombre d'élèves face à un professeur, en particulier dans les groupes à plus forts besoins, mais les inspecteurs n'ont pas observé de modification des pratiques pédagogiques associée à cette réduction d'effectifs ou à la plus grande homogénéité des élèves au sein de ces groupes.

⁸ Compte rendu de la commission d'évaluation des politiques publiques relatives à la mission *Enseignement scolaire* de l'Assemblée nationale du 29 mai 2024.

Enfin, en termes d'effets, le rapport indique que la mesure a donné une plus grande visibilité à la difficulté scolaire, mais sans donner pour autant les moyens aux enseignants de la gérer. Il y a donc une crainte que le retour en classe de référence en 4^e soit très difficile pour les élèves des groupes les moins avancés, qui n'auront eu accès qu'à des enseignements moins exigeants que le reste de leur classe. Concernant les élèves les plus compétents, le rapport aboutit à des conclusions mitigées, entre d'une part un enseignement facilité pour les enseignants, qui s'en satisfont et peuvent aller plus loin dans le programme, et d'autre part un sentiment de compétition plus fort, voire même de déclassement pour les élèves les plus fragiles au sein de ces groupes globalement plus performants.

Référence : Claval, B., Darnault, C., Fliche, F., Quéré, M. & Sayac, N., « Mise en place des groupes de besoins en français et mathématiques au collège », Rapport de l'Inspection générale de l'éducation, du sport et de la recherche, N° 24-25 007, mai 2025.

<https://www.education.gouv.fr/sites/default/files/2025-06/mise-en-place-des-groupes-de-besoins-en-fran-ais-et-math-matiques-au-coll-ge--440628.pdf>.

Les regroupements d'élèves ont été reconduits à la rentrée scolaire de septembre 2025, mais leur caractère obligatoire a été levé en décembre 2025. Dans le cadre de la démarche « Collèges en progrès », le ministère invite les établissements scolaires à adopter des solutions locales, adaptées à leur contexte, plutôt qu'à suivre des directives globales. Les « groupes de besoins » seraient donc une option offerte aux établissements dans un plus large éventail de mesures, à considérer dans les situations où ces organisations représentent, selon les établissements, une solution efficace face aux difficultés scolaires des élèves. L'un des enjeux d'évaluation des nouvelles organisations pédagogiques mises en place en 2024-2025 et 2025-2026 n'est donc pas seulement de savoir si les groupes de besoins ont eu un effet sur la réussite des élèves, mais également d'identifier l'hétérogénéité éventuelle de ces effets selon les contextes scolaires, et de mettre en lumière les leviers qui peuvent conduire à des impacts positifs ou négatifs de ces dispositifs.

Mission d'évaluation

Dès le mois de mai 2024, le ministère chargé de l'éducation nationale a souligné la nécessité d'évaluer les effets de ces futures organisations pédagogiques et a confié cette mission conjointe à l'IGÉSR et à la DEPP. L'objectif de ce travail d'évaluation est de documenter les différentes formes de regroupements qui ont pu être mises en œuvre pour les années scolaires 2024-2025 et 2025-2026, de recueillir les perceptions des différents acteurs aux différents temps de la réforme, de mesurer l'évolution des pratiques enseignantes, en particulier autour de la question de la gestion de l'hétérogénéité scolaire en classe, de faire état de l'accompagnement proposé aux équipes, de documenter l'évolution des dynamiques intra-établissements suite à la mise en œuvre de ce dispositif et de mesurer de premiers effets sur les apprentissages.

L'IGÉSR a été chargée de réaliser une évaluation qualitative auprès de trente-neuf collèges (voir **Encadré 2**), tandis que la DEPP a été missionnée pour compléter ce volet qualitatif par une étude quantitative s'appuyant sur des questionnaires auto-administrés à destination d'un échantillon national représentatif d'établissements scolaires et sur les données administratives présentes dans son système d'information statistique, dont notamment les résultats aux évaluations nationales de rentrée. Pour ce faire, la DEPP a mis en place l'Étude relative à l'organisation des enseignements au collège (EEC), menée sur un échantillon national représentatif d'environ 1 500 collèges publics en France (hors Mayotte⁹), répartis dans 29 académies. Appuyée sur des questionnaires en ligne adressés aux chefs d'établissement et coordonnateurs des organisations pédagogiques¹⁰, ainsi que sur un

⁹ En raison des conséquences du cyclone Chido, les établissements de Mayotte n'ont pas été sollicités dans le cadre de cette enquête.

¹⁰ Pour cette enquête, le terme de coordonnateur renvoie aux enseignants de français ou de mathématiques qui interviennent pour les organisations pédagogiques mises en place en 6^e et/ou en 5^e pour l'année scolaire 2024-2025, ou désignés comme tels par leur chef d'établissement.

recueil de données de contexte extraites des logiciels de gestion scolaire des collèges, l'étude EEC vise à apprécier l'hétérogénéité des contextes scolaires concernés par la mesure des « groupes de besoins », à rendre compte des organisations pédagogiques retenues pour les années scolaires 2024-2025 et 2025-2026 en français et en mathématiques en classes de 6^e et 5^e et à documenter le ressenti des personnels s'agissant de la mise en œuvre et des potentiels premiers effets perçus des actions mises en œuvre. L'étude EEC cherche également à apprécier de façon causale les effets des organisations pédagogiques sur les apprentissages des élèves. Cette étude est menée en collaboration avec une équipe de recherche composée de (par ordre alphabétique) Carlo Barone (Sciences Po), Julien Grenet (CNRS, École d'économie de Paris), Marc Gurgand (CNRS, École d'économie de Paris), Nina Guyon (École normale supérieure-PSL, École d'économie de Paris), Elise Huillery (Université Paris Dauphine-PSL, École normale supérieure-PSL), Laïla Souali (Université Paris Dauphine-PSL) et Camille Voisin (Sciences Po), ainsi qu'avec l'IGÉSR.

Ce présent rapport résume les observations issues de la première collecte de données portant sur l'année scolaire 2024-2025 et permet de documenter la mise en œuvre du dispositif et ses effets pour sa première année. Une nouvelle collecte de données est conduite au printemps 2026 auprès des établissements de l'échantillon EEC, afin d'analyser l'évolution des organisations pédagogiques et d'apprécier les conditions de mise en œuvre et les effets de la réforme au cours de sa deuxième année.

➤ Données et taux de réponse

Matériel d'enquête

Ce document de travail s'appuie en premier lieu sur les réponses aux questionnaires auto-administrés adressés au printemps 2025 aux chefs d'établissement et aux coordonnateurs intervenant pour les organisations pédagogiques mises en place en 6^e et/ou en 5^e pour l'année scolaire 2024-2025 en français et/ou en mathématiques (appelés « coordonnateurs » dans la suite du rapport). Les coordonnateurs ont été ciblés par l'enquête du fait de leur rôle de relais entre les équipes de direction et les équipes pédagogiques. Toutefois, ce statut varie fortement d'un établissement à l'autre, certains collèges de petite taille ne déclarent aucun coordonnateur alors que d'autres établissements avec des effectifs plus importants déclarent jusqu'à quatre coordonnateurs pour une même discipline.

L'échantillon initial des établissements contactés pour participer à l'étude EEC était composé de 1 676 collèges publics de France hors Mayotte. Ces collèges sont issus d'un échantillon principal stratifié de 1 500 établissements et de 176 établissements tirés aléatoirement à partir d'un échantillon de remplacement stratifié de 1 500 collèges, constitué pour pallier d'éventuels refus de participation des collèges de l'échantillon principal (voir **Annexe 1** pour plus de détails). Parmi ces collèges, 1 500 établissements ont répondu lors de la première sollicitation de la DEPP en février-mars 2025 les invitant à renseigner les informations de contact des coordonnateurs de français et de mathématiques intervenant pour les organisations pédagogiques mises en place en 6^e et en 5^e pour l'année scolaire 2024-2025, ce qui représente un taux de réponse de 89 %. Parmi les établissements qui ont renvoyé le formulaire de contact, 21 % ne déclarent aucun coordonnateur, 2 % n'en inscrivent qu'un seul, 58 % en indiquent exactement deux et 19 % en désignent au moins trois. Au total, 2 774 coordonnateurs ont été renseignés par les chefs d'établissements, dont 1 371 en lettres et 1 403 en mathématiques. Pour le français, 23 % des établissements qui ont renvoyé le formulaire ne déclarent aucun coordonnateur, 63 % en renseignent un seul et 14 % en indiquent deux ou plus. Pour les mathématiques, ces proportions sont respectivement de 22 %, 63 % et 15 %.

À la différence des collèges de l'échantillon de remplacement, les collèges de l'échantillon principal n'ayant pas répondu à cette première sollicitation n'ont pas été retenus pour la suite de l'enquête. L'échantillon auquel le questionnaire à destination des chefs d'établissement a été adressé est ainsi composé de 1 525 établissements, et 1 388 réponses ont été reçues, soit un taux de réponse de 91 %. Les questionnaires trop incomplets pour être utilisés à des fins statistiques ont été retirés de l'échantillon final, qui est composé de 1 315 questionnaires exploitables, soit 78 % de l'échantillon initial de 1 676 collèges. Une repondération de l'échantillon final a été réalisée afin qu'il reste représentatif de la base de sondage (voir **Annexe 2**). Dans l'ensemble, la non-réponse aux différentes étapes de l'enquête n'a pas déformé la structure de l'échantillon : les différences de moyennes standardisées entre les valeurs de l'échantillon initial pondéré et les valeurs de l'échantillon final avec non-réponse et sans pondération sont presque toutes inférieures au seuil de 0,1, qui représente un déséquilibre léger (**Figure 28** de l'**Annexe 2**). La seule exception concerne les collèges qui participent à l'expérimentation des évaluations nationales de début de 5^e en 2024, qui sont sur-représentés dans l'échantillon final non-pondéré par rapport à l'échantillon initial pondéré. La procédure de repondération appliquée à l'échantillon permet presque toujours de se rapprocher des valeurs moyennes de l'échantillon initial pondéré.

Les 2 774 coordonnateurs renseignés par les chefs d'établissement ont également été invités à répondre au questionnaire de l'enquête. Parmi eux, 1 587 ont renvoyé un questionnaire, ce qui représente un taux de réponse de 57 %. Cependant, seuls 1 421 questionnaires sont suffisamment complets pour être analysés, ce qui représente un taux de réponse final de 51 % par rapport à l'ensemble des coordonnateurs contactés. Sur les 1 371 coordonnateurs de français identifiés par les chefs d'établissement, 614 ont renvoyé un questionnaire suffisamment complet pour l'analyse, soit

un taux de réponse de 45 %. Pour les mathématiques, 807 réponses analysables ont été reçues pour les 1 403 questionnaires envoyés, soit un taux de réponse de 58 %. Au moins un coordonnateur a répondu dans 76 % des établissements où un coordonnateur ou plus a été identifié. Comme pour les chefs d'établissement, un modèle de repondération a également été appliqué aux questionnaires des coordonnateurs afin qu'ils restent représentatifs des 2 774 coordonnateurs désignés initialement (**Annexe 2**). Malgré des taux de réponse moindres que pour les chefs d'établissement, l'échantillon final des coordonnateurs ne diverge pas beaucoup de l'échantillon initial pondéré en termes de valeurs moyennes (**Figure 29** de l'**Annexe 2**). En effet, les seules différences de moyennes standardisées qui dépassent le seuil de 0,1 sont la discipline, qui présente un léger déséquilibre en raison des taux de réponse plus importants chez les coordonnateurs de mathématiques, et à nouveau les collègues qui participent à l'expérimentation des évaluations nationales de 5^e, qui répondent davantage. Là aussi, la procédure de repondération permet de se rapprocher des valeurs de l'échantillon initial pondéré.

Données administratives

Ce document de travail mobilise également des données produites par la DEPP et hébergées dans différents systèmes d'information. Les données de contexte mobilisées pour décrire les établissements (appartenance à l'éducation prioritaire, taille, résultats au diplôme national du brevet ou aux évaluations nationales de début de 6^e, etc.) sont les mêmes que celles qui ont été utilisées pour la construction de la base de sondage et pour le redressement statistique réalisé pour corriger la non-réponse. Il s'agit principalement des données du système d'information CYCLADES-DNB, des fichiers Scolarité, du Constat de rentrée, des évaluations nationales de 6^e et des bases Relais.

Afin de documenter les éventuels effets des regroupements d'élèves sur les performances scolaires des élèves, les données des évaluations nationales de début d'année sont également mobilisées. En particulier, toutes les évaluations nationales de début de 4^e disponibles sont mobilisées, ce qui couvre la période 2023-2025. Les évaluations de 6^e de 2021 à 2023 sont également exploitées afin d'avoir une idée du niveau des cohortes entrant en 4^e entre 2023 et 2025 lorsqu'elles sont entrées en 6^e, deux ans plus tôt. Enfin, afin de décrire l'évolution des compétences des élèves qui entrent en 5^e en septembre 2024, les données des évaluations Repères de début de CP de 2018, le point de mi-CP de 2019 et les évaluations Repères de début de CE1 de 2019 sont également utilisées (voir **Annexe 6** sur l'utilisation des données des évaluations nationales et la construction des scores standardisés).

Données extraites des logiciels de vie scolaire

Outre les questionnaires auto-administrés, il était demandé aux établissements dans le cadre du protocole d'enquête de transmettre des extractions de leur logiciel de vie scolaire afin de pouvoir associer chaque collégien de 6^e et de 5^e à ses éventuels groupes d'appartenance pour les enseignements de français et de mathématiques au cours de l'année scolaire 2024-2025. Il s'agit d'une source inédite de données qui n'avait jamais encore été exploitée dans les études de la DEPP, et qui nécessite un traitement statistique important afin d'harmoniser le format des données et de les apparier avec les autres données disponibles dans les systèmes d'information statistique de la DEPP.

Ce travail de préparation des données est encore en cours, leur exploitation dépasse donc le périmètre des analyses proposées dans ce document de travail. De futurs travaux de recherche permettront, à partir de ces données, de décrire plus finement la composition des différents regroupements selon les caractéristiques des élèves qui les composent. Ils permettront également de mesurer non seulement l'effet, sur les performances scolaires, de la scolarisation dans un établissement mettant en œuvre des regroupements d'élèves par niveau, mais aussi l'effet spécifique de l'appartenance à un groupe composé des élèves les plus fragiles ou, à l'inverse, des plus performants.

➤ Les organisations pédagogiques mises en œuvre en 2024-2025

Premiers éléments descriptifs

En raison de la grande flexibilité laissée aux établissements, malgré le cadre relativement strict présenté dans l'arrêté du 15 mars 2024, une première question porte sur la mise en œuvre effective du dispositif par les collèges et sur les formes d'organisation privilégiées pour l'année scolaire 2024-2025. Sauf indication contraire, la plupart des éléments présentés dans cette partie sont tirés de l'analyse des réponses des chefs d'établissement aux questionnaires qui leur ont été adressés.

De nouvelles organisations pédagogiques dans la quasi-totalité des établissements à la rentrée 2024

D'après les réponses des chefs d'établissement, les nouvelles organisations pédagogiques ont été très largement mises en œuvre : 95 % des collèges publics de France hors Mayotte ont ainsi mis en place des regroupements d'élèves à la rentrée 2024 tandis que 5 % ne déclarent aucune forme de regroupement pour les classes de 6^e et de 5^e, ni en français, ni en mathématiques (**Encadré 3**). Ces proportions sont cohérentes avec celles observées pour l'année scolaire 2024-2025 dans l'évaluation qualitative menée par l'IGÉSR (cf. **Encadré 2**). Ici, le terme « regroupements d'élèves » recouvre toutes les formes d'organisations pédagogiques impliquant que certains élèves des classes de 6^e ou de 5^e soient rassemblés au sein de groupes se distinguant des classes de référence pour tout ou partie des enseignements de français ou de mathématiques. Dans trois établissements sur dix, les élèves sont dans les mêmes groupes pour les deux disciplines, suivant une logique plus proche des classes de niveau que des groupes de besoins.

ENCADRE 3 : Les établissements qui n'ont pas mis en place d'organisation pédagogique particulière à la rentrée 2024

Une minorité d'établissements déclarent ne pas avoir mis en place de regroupements d'élèves à la rentrée 2024. D'après les questionnaires renseignés par les chefs d'établissement, cela concernerait 5 % des collèges du champ de l'étude.

Ces établissements sont relativement plus présents dans les communes rurales que ceux qui ont mis en place des groupes, et relativement moins implantés dans les communes de densité intermédiaire. Enfin, aucun de ces collèges ne se situent dans les départements et régions d'outre-mer.

En moyenne, les établissements qui n'ont pas mis en œuvre de regroupements sont de plus petite taille : ces collèges comptent en moyenne une soixantaine d'élèves de moins que ceux qui ont mis en place des regroupements. Leur indice de position sociale (IPS)* moyen est très légèrement supérieur, mais la différence n'est que marginalement significative (au seuil de 10 %). Enfin, leurs résultats au diplôme national du brevet sont supérieurs de presque un demi-point en moyenne par rapport aux établissements qui ont mis en place des regroupements d'élèves. D'après les réponses à l'enquête, les établissements qui n'ont pas mis en œuvre de regroupements d'élèves à la rentrée 2024 fonctionnaient également moins avec des répartitions d'élèves dans d'autres formes de groupes l'année précédente.

Parmi les motifs choisis par les chefs d'établissements pour expliquer l'absence de regroupements d'élèves, le refus des équipes pédagogiques est le plus fréquent : il est mentionné par 61 % d'entre eux. Viennent ensuite les contraintes d'emploi du temps (45 %), l'absence de besoin particulier (32 %) et l'incertitude sur la mise en œuvre effective de la mesure (32 %), puis l'inquiétude des

familles (30 %) et enfin le manque d'enseignants (23 %). Le manque de salles est plus rarement mentionné (11 %), tandis que le manque d'accompagnement des équipes ne l'est presque jamais. Enfin, 22 % des chefs d'établissement renseignent d'autres motifs que ceux proposés dans le questionnaire, principalement l'absence de moyens supplémentaires pour mettre en œuvre la réforme.

Cette question des moyens est abordée à travers d'autres questions dans l'enquête, en particulier autour de la présence ou non d'heures supplémentaires dans la dotation horaire globale (DHG). Ainsi, 51 % des établissements qui n'ont pas mis en place de regroupements d'élèves déclarent avoir reçu des heures supplémentaires, contre 71 % pour les autres établissements. Les établissements qui n'ont pas mis en place de regroupements et qui ont reçu des heures supplémentaires déclarent également avoir reçu deux heures de moins que les autres établissements, mais cette différence n'est significative qu'au seuil de 10 %.

Dans leurs commentaires, certains chefs d'établissements expliquent la non-mise en œuvre de la mesure en raison de leur arrivée récente ou de leur situation d'intérim dans l'établissement. Les chefs d'établissements ou adjoints qui répondent à l'enquête ont en moyenne une année d'expérience de moins au sein des établissements qui n'ont pas mis en place de groupes, que ce soit en termes d'expérience dans l'établissement ou dans leur fonction. Ce constat reste valide lorsque l'échantillon est restreint aux questionnaires effectivement remplis par les chefs d'établissements eux-mêmes. Parmi eux, 42 % sont en poste depuis un an ou moins dans leur établissement au moment de l'enquête, contre 24 % pour les chefs d'établissements dans les collèges qui ont mis en place des regroupements d'élèves.

Enfin, dans leurs commentaires, les chefs des établissements qui n'ont pas mis en place de regroupements indiquent parfois avoir privilégié d'autres formes d'organisation, comme la continuité de dispositifs expérimentés au cours des années précédentes ou la mise en œuvre de pratiques de co-intervention (deux enseignants prenant en charge simultanément les élèves d'une même classe de référence) dans les classes de 6^e ou de 5^e. D'après leurs réponses, 95 % des établissements qui n'ont pas mis en œuvre de regroupements d'élèves ont opté pour une autre organisation, contre 80 % pour les autres établissements. Si les deux types d'établissements ont mis en place des heures de soutien supplémentaires ou ont étendu le dispositif « Devoirs faits » dans des proportions similaires, les établissements qui n'ont pas mis en place de regroupements ont davantage renforcé l'accompagnement personnalisé ou mis en œuvre des pratiques de co-intervention.

*L'IPS d'un établissement scolaire est un indicateur permettant de résumer les conditions socio-économiques et culturelles des élèves qu'il accueille. Il permet de rendre compte des disparités sociales entre établissements scolaires ainsi qu'à l'intérieur d'entre eux. Pour plus d'informations sur l'IPS et la construction de cet indicateur, voir Dauphant et al. (2023).

Un tiers des établissements avaient déjà adopté un fonctionnement avec des regroupements d'élèves au cours des années précédentes. Cependant, dans la très grande majorité des cas, l'entrée en vigueur de la réforme dite des « groupes de besoins » a induit une modification des modalités de regroupement : seuls 14 % des collèges qui constituaient déjà des groupes l'année précédente les ont conservés sans en modifier la nature.

Malgré le caractère relativement nouveau de ces regroupements d'élèves pour la majorité des établissements, l'enquête montre que les nouvelles organisations pédagogiques ont souvent été mises en œuvre dès la rentrée 2024 et préparées en amont. Dans la plupart des cas, il n'a donc pas été jugé nécessaire de maintenir les élèves en classe de référence au cours des premières semaines, le temps de définir les groupes. Ainsi, dans plus de huit établissements sur dix, ces regroupements

ont été définis dès la rentrée ou en amont de celle-ci pour les classes de 6^e. En 5^e, cette proportion est plus proche de neuf établissements sur dix.

La question des critères d'affectation des élèves à chacun des groupes a été posée aux coordonnateurs. Ainsi, neuf coordonnateurs sur dix indiquent que les équipes pédagogiques ont utilisé leur connaissance des élèves pour les répartir dans les groupes en 5^e, contre cinq sur dix pour la 6^e, ce qui souligne le rôle prédominant de ce facteur sur d'autres éléments comme les résultats scolaires obtenus en classe entière, mobilisés par six coordonnateurs sur dix pour les deux niveaux. D'après les coordonnateurs, les enseignants de 6^e se sont d'ailleurs appuyés davantage sur leurs propres évaluations que sur les résultats obtenus aux évaluations nationales de début d'année en 6^e, qui sont mentionnés par moins de cinq coordonnateurs sur dix pour le français et moins de quatre coordonnateurs sur dix pour les mathématiques. En 5^e, les évaluations nationales de début d'année n'avaient pas été généralisées à la rentrée 2024, ce qui explique qu'elles ne soient mentionnées que par près d'un coordonnateur sur dix dans les deux matières. Pour la répartition des élèves en 6^e, les enseignants se sont aussi appuyés sur les éléments de connaissance des élèves transmis par les professeurs des écoles *via* le livret scolaire unique ou sur des échanges directs avec les équipes du premier degré, notamment dans le cadre des conseils école-collège : c'est ce qui est rapporté par cinq coordonnateurs sur dix pour le français et par quatre coordonnateurs sur dix pour les mathématiques. Dans leurs commentaires¹¹, certains coordonnateurs soulignent qu'une transmission d'information très institutionnalisée a été mise en place dès la fin de l'année scolaire 2023-2024, à travers par exemple des tests spécifiques organisés au mois de juin pour identifier le niveau des futurs collégiens. Toujours autour de l'enjeu de connaissance des élèves, les coordonnateurs indiquent également que les temps d'observation en classe entière ont davantage été mobilisés par les enseignants de 6^e que par les enseignants de 5^e : pour le français, cinq coordonnateurs sur dix estiment que ces temps d'observation ont été utilisés en 6^e contre quatre sur dix pour la 5^e ; et pour les mathématiques ces proportions sont respectivement de cinq coordonnateurs sur dix et de trois sur dix. Enfin, d'après quatre coordonnateurs sur dix, le comportement des élèves a joué un rôle dans leur affectation aux différents groupes, et plus fréquemment en classe de 5^e qu'en 6^e. C'est également un point qui ressort des commentaires des coordonnateurs, notamment à travers l'attention portée à la séparation d'élèves impliqués dans des situations de harcèlement.

L'exploitation des commentaires libres des coordonnateurs autour de la question des critères de regroupement dévoile d'autres facteurs qui n'étaient pas proposés parmi les modalités de réponse du questionnaire, comme la prise en compte du genre des élèves afin de garantir la mixité dans les groupes, mais également des considérations liées aux compétences psychosociales des élèves (autonomie, confiance en soi, motivation, aisance dans un groupe), au choix des élèves ou des familles, notamment lors du passage de la 6^e à la 5^e, ou encore à la prise en charge des élèves présentant des besoins éducatifs particuliers. Enfin, environ un commentaire sur six concernait des coordonnateurs indiquant ne pas avoir utilisé d'autre critère de regroupement que l'ordre alphabétique afin de garantir l'hétérogénéité des élèves au sein des groupes¹².

L'arrêté du 15 mars 2024 modifiant l'arrêté du 19 mai 2015 relatif à l'organisation des enseignements dans les classes de collège mentionnait des groupes pour l'ensemble des classes et des niveaux du collège (6^e et 5^e à compter de la rentrée 2024, puis 4^e et 3^e à compter de la rentrée 2025), sur tout l'horaire des enseignements communs de français et de mathématiques. Conformément à ces directives, les résultats de l'enquête indiquent que les établissements qui ont mis en œuvre des regroupements d'élèves les ont déployés de manière quasi généralisée à leur échelle pour les niveaux 6^e et 5^e. Plus de trois quarts des collèges ont ainsi adopté ces nouvelles organisations pédagogiques pour l'ensemble de leurs classes de 6^e et de 5^e, et cette proportion est relativement stable à travers

¹¹ L'enquête comportait des champs libres permettant aux répondants de compléter ou de préciser certaines réponses.

¹² Les résultats de la recherche soulignent cependant que des règles d'affectation des élèves entre les classes qui paraissent aussi arbitraire que l'ordre alphabétique ne suffisent pas à limiter la ségrégation au sein des établissements en l'absence d'autres mesures plus volontaristes pour promouvoir la mixité au sein des classes (Ly et al., 2014).

les niveaux et disciplines. Dans de nombreux cas, les dispositifs à destination des élèves à besoins éducatifs particuliers sont également intégrés à ces nouvelles organisations pédagogiques. Lorsque les sections d'enseignement général et professionnel adapté (Segpa) sont présentes dans les établissements, leurs élèves sont inclus dans les regroupements d'élèves dans deux établissements sur dix en 6^e et dans un établissement sur dix en 5^e. Concernant les élèves des unités localisées pour l'inclusion scolaire (Ulis) ou des unités pédagogiques pour élèves allophones nouvellement arrivés (UPE2A), leurs élèves sont intégrés aux regroupements d'élèves pour les deux niveaux dans la moitié des établissements où ces unités existent¹³.

Dans huit établissements sur dix, ces nouveaux regroupements d'élèves ont reçu des dénominations neutres, par exemple sous la forme de noms propres (« Thalès » et « Pythagore » en mathématiques, « Baudelaire » et « Molière » en français). Dans les autres établissements, les libellés étaient plus souvent explicites (« à besoin », « fragiles » ou « satisfaisants ») qu'implicites (« racine », « bourgeon », « fleur »).

En ce qui concerne les heures d'enseignement du français et des mathématiques en 6^e et 5^e, celles-ci ont toutes été réalisées en groupe dans plus des trois quarts des établissements qui ont mis en œuvre des regroupements d'élèves. Cette proportion est d'autant plus importante que ces regroupements concernent l'ensemble des élèves de chaque niveau. Plus précisément, pour chaque niveau et discipline, neuf chefs d'établissement sur dix indiquent que la totalité des heures ont été réalisées en groupes lorsque toutes les classes du niveau et de la discipline ont été concernées par les regroupements d'élèves.

L'arrêté du 15 mars 2024 prévoyait par ailleurs une réduction de la taille des groupes pour ceux qui accueillent les élèves les plus en difficulté. D'après neuf chefs d'établissement sur dix, les regroupements ont induit des réductions d'effectifs par rapport aux classes de référence, mais cette réduction n'a été dirigée exclusivement vers les élèves les plus en difficulté que dans un établissement sur deux.

La mise en place des organisations pédagogiques a nécessité des moyens supplémentaires

La mise en place des nouvelles organisations pédagogiques à la rentrée 2024 a nécessité de consacrer davantage d'heures aux enseignements de français et de mathématiques pour les classes de 6^e et de 5^e, en particulier lorsque les regroupements d'élèves se sont traduits par une réduction du nombre d'élèves en face des professeurs¹⁴. À titre indicatif, la création d'un groupe supplémentaire par rapport au nombre de classes de référence pour toute la durée de l'enseignement nécessite 4,5 heures hebdomadaires supplémentaires en français et en mathématiques pour le niveau 6^e, 4,5 heures également pour les français en 5^e et 3,5 heures pour les mathématiques en 5^e. Parmi les établissements qui ont mis en œuvre des regroupements d'élèves, sept sur dix déclarent avoir obtenu des heures supplémentaires dans leur dotation horaire globale (DHG), à hauteur de 13 heures annualisées par semaine en moyenne. Dans presque tous les établissements, ces heures supplémentaires ont été intégralement consommées par la mise en place des nouvelles organisations pédagogiques. Les établissements de plus grande taille et ceux dont les résultats aux évaluations nationales de début de 6^e en mathématiques sont les plus faibles sont davantage représentés parmi ceux qui ont obtenu des heures supplémentaires.

Le choix de réduire ou non les effectifs semble lié à la question des heures supplémentaires de DHG : d'après les réponses des chefs d'établissements où ont été mis en place des regroupements d'élèves, les collèges qui ont reçu des heures supplémentaires sont légèrement plus susceptibles d'avoir utilisé

¹³ Les taux de non-réponse aux questions portant sur ces sections et dispositifs sont particulièrement élevés : entre 15 et 20 % pour les Segpa et UPE2A selon le niveau, et autour de 10 % pour les Ulis. Ces résultats sont donc à interpréter avec prudence.

¹⁴ Baradji (2025) relève, au niveau national, pour l'année 2024-2025, une baisse de l'indicateur E/S comprise entre 1,5 et 2,4 élèves pour les enseignements de français et de mathématiques en 6^e et en 5^e.

les nouvelles organisations pédagogiques pour réduire le nombre d'élèves dans les groupes par rapport à leur classe de référence. De plus, les établissements qui ont obtenu des heures supplémentaires dans leur DHG sont relativement plus nombreux à avoir fait bénéficier l'ensemble des élèves de ces réductions d'effectif, tandis que les établissements qui n'ont pas obtenu d'heures supplémentaires ont plus souvent orienté ces baisses d'effectifs en direction des élèves les plus fragiles.

En raison des moyens nécessaires à leur mise en place, les regroupements ont parfois été instaurés aux dépens d'autres dispositifs. Les chefs d'établissements ont été interrogés sur la substitution d'heures nécessaires à la mise en place des regroupements d'élèves, en particulier concernant l'accompagnement personnalisé, la mise en place de demi-groupes en sciences et en technologie ou en langue et les options. Les taux de non-réponse sont particulièrement importants pour ces questions (de 10 à 30 %), peut-être en raison de l'absence de ces différents dispositifs pédagogiques au sein des établissements. Sur le sous-échantillon des répondants, 64 % des chefs d'établissement indiquent avoir supprimé des heures dédiées à l'accompagnement personnalisé, 50 % déclarent avoir supprimé des demi-groupes en sciences et technologie, contre 40 % pour les langues, et 17 % affirment avoir supprimé des options. Il n'y a pas de différence significative ici entre les établissements qui ont obtenu des heures supplémentaires de DHG et les autres, à l'exception des demi-groupes de sciences et de technologie, qui ont été un peu plus fréquemment abandonnés en l'absence d'heures supplémentaires de DHG.

Les réponses concernant l'accompagnement personnalisé sont cependant à manipuler avec prudence : près d'un quart des établissements qui répondent avoir diminué le nombre d'heures dédiées à l'accompagnement personnalisé indiquent par ailleurs avoir renforcé l'accompagnement personnalisé en 2024-2025. Parmi eux, 50 % déclarent avoir diminué le nombre d'heures d'accompagnement personnalisé exclusivement pour les élèves de 6^e ou de 5^e, et 14 % uniquement pour les élèves de 4^e et de 3^e. Cependant, dans 20 % des cas, les chefs d'établissement reconnaissent avoir renforcé l'accompagnement personnalisé tout en diminuant le nombre d'heures dédiées à ce dispositif pour l'ensemble des niveaux, ce qui souligne des subtilités d'organisation qu'il n'est pas possible de capturer à l'aide du questionnaire de l'enquête.

Les chefs d'établissements pouvaient également renseigner d'autres choix organisationnels réalisés afin de rendre possible les regroupements d'élèves. Parmi les éléments cités, on retrouve la suppression des dédoublements et des groupes dans d'autres disciplines, en particulier en histoire-géographie, la disparition d'ateliers culturels ou de dispositifs sportifs, l'abandon des enseignements de découverte des langues anciennes ou du numérique ou encore l'annulation d'activités proposées sur la pause méridienne. Par ailleurs, certains chefs d'établissement ont revu le fonctionnement interne de leur établissement en limitant les missions associées au rôle de professeur principal pour dégager des heures, en réduisant le nombre d'heures de DHG consacrées à la gestion administrative ou en fusionnant certains dispositifs comme les Segpa ou les classes bilingues dans les regroupements d'élèves.

Enfin, les moyens mobilisés pour mettre en place les regroupements d'élèves recouvrent également les actions d'accompagnement ou de formation déployées en faveur des enseignants pour faciliter leur prise en charge pédagogique du dispositif. D'après les questionnaires qui leur étaient adressés, 40 % des coordonnateurs estiment que les enseignants de leur établissement n'ont bénéficié d'aucune mesure de ce type, tandis que 45 % soulignent qu'au moins la moitié des enseignants ont été accompagnés ou formés et 15 % que seule une minorité d'enseignants a été formée.

Des regroupements moins flexibles que dans les consignes de l'arrêté du 15 mars 2024

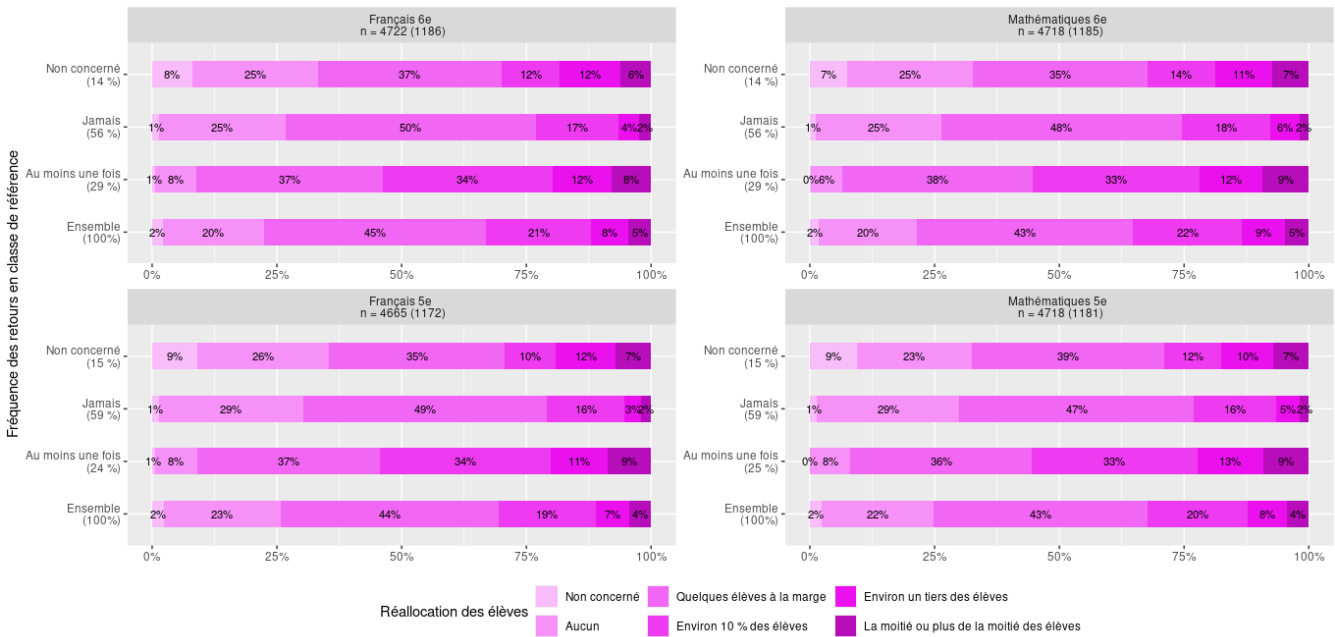
À la lecture des réponses aux questionnaires adressés aux chefs d'établissement, il apparaît que les formes de regroupement mises en œuvre en début d'année scolaire ont perduré jusqu'au moment

de l'enquête sans changements majeurs. Seule une minorité d'établissements a ainsi modifié les critères de regroupement en cours d'année (16 %).

Par ailleurs, l'arrêté du 15 mars 2024 permettait que les élèves soient rassemblés par classe de référence jusqu'à dix semaines dans l'année afin de garantir la cohérence des progressions pédagogiques et de réexaminer la composition des différents groupes pour tenir compte de la progression et des besoins des élèves, mais ces retours en classe de référence sont finalement assez rares. Les chefs d'établissement ont en effet été interrogés sur la fréquence des retours en classe de référence à l'échelle du niveau et de la discipline. D'après leurs réponses, 30 % des établissements ont réalisé des retours en classe de référence en 6^e et 25 % en 5^e. Dans huit établissements sur dix, il s'agit de retours en classe de référence pour une durée inférieure à 10 semaines. Ces proportions doivent toutefois être interprétées avec prudence. Elles ont été calculées sur l'ensemble des établissements ayant répondu à ces questions, y compris ceux qui ont choisi la modalité « Non concerné », proposée initialement pour les établissements n'ayant pas mis en place de regroupements d'élèves pour un niveau ou une discipline donnée. Or cette modalité représente ici une part non négligeable des réponses, environ 15 %, et a pu être comprise par certains répondants comme signifiant l'absence de retour en classe de référence. Cette ambiguïté invite donc à considérer ces résultats comme des ordres de grandeur plutôt que comme une mesure précise de la fréquence des retours.

En ce qui concerne les réallocations d'élèves entre les groupes, il apparaît qu'elles ne sont pas la norme, comme les retours en classe de référence (Figure 1). Selon les niveaux et disciplines, entre 20 % et 25 % des établissements n'ont réalisé aucun changement de groupe, et dans 44 % des établissements il n'y a eu que quelques changements à la marge. Dans 20 % des établissements, il y a eu des changements à hauteur d'un dixième des effectifs, et des changements plus importants n'ont été observés que dans 10 % à 15 % des collèges.

FIGURE 1 • Proportion d'élèves concernés par les changements de groupes selon l'organisation de retours en classe de référence, par niveau et discipline



Lecture : Parmi l'ensemble des établissements ayant répondu à l'enquête, 20 % déclarent n'avoir opéré aucune réallocation d'élèves entre les groupes pour les enseignements de français en 6^e. Cette proportion est de 25 % pour les établissements qui n'ont jamais mis en œuvre de retour en classe de référence et de 8 % pour les établissements au sein desquels ces retours ont eu lieu au moins une fois.

Champ : Collèges publics de France hors Mayotte, ouverts à la rentrée 2024 et accueillant au moins 20 élèves de niveau 6^e et 20 élèves de niveau 5^e. Pour ce graphique, seuls les établissements ayant déclaré avoir mis en place au moins une forme de regroupement ont été conservés.

Note : Les établissements non-répondants à l'une des questions pour un niveau et une discipline n'ont pas été pris en compte pour ce niveau et cette discipline, mais ont été conservés pour les autres niveaux et disciplines en cas de réponse. Les effectifs de répondants sont indiqués pour chaque niveau et discipline (les effectifs non pondérés figurent dans les parenthèses).

Source : DEPP, enquête EEC, questionnaires Chefs d'établissement, printemps 2025, données finales repondérées.

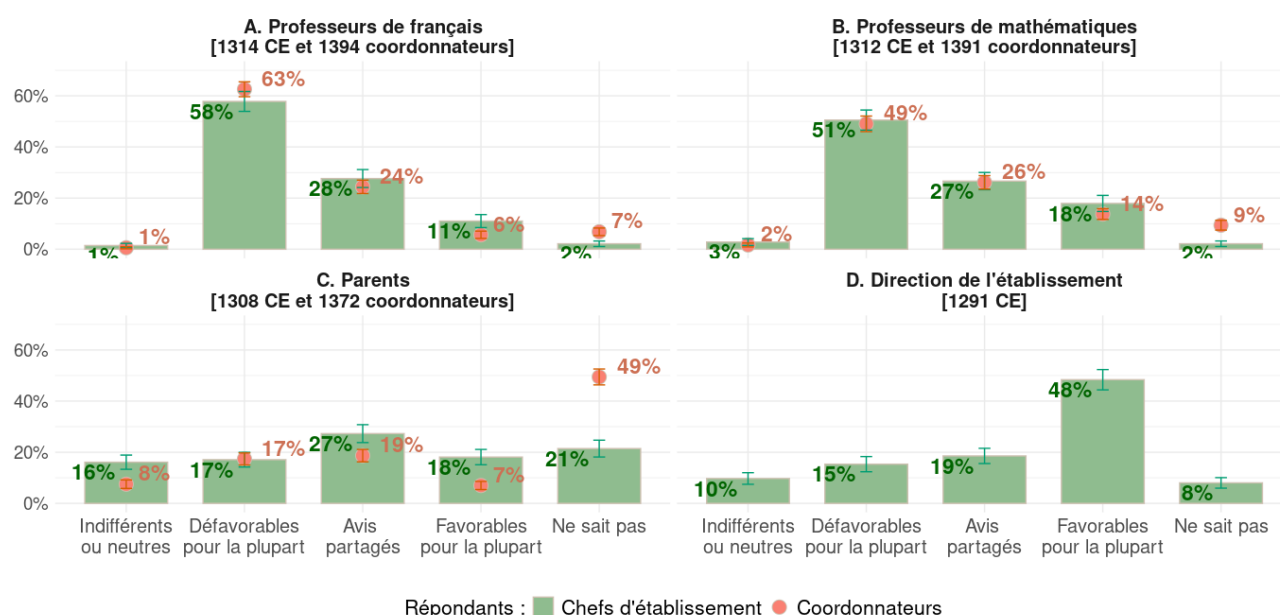
En pratique, les regroupements mis en place dès le début de l'année 2024 ont donc été assez peu remaniés en cours d'année. Cependant, ces résultats cachent une certaine hétérogénéité : les établissements qui ont rassemblé les élèves des différents groupes en classes de référence ont organisé davantage de réallocations d'élèves entre les groupes. Ainsi, lorsqu'il y a eu des retours en classe de référence au cours de l'année, des réallocations d'élèves ont été observées dans 92 % des collèges et, dans ce cas, la majorité des établissements souligne que ces réallocations ont concerné plus de 10 % des élèves. Par ailleurs, lorsqu'il y a eu des rassemblements en classe de référence, les changements de critères de regroupement en cours d'année sont également plus fréquents : ils sont observés dans plus d'un établissement sur quatre. Pour ces établissements, le retour en classe de référence pourrait donc constituer un moyen de redéfinir les besoins des élèves afin d'effectuer des changements de groupes et de s'emparer de la flexibilité allouée au dispositif.

En excluant les collèges qui se déclarent « non concernés » par les retours en classe de référence, il apparaît que les établissements qui ont organisé des retours en classe de référence sont plutôt de petits établissements situés dans des communes rurales et, dans une moindre mesure, dans des grands centres urbains, qui se caractérisent par une faible hétérogénéité des acquis scolaires à l'entrée en 6^e, et par des notes élevées ou basses aux évaluations de rentrée de 6^e en mathématiques ainsi qu'au diplôme national du brevet. Ces retours en classe de référence sont également plus fréquents lorsque le ou la cheffe d'établissement n'est pas dans sa première année au sein du collège et que sa prise de fonction à ce poste dans l'éducation nationale a débuté il y a entre six et vingt ans. Enfin, les établissements qui indiquent ne pas avoir reçu d'heures supplémentaires dans leur DHG sont légèrement plus nombreux à déclarer des retours en classe de référence pour une période d'une à dix semaines au cours de l'année.

Une perception de la réforme initialement très défavorable chez les enseignants

Les réponses des chefs d'établissement et des coordonnateurs concernant la perception de la mesure par les enseignants sont globalement cohérentes et soulignent que les équipes enseignantes étaient plutôt défavorables à la réforme lors de son annonce en décembre 2023. Ainsi, 58 % des chefs d'établissement et 63 % des coordonnateurs de français ou de mathématiques indiquent que la plupart des professeurs de français en classe de 6^e et de 5^e étaient initialement défavorables à la réforme (**Figure 2**). Pour les mathématiques, ces proportions sont légèrement inférieures, respectivement de 51 % et 49 %, ce qui souligne une adhésion légèrement supérieure à celle des professeurs de français dès l'annonce de la réforme.

FIGURE 2 • Perception des acteurs lors de l'annonce de la mesure selon les chefs d'établissement et les coordonnateurs de français et de mathématiques



Lecture : 63 % des coordonnateurs et 58 % des chefs d'établissement considèrent qu'au moment de l'annonce de la réforme, les professeurs de français de leur établissement avaient pour la plupart une perception défavorable des regroupements d'élèves.

Note : Les effectifs de répondants sont indiqués indépendamment pour chaque question. Il s'agit des effectifs non pondérés. Les segments autour des points représentent les intervalles de confiance à 95 %.

Champ : Collèges publics de France hors Mayotte, ouverts à la rentrée 2024 et accueillant au moins 20 élèves de niveau 6^e et 20 élèves de niveau 5^e.

Source : DEPP, enquête EEC, questionnaires Chefs d'établissement et Coordonnateurs, printemps 2025, données finales repondérées.

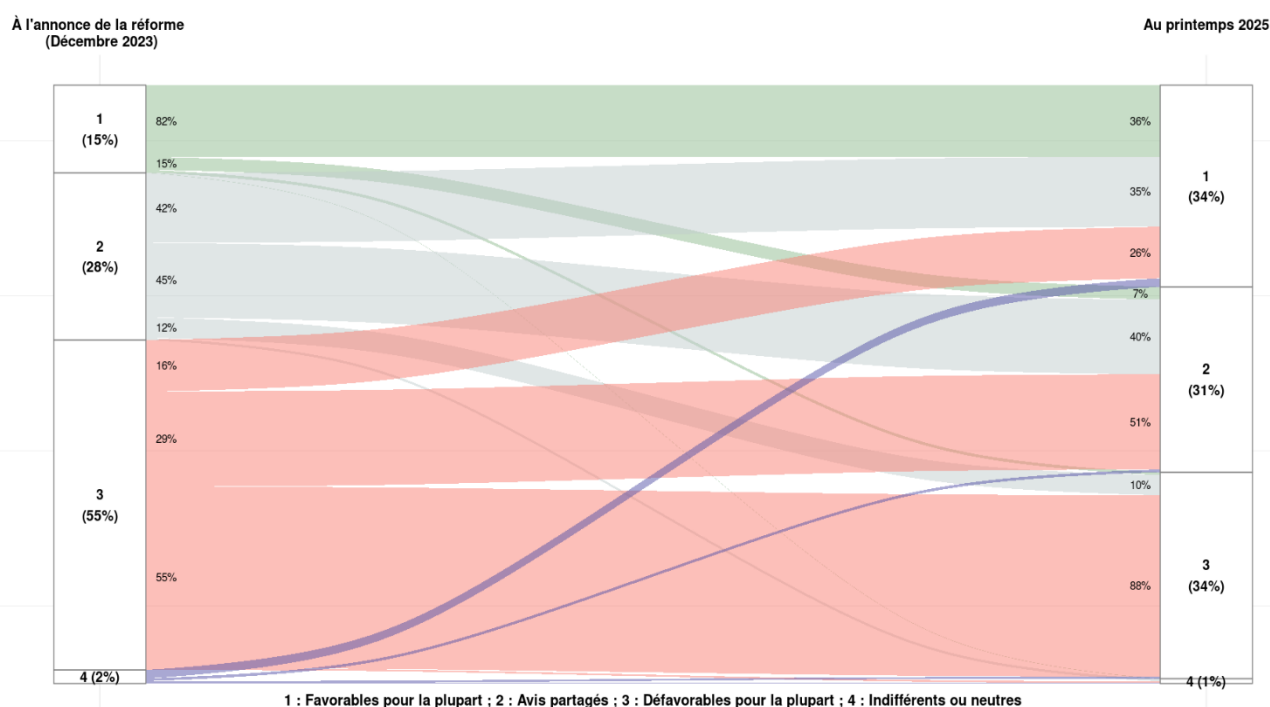
Les coordonnateurs comme les chefs d'établissement indiquent que les parents avaient une perception nettement moins défavorable à l'égard de la réforme, lors de son annonce, par rapport à celle des enseignants¹⁵. Enfin, les chefs d'établissement estiment, pour près de la moitié d'entre eux, que les équipes de direction étaient majoritairement favorables à cette réforme lorsqu'elle a été annoncée. Seuls 15 % des chefs d'établissement déclarent qu'elles y étaient alors majoritairement défavorables.

Une amélioration nette de la perception de la réforme chez les enseignants à la fin de la première année

D'après les chefs d'établissement, la perception de la réforme par les enseignants de français et de mathématiques intervenant en classes de 6^e ou de 5^e a été globalement meilleure à la fin de la première année qu'elle ne l'était lors de son annonce, en décembre 2023 : ils n'étaient plus que 34 % à penser que la plupart des enseignants de français ou de mathématiques de 6^e et de 5^e y étaient défavorables lors de l'enquête au printemps 2025, contre 55 % au moment de l'annonce de la réforme (Figure 3). Plus précisément, 16 % des chefs d'établissement qui déclaraient que les enseignants étaient pour la plupart défavorables à la réforme lors de son annonce ont rapporté que la plupart d'entre eux étaient favorables à la réforme au printemps 2025, et 29 % estimaient que les perceptions de la majorité des enseignants ont évolué de « défavorables » à « partagées ».

¹⁵ Il convient néanmoins de noter que près d'un coordonnateur sur deux déclare ne pas savoir se prononcer sur la perception des parents au sujet de la réforme, contre 21 % des chefs d'établissement.

FIGURE 3 • Évolution de la description par les chefs d'établissement de la perception des professeurs de français et de mathématiques en 6^e et en 5^e vis-à-vis des regroupements d'élèves



Lecture : 15 % des principaux des collèges ayant mis en place des regroupements d'élèves considèrent que les enseignants de français et de mathématiques en 6^e et en 5^e avaient une perception favorable vis-à-vis des regroupements d'élèves au moment de l'annonce de la mesure. La majeure partie d'entre eux déclarent que ces enseignants ont toujours une perception favorable de la mesure lorsqu'ils répondent au questionnaire au cours du printemps 2025 (82 %). Du fait de l'évolution des perceptions décrites par les chefs d'établissement, 34 % d'entre eux estiment que les enseignants de français et de mathématiques en 6^e et en 5^e ont une opinion favorable vis-à-vis de la mesure au moment où ils répondent au questionnaire. Les coordonnateurs qui indiquent que les enseignants de français et de mathématiques avaient déjà une perception favorable de cette réforme lors de son annonce représentent 36 % de ce groupe.

Note : Les proportions indiquées dans les étiquettes représentent les réponses des chefs d'établissement concernant la position des enseignants de français et de mathématiques au moment de l'annonce de la réforme et lors de la réponse au questionnaire. Il y a jusqu'à deux réponses par chef d'établissement (une par discipline). À gauche, les proportions présentées dans les flux correspondent à la distribution des perceptions déclarées au printemps 2025 (moment de la réponse au questionnaire) pour des chefs d'établissement qui partagent une même réponse quant aux perceptions à l'annonce de la réforme. À droite, les proportions présentées dans les flux correspondent à la représentation dans les perceptions déclarées lors de la réponse au questionnaire des réponses des chefs d'établissement quant à la situation à l'annonce de la réforme. Ces deux proportions ne sont pas affichées pour les flux qui représentent moins de 2 % des réponses des chefs d'établissement.

Champ : Chefs d'établissement de l'échantillon EEC s'étant prononcés à la fois sur la perception des enseignants de français ou de mathématiques lors de l'annonce de la mesure et sur leur perception au moment du questionnaire (hors « Je ne sais pas »). La question de la perception n'étant posée que pour les établissements déclarant avoir mis en place des regroupements d'élève, seuls ces établissements sont conservés pour ce graphique. Effectif : 1 199 chefs d'établissement sur 1 315, représentant 4 807 établissements après pondération.

Source : DEPP, enquête EEC questionnaires Chefs d'établissement, printemps 2025, données finales repondérées.

Inversement, 34 % des chefs d'établissement ont estimé que la plupart des enseignants de français et de mathématiques en 6^e et en 5^e étaient favorables à la réforme lors de l'enquête, alors qu'ils n'étaient que 15 % à indiquer que c'était le cas lors de l'annonce de la réforme. Certains chefs d'établissement ont cependant identifié une dégradation dans la perception de la réforme chez les enseignants de leur établissement : 18 % parmi ceux qui estimaient que les enseignants étaient majoritairement favorables et 13 % parmi ceux qui témoignaient d'avis majoritairement partagés lors de l'annonce de la réforme.

Enfin, au moment de répondre à l'enquête, 62 % des chefs d'établissement qui déclaraient ne pas être en mesure de prononcer sur les perceptions des enseignants lors de l'annonce de la réforme ont indiqué que la majorité des enseignants étaient favorables à cette réforme lors de l'enquête réalisée au printemps 2025.

Les coordonnateurs ont également été interrogés sur leur satisfaction concernant l'organisation retenue pour les regroupements d'élèves, selon des modalités différentes ne permettant pas de mesurer une progression directe de leur perception. Au moment où ils ont été interrogés, 57 % des coordonnateurs se déclaraient satisfaits par les organisations retenues, et une proportion similaire estimait que c'est également le cas pour les autres enseignants de français et de mathématiques en

6^e et en 5^e. En revanche, seuls 31 % des coordonnateurs ont répondu que les autres enseignants de l'établissement étaient satisfaits par ces nouvelles organisations pédagogiques.

D'après les commentaires libres des coordonnateurs et des chefs d'établissement, l'élément principal qui a contribué à la satisfaction des enseignants est la réduction des effectifs induite par la réforme¹⁶, ce qui aurait favorisé les échanges entre les élèves et les professeurs et aurait permis la mise en place de nouvelles pratiques pédagogiques. Pour certains, cette réforme a par ailleurs permis un meilleur accompagnement des élèves en difficulté et un plus grand respect de leur rythme d'apprentissage. Pour d'autres en revanche, la réforme a contribué à donner plus de visibilité à la grande difficulté scolaire, sans fournir aux enseignants les outils nécessaires pour mieux l'accompagner. À l'inverse, certains commentaires soulignent que la réforme a été appréciée parce qu'elle a introduit davantage de collaboration dans les équipes enseignantes. Toujours d'après l'analyse des commentaires transmis par les répondants, la moindre satisfaction des autres enseignants viendrait du fait qu'ils n'ont pas pu bénéficier de ces réductions d'effectifs, et qu'au contraire certains dispositifs existants, comme les demi-groupes, ont été abandonnés afin de dégager des moyens supplémentaires pour les enseignements de français et de mathématiques en 6^e et en 5^e.

Néanmoins, plusieurs chefs d'établissement soulignent dans leurs commentaires que leur établissement a bénéficié de moyens supplémentaires en 2024-2025 ou de conditions particulièrement favorables pour la mise en place de la réforme cette année-là, et qu'il n'est pas garanti que la réforme aurait été aussi appréciée sans ces moyens ou conditions favorables. De même, les enseignants mentionnent des arbitrages qui ne pourront pas être reconduits, ou des moyens qui risquent de ne pas être réattribués pour l'année scolaire 2025-2026.

Un impact perçu comme négatif sur le rôle de professeur principal et le suivi des élèves

Un commentaire récurrent, à la fois chez les chefs d'établissement et chez les coordonnateurs, porte sur les difficultés à désigner un professeur principal en classes de 6^e ou de 5^e après l'entrée en vigueur de la réforme. En effet, les enseignants de français et de mathématiques ne se sentent pas bien placés pour occuper ce rôle car, selon les organisations retenues, ils peuvent ne jamais enseigner devant la classe de référence, et donc ne pas connaître tous les élèves et avoir des difficultés à saisir le climat de classe.

Parmi les chefs d'établissements, 74 % déclarent que la mise en œuvre des regroupements d'élèves a eu un impact négatif sur le rôle du professeur principal. De plus, 57 % d'entre eux identifient également un impact négatif sur l'organisation des conseils de classe. Cela trouve un écho dans les commentaires des coordonnateurs, qui sont nombreux à souligner que la réforme a rendu difficile le suivi des élèves en conseil de classe. En effet, les professeurs ne peuvent souvent pas suivre l'ensemble des conseils de classe des classes dont sont issus les élèves présents dans leur groupe, et ne peuvent donc pas s'exprimer à leur sujet. Lorsque les professeurs participent à l'ensemble des conseils de classe, cela représente une charge très importante pour eux, alors qu'ils n'enseignent parfois qu'à un petit nombre d'élèves dans la classe de référence.

Ces différents éléments ont pu avoir une répercussion sur la communication au sein des équipes enseignantes, sur laquelle les avis des différents acteurs sont partagés. D'après 49 % des chefs d'établissement, les nouvelles organisations pédagogiques n'ont pas eu d'impact négatif sur la communication au sein de l'équipe pédagogique, 10 % ne se prononcent pas et 40 % estiment au

¹⁶ À la rentrée 2024, la part d'heures d'enseignement effectuées en groupes est de 31 % dans les collèges publics (+8,6 points par rapport à 2023) avec en 6^e et 5^e trois quarts des heures de français et de mathématiques dispensées en groupes. L'indicateur E/S (moyenne des effectifs par classes ou groupes) était de 21 élèves en moyenne pour le français et les mathématiques contre 23 élèves en moyenne toutes disciplines confondues pour les niveaux 6^e et 5^e (Baradji, 2025). L'exploitation des données issues de logiciels de vie scolaire permettra d'objectiver cette réduction des effectifs rapportée par les répondants aux questionnaires.

contraire qu'elles ont eu un impact négatif. D'après les commentaires des coordonnateurs et des chefs d'établissement, la mise en place des nouvelles organisations pédagogiques a entraîné une amélioration de la communication, notamment au sein de chaque discipline. En revanche, elle a pu se traduire par une surcharge de communication, notamment par courriel, puisque beaucoup plus d'acteurs se retrouvent concernés par la gestion des élèves d'une même classe ; certains coordonnateurs identifient également une complexification de la communication avec les enseignants des autres disciplines, qui se retrouvent eux-aussi face à une multiplicité d'interlocuteurs pour une même classe.

Enfin, la communication avec les enseignants des autres disciplines a pu être rendue plus difficile par l'abandon des projets interdisciplinaires au niveau de la classe, devenus trop compliqués à mettre en place en plus des nouvelles organisations pédagogiques. Ces abandons sont mentionnés par plusieurs coordonnateurs dans les commentaires, et 43 % des chefs d'établissement soulignent que la mise en œuvre des regroupements d'élèves a eu un impact négatif sur la réalisation de projets interdisciplinaires. Dans leurs commentaires, les chefs d'établissement ajoutent enfin comme difficulté supplémentaire le remplacement des enseignants en cas d'absence, plus difficile à mettre en œuvre du fait de la moindre disponibilité des professeurs pour des remplacements au niveau de l'établissement ou de l'académie.

Pour autant, ces nouvelles organisations pédagogiques ne semblent pas avoir détérioré la qualité des relations au sein des équipes pédagogiques dans la plupart des établissements : d'après la moitié des coordonnateurs, il n'y a pas eu de changement dans les relations entre les enseignants, tandis que 21 % d'entre eux perçoivent une évolution positive et 22 % une évolution négative, les coordonnateurs restants ne se prononçant pas.

Une collaboration entre les enseignants perçue comme renforcée, en particulier en français

Les coordonnateurs ont été interrogés sur les pratiques relatives à la collaboration entre les enseignants de leur discipline lors de l'année 2023-2024 et lors de l'année 2024-2025. Le premier constat est que les pratiques collaboratives en 2023-2024 étaient plus répandues chez les enseignants de mathématiques que chez les enseignants de français. En effet, 74 % des coordonnateurs de mathématiques indiquaient définir une progression commune à l'équipe disciplinaire pour un même niveau au cours de l'année 2023-2024, contre 38 % des coordonnateurs de français. Par ailleurs, 26 % des coordonnateurs de mathématiques déclaraient réaliser des évaluations communes régulières, contre 17 % des coordonnateurs de français. Les coordonnateurs se retrouvent cependant sur la question des progressions communes en matière de compétences majeures travaillées en français ou en matière de grandes thématiques travaillées en mathématiques : c'est le cas pour respectivement 59 % des coordonnateurs de français et 58 % des coordonnateurs de mathématiques.

Après la mise en place des nouvelles organisations pédagogiques, la collaboration entre les enseignants est perçue comme s'étant accrue, notamment en français. Ainsi, pour l'année scolaire 2024-2025, 76 % des coordonnateurs de français et 85 % des coordonnateurs de mathématiques déclarent avoir mis en place une progression commune à l'équipe disciplinaire pour un même niveau, contre respectivement 38 % et 74 % lors de l'année précédente. Par ailleurs, la proportion de coordonnateurs de français qui indiquent avoir défini une progression commune en matière de compétences majeures travaillées est passée de 59 % à 76 %, et la proportion de coordonnateurs de français réalisant en commun le choix des œuvres étudiées pour un même niveau est passée de 25 % en 2023-2024 à 37 % en 2024-2025. Cependant, d'après les commentaires des coordonnateurs, ce choix concerté des œuvres peut être rendu compliqué par le manque de moyens lorsqu'il n'est pas possible de fournir le même livre ou le même manuel à l'ensemble des élèves d'un même niveau.

Une exception importante à cette progression perçue des pratiques collaboratives concerne les grandes thématiques travaillées pour un même niveau chez les coordonnateurs de mathématiques.

La proportion de coordonnateurs déclarant mettre en place une progression commune autour de ces thématiques s'établit à 40 % en 2024-2025, contre 58 % l'année précédente. De même, la proportion de coordonnateurs de mathématiques qui déclarent des points de rencontre réguliers avec les autres enseignants de leur discipline est passée de 53 % en 2023-2024 à 43 % en 2024-2025. Les nouvelles organisations pédagogiques semblent donc avoir favorisé la collaboration en français, mais pas dans tous les domaines pour les mathématiques, ce qui pourrait traduire une différenciation plus marquée des enseignements dans cette seconde discipline.

Après la réforme, la difficulté scolaire perçue comme mieux prise en charge en mathématiques

La difficulté scolaire était bien identifiée avant la réforme : 83 % des coordonnateurs estiment que la nature des difficultés scolaires rencontrées par les élèves était déjà connue lors de l'année scolaire 2023-2024. En cela, il n'est pas possible de dire que les nouvelles organisations pédagogiques ont davantage mis en lumière la difficulté scolaire : la proportion de coordonnateurs qui fait le même constat en 2024-2025 n'a progressé que de quatre points de pourcentage.

En mathématiques, les coordonnateurs ont le sentiment que la réforme aurait été accompagnée d'un changement dans la gestion de cette difficulté scolaire. En effet, 74 % des coordonnateurs de mathématiques estiment que les élèves en difficulté sont bien pris en charge au cours de l'année 2024-2025, alors qu'ils ne sont que 65 % à faire ce constat pour l'année précédente. Par ailleurs, concernant l'année scolaire 2024-2025, 79 % des coordonnateurs de mathématiques déclarent que des modalités pédagogiques et didactiques spécifiques peuvent être mises en œuvre pour répondre aux difficultés des élèves, contre 71 % pour l'année 2023-2024. Ces déclarations progressent également chez les coordonnateurs de français, mais dans une moindre mesure. Cette progression doit toutefois être interprétée avec prudence : l'appréciation de l'année 2023-2024 est rétrospective et peut être affectée par un biais de mémoire ou de désirabilité sociale. Les écarts observés restent en outre modérés et ne permettent pas d'isoler un effet propre de la réforme.

Pour les coordonnateurs de français, les nouvelles organisations pédagogiques semblent avoir légèrement renforcé l'idée selon laquelle la prise en charge des élèves les plus fragiles soulève des enjeux pédagogiques ou didactiques spécifiques. Ils sont ainsi 79 % à partager ce constat pour l'année 2024-2025, contre 71 % pour l'année 2023-2024. La question de la prise en charge de la grande difficulté scolaire revient beaucoup dans les commentaires des coordonnateurs : certains enseignants se retrouvent face à des groupes composés des élèves les plus en difficulté sans bénéficier de formation particulière ou d'un accompagnement par du personnel spécialisé, et se voient eux-mêmes en grande difficulté pour faire progresser les élèves. Cette difficulté tient notamment au fait que ces groupes peuvent réunir des élèves à besoins éducatifs particuliers, des élèves présentant des problèmes de comportements, ou encore des élèves démotivés ou résignés. Ces configurations génèrent des situations auxquelles la seule différenciation pédagogique ne permet pas toujours de répondre.

Plus de 70 % des établissements ont mis en place un seul type de regroupement

La question de l'hétérogénéité des pratiques entre établissements souligne des différences d'appropriation de la réforme des organisations pédagogiques par les établissements. Pour documenter la diversité des modalités de déploiement de la mesure, le questionnaire administré demandait aux chefs d'établissement d'indiquer, pour chaque niveau et discipline concernés (6^e en français, 6^e en mathématiques, 5^e en français et 5^e en mathématiques) la ou les formes de regroupement mises en œuvre dans leur établissement. Les modalités proposées distinguaient : les groupes homogènes par discipline, les groupes réunissant des élèves de niveau similaire à la fois en français et en mathématiques, les groupes constitués autour d'une compétence ou d'un domaine spécifique des programmes, les groupes hétérogènes, les groupes destinés aux élèves les plus performants et les groupes spécifiques pour les élèves rencontrant le plus de difficultés dans la

matière (voir **Annexe 3**). Les éléments présentés dans cette partie s'appuient donc uniquement sur les réponses des chefs d'établissement.

Parmi les collèges indiquant avoir mis en place au moins une forme de regroupement à la rentrée 2024, 1 % ne se reconnaissent pas dans les catégories proposées par le questionnaire. D'après les commentaires laissés par ces établissements, ils auraient opté pour des formes plus ponctuelles de regroupement ou de co-intervention, uniquement pour une petite partie des heures ou des élèves.

Parmi les établissements qui n'ont mis en place qu'un type de regroupement pour les deux niveaux (6^e et 5^e) et les deux disciplines (français et mathématiques), trois formes d'organisation dominent nettement (**Tableau 2**, première colonne). La plus fréquente est la constitution de groupes hétérogènes, composés d'élèves de niveaux variés : 20,6 % des collèges l'ont retenue comme unique modalité de regroupement. Une proportion très proche, 19,2 %, a uniquement constitué des groupes homogènes selon le niveau des élèves dans la discipline. Les groupes destinés aux élèves rencontrant le plus de difficultés constituent la troisième modalité exclusive la plus fréquente, avec 15,3 % des établissements.

Les autres formes d'organisation sont nettement moins représentées lorsqu'elles sont considérées comme modalités uniques. Ainsi, près de 8 % des établissements ont mis en place des groupes fondés sur le niveau global des élèves dans les deux disciplines, et 7 % des groupes définis selon une compétence ou un domaine particulier. Les regroupements exclusivement destinés aux élèves les plus performants restent marginaux, avec moins de 1 % des établissements. Enfin, 6 % des collèges n'ont pas mis en place de regroupements identifiables dans les modalités proposées : 5 % déclarent n'avoir opéré aucun regroupement (**Encadré 3**) et 1 % indiquent des regroupements qui ne correspondent pas aux catégories du questionnaire.

Par ailleurs, 23 % des collèges ont expérimenté plusieurs modalités de regroupement au cours de l'année. Dans près de la moitié de ces établissements, les différences d'organisation tiennent uniquement à la discipline : les classes de 6^e et de 5^e suivent alors une même organisation au sein d'une discipline, mais celle-ci diffère entre le français et les mathématiques. Dans 30 % des cas, les regroupements varient selon le niveau, tout en conservant une organisation identique dans les deux disciplines. Les autres établissements combinent plusieurs formes de regroupement sans découpage aussi systématique par niveau ou par discipline.

TABLEAU 2 • Part d'établissements mettant en place les différents types de regroupements d'élèves

Type de regroupement	Part d'établissements (ensemble = 100 %)	Part d'établissements mettant en place la forme de regroupement pour au moins un niveau et une discipline
Plusieurs modalités de regroupements dans l'établissement	23,3 %	
Hétérogène	20,6 %	35,5 %
Homogène	19,2 %	32,0 %
Élèves en difficulté	15,3 %	28,8 %
Niveau scolaire dans les deux matières	7,7 %	11,4 %
Par compétence	7,0 %	14,7 %
Pas de groupe	6,2 %	14,0 %†
Élèves les plus performants	0,7 %	3,3 %
Modalités multiples pour un même niveau et une même discipline		2,9 %†

Lecture : Les établissements qui n'ont mis en place que des regroupements hétérogènes pour les classes de 6^e et de 5^e en français et en mathématiques représentent 20,6 % des établissements. Dans l'ensemble, 35,5 % des établissements ont mis en place au moins une forme de regroupement hétérogène pour un niveau ou une discipline. Près d'un quart des établissements ont mis en place plus d'une modalité de regroupement (23,3 %).

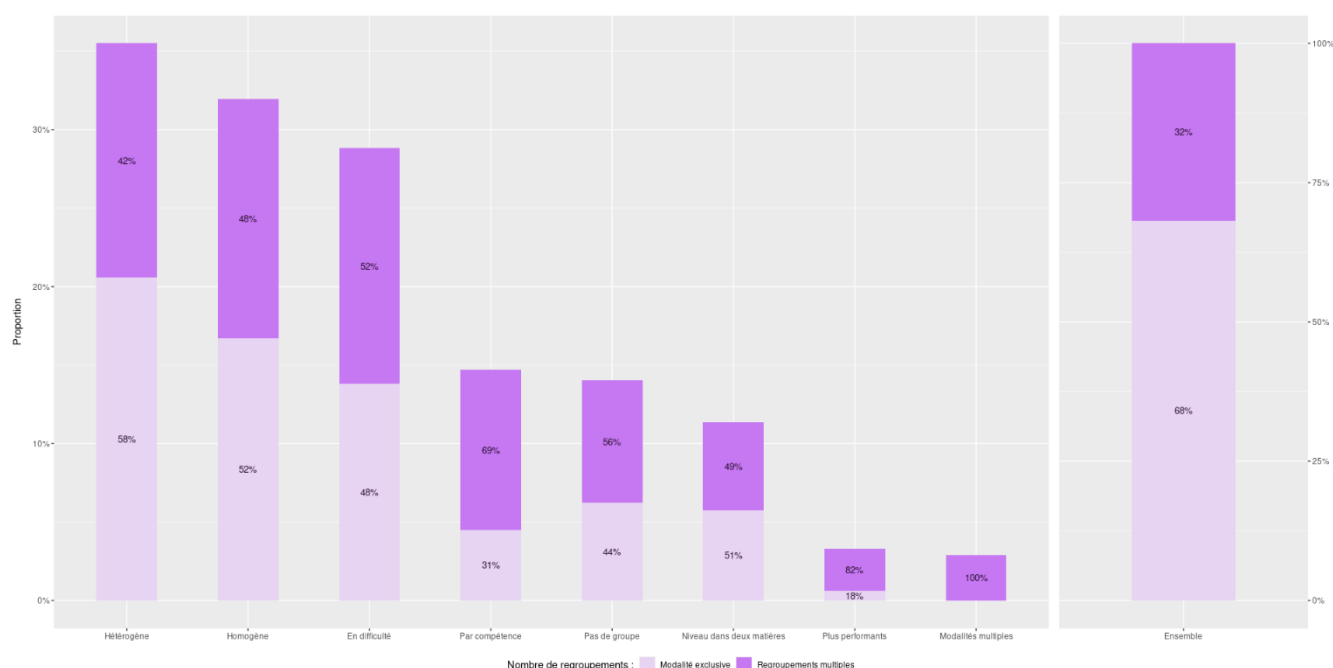
Champ : Collèges publics de France hors Mayotte, ouverts à la rentrée 2024 et accueillant au moins 20 élèves de niveau 6^e et 20 élèves de niveau 5^e. Effectif pondéré : 5 250. Effectif non pondéré : 1 315.

Note : † Dans la deuxième colonne, les proportions associées aux modalités « Pas de groupe » et « Modalités multiples » relèvent d'une définition légèrement différente du reste de la colonne. En effet, les autres chiffres indiquent la proportion d'établissements pour lesquels la modalité de regroupement a été choisie au moins une fois. Comme les réponses « Pas de groupe » et « Modalités multiples » n'étaient pas proposées dans le questionnaire, elles sont associées aux modalités construites dans cette étude à partir des réponses des chefs d'établissement. « Modalités multiples » représente ici la proportion d'établissements ayant mis en place des regroupements relevant de logiques à la fois homogènes et hétérogènes à l'échelle d'au moins un niveau et une discipline, comme l'enseignement de français en 5^e par exemple. Cette définition est également retenue dans le [Tableau 4](#). Un établissement qui déclare avoir mis en place à la fois des groupes avec les élèves en difficulté et des groupes homogènes pour un même niveau et une même discipline ne se retrouve pas dans cette modalité. La modalité « Pas de groupe » correspond ici aux établissements qui indiquent ne pas avoir mis en place de groupes du tout pour au moins une intersection entre le niveau et la discipline. Par exemple, un établissement qui a mis en place des regroupements pour toutes les combinaisons de niveau et de discipline, à l'exception du français en 6^e, est classé dans cette catégorie. En revanche, un établissement ayant constitué des groupes homogènes pour tous les niveaux et toutes les disciplines, mais pas nécessairement pour toutes les classes de chaque niveau et discipline, n'est pas classé dans cette catégorie, car cette information n'est pas disponible dans les données de l'étude.

Source : DEPP, enquête EEC, questionnaires Chefs d'établissement, printemps 2025, données finales repondérées.

Une lecture complémentaire consiste à identifier les modalités expérimentées au moins une fois dans l'établissement, indépendamment du fait qu'elles aient été ou non appliquées à tous les niveaux et dans les deux disciplines. Cette approche conduit logiquement à augmenter la part de chaque modalité, puisque celles-ci ne sont plus exclusives les unes des autres, mais elle ne modifie pas substantiellement la hiérarchie d'ensemble ([Tableau 2](#), colonne 2 ; [Figure 4](#)).

FIGURE 4 • Distribution des types de regroupements selon leur mise en œuvre exclusive ou avec d'autres modalités de regroupement



Lecture : 36 % des établissements déclarent avoir mis en place des groupes hétérogènes parmi les modalités pédagogiques présentes pour les enseignements de français et de mathématiques en classes de 6e ou de 5e à la rentrée 2024. Près de 60 % d'entre eux ont mis en place uniquement des groupes hétérogènes au niveau de l'établissement, tandis que 40 % d'entre eux ont également mis en place au moins une seconde forme de regroupement d'élèves.

Note : L'absence de regroupement pour un niveau ou une discipline n'est pas considérée comme une modalité de regroupement dans la définition des regroupements multiples. La modalité « Aucun regroupement » renvoie aux établissements qui déclarent avoir mis en place des regroupements d'élèves mais qui ne se reconnaissent pas dans les six modalités proposées.

Champ : Collèges publics de France hors Mayotte, ouverts à la rentrée 2024 et accueillant au moins 20 élèves de niveau 6^e et 20 élèves de niveau 5^e. Pour ce graphique, seuls les établissements ayant déclaré avoir mis en place au moins une forme de regroupement ont été conservés. Effectif pondéré : 5 250. Effectif non pondéré : 1 315.

Source : DEPP, enquête EEC, questionnaires Chefs d'établissement, printemps 2025, données finales repondérées.

Ainsi, 36 % des établissements ont expérimenté au moins une forme de regroupement hétérogène, contre 32 % pour les groupes homogènes et 29 % pour les groupes destinés aux élèves rencontrant le plus de difficultés. Ces résultats montrent que les formes dominantes de regroupement ne s'excluent pas nécessairement : elles sont souvent combinées avec d'autres organisations pédagogiques (**Tableau 3**). C'est le cas de 42 % des établissements ayant mis en place des groupes hétérogènes, de 48 % de ceux ayant constitué des groupes homogènes et de 52 % de ceux ayant créé des groupes pour les élèves les plus en difficulté. Les groupes définis selon une compétence ou un domaine particulier sont encore plus fréquemment associés à d'autres modalités : 69 % des établissements concernés ont combiné cette organisation avec au moins une autre forme de regroupement.

Les regroupements destinés aux élèves les plus performants restent, en revanche, très peu fréquents. À peine plus de 3 % des établissements en ont expérimenté au moins une fois, et ils sont presque toujours associés à d'autres dispositifs, le plus souvent à des groupes destinés aux élèves rencontrant le plus de difficultés (**Tableau 3**).

Les déclarations des coordonnateurs confirment globalement les tendances observées à partir des réponses des chefs d'établissement. D'après 93 % des coordonnateurs, des regroupements d'élèves ont été mis en œuvre au sein de leur établissement, et près des deux tiers d'entre eux déclarent qu'une seule organisation pédagogique a été instaurée.

TABLEAU 3 • Interaction entre les différentes formes de groupes au sein des établissements

Type de regroupement	Hétérogène	Homogène	Élèves en difficulté	Pas de groupe	Par compétence	Niveau scolaire dans les deux matières	Modalités multiples	Élèves les plus performants
Hétérogène	100 %	14 %	22 %	24 %	24 %	14 %	100 %	23 %
Homogène	12 %	100 %	18 %	21 %	31 %	26 %	41 %	13 %
Élèves en difficulté	18 %	16 %	100 %	15 %	26 %	14 %	36 %	61 %
Pas de groupe	10 %	9 %	7 %	100 %	8 %	6 %	15 %	6 %
Par compétence	10 %	14 %	14 %	8 %	100 %	8 %	52 %	24 %
Niveau scolaire dans les deux matières	4 %	9 %	5 %	5 %	6 %	100 %	29 %	15 %
Modalités multiples	8 %	4 %	4 %	3 %	10 %	8 %	100 %	4 %
Élèves les plus performants	2 %	7 %	7 %	5 %	5 %	5 %	5 %	100 %

Lecture : Parmi les établissements qui ont mis en place des groupes hétérogènes, 12 % au moins un groupe homogène, 18 % au moins un groupe avec des élèves en difficulté, 10 % n'ont pas défini de regroupement d'élèves pour au moins une discipline et/ou un niveau, 10 % ont constitué au moins un groupe selon le niveau pour une compétence ou un domaine particulier, 4 % ont défini au moins un groupe selon le niveau des élèves dans les deux disciplines, 8 % ont adopté des modalités de regroupement multiple pour au moins un même niveau et une même discipline, et 2 % ont constitué des groupes en rassemblant les élèves les plus performants.

Note : Voir la note du [Tableau 2](#) pour la définition de la catégorie « Modalités multiples ».

Champ : Collèges publics de France hors Mayotte, ouverts à la rentrée 2024 et accueillant au moins 20 élèves de niveau 6^e et 20 élèves de niveau 5^e. Effectif pondéré : 5 250. Effectif non pondéré : 1 315.

Source : DEPP, enquête EEC, questionnaires Chefs d'établissement, printemps 2025, données finales repondérées.

L'analyse par niveau et par discipline conduit à un constat similaire, tout en faisant apparaître une moindre importance des configurations faisant intervenir des modalités multiples qu'à l'échelle de l'établissement dans son ensemble ([Tableau 4](#)). Pour chaque niveau et chaque discipline, environ un quart des établissements ont mis en place des groupes hétérogènes, et une proportion légèrement inférieure des groupes homogènes. Ensemble, ces deux modalités concernent près de la moitié des collèges. Les groupes destinés aux élèves les plus en difficulté représentent une autre modalité importante, avec environ 19 % des établissements selon les niveaux et les disciplines.

Les autres formes de regroupement sont moins fréquentes. Environ 10 % des établissements ont constitué des groupes selon une compétence ou un domaine particulier, et une proportion légèrement inférieure des groupes fondés sur le niveau global des élèves dans les deux disciplines. Les groupes destinés aux élèves les plus performants demeurent très rares, tandis qu'environ 10 % des établissements déclarent ne pas avoir mis en place de regroupement pour le niveau et la discipline considérés, ou ne pas se reconnaître dans les modalités proposées par le questionnaire.

Les différences observées entre les niveaux et entre les disciplines sont limitées et ne sont pas statistiquement significatives. Il n'est donc pas possible d'établir l'existence de variations réelles selon que les regroupements concernent la 6^e ou la 5^e, le français ou les mathématiques. C'est cette nomenclature par niveau et par discipline qui sera retenue dans la suite du document de travail pour étudier la relation entre la forme de regroupement adoptée et l'évolution des performances des élèves.

TABLEAU 4 • Types de regroupements mis en place selon la discipline (français/mathématiques) et le niveau (6^e/5^e)

Type de regroupement	Français 6 ^e	Français 5 ^e	Mathématiques 6 ^e	Mathématiques 5 ^e
Hétérogène	25,9 %	25,2 %	24,8 %	25,0 %
Homogène	23,8 %	22,8 %	25,2 %	24,1 %
Élèves en difficulté	18,9 %	18,7 %	19,1 %	19,0 %
Par compétence	10,1 %	9,9 %	9,9 %	9,4 %
Pas de groupe	9,1 %	10,9 %	9,5 %	10,4 %
Niveau scolaire dans les deux matières	8,8 %	9,3 %	8,6 %	9,2 %
Modalités multiples	2,5 %	2,2 %	2,2 %	2,0 %
Élèves les plus performants	0,9 %	1,0 %	0,9 %	1,0 %
Ensemble	100 %	100 %	100 %	100 %

Lecture : Pour les classes de 6^e en français, 25,9 % des chefs d'établissements indiquent que seuls des regroupements hétérogènes ont été mis en place. La part des chefs d'établissement qui déclarent plusieurs modalités de regroupement pour les classes de 6^e en français est égale à 2,5 %.

Champ : Collèges publics de France hors Mayotte, ouverts à la rentrée 2024 et accueillant au moins 20 élèves de niveau 6^e et 20 élèves de niveau 5^e. Effectif pondéré : 5 250. Effectif non pondéré : 1 315.

Source : DEPP, enquête EEC, questionnaires Chefs d'établissement, printemps 2025, données finales repondérées.

Présentation de la typologie

Des choix organisationnels qui varient selon la taille des établissements et leur indice de position sociale

Attribuer une forme principale de regroupement aux établissements permet d'identifier les caractéristiques des collèges selon le type d'organisation pédagogique retenu (**Tableau 5**). Globalement, plusieurs facteurs semblent être corrélés aux choix des types de regroupements par les établissements : la densité de la commune d'implantation, la taille de l'établissement, le degré d'hétérogénéité de niveau des élèves à l'entrée en 6^e, la zone scolaire, l'indice de position sociale (IPS), les résultats moyens au diplôme national du brevet (DNB) et les résultats aux évaluations nationales de 6^e en français et en mathématiques.

Des tests du Chi-deux ont également été réalisés pour vérifier si les caractéristiques des établissements qui ont expérimenté une modalité de regroupement, de manière exclusive ou non, se distinguent de ceux qui ne mettent pas ce genre de regroupement en place. La p-value associée à ces tests d'une modalité de regroupement contre l'ensemble des autres peut être plus faible que celle présentée dans le **Tableau 5**, qui correspond au test du Chi-deux entre les caractéristiques des établissements et le détail des modalités exclusives mises en place au sein des établissements. Tous les résultats commentés dans le texte sont significatifs au moins au seuil de 10 %.

Il apparaît ainsi que les établissements qui ont mis en place des groupes hétérogènes se distinguent par une hétérogénéité élevée des scores à l'entrée en 6^e. De plus, ils sont peu représentés dans les DROM et leur IPS est plus élevé que la moyenne.

TABEAU 5 • Caractéristiques des établissements selon le type de regroupement

Caractéristique de l'établissement		Hétérogène (21 %)	Homogène (19 %)	Élèves en difficulté (15 %)	Par compétence (7 %)	Niveau scolaire dans les deux matières (8 %)	Modalités multiples (23 %)	Pas de groupe (6 %)	Ensemble† (100 %)	P-value (Chi-deux)
Appartenance à l'éducation prioritaire (EP)	Hors EP	86 %	80 %	80 %	83 %	69 %	75 %	84 %	80 %	0,32
	REP	9 %	13 %	13 %	12 %	22 %	17 %	13 %	14 %	
	REP +	5 %	7 %	7 %	6 %	9 %	8 %	4 %	7 %	
Densité de la commune du collège	Commune rurale	30 %	30 %	24 %	26 %	38 %	34 %	44 %	31 %	0,01
	Densité intermédiaire	35 %	42 %	33 %	39 %	35 %	31 %	23 %	35 %	
	Grands centres urbains	35 %	29 %	43 %	35 %	27 %	36 %	33 %	34 %	
Taille du collège (terciles)	Petite	29 %	34 %	22 %	26 %	47 %	38 %	46 %	33 %	0
	Moyenne	36 %	33 %	33 %	29 %	28 %	37 %	25 %	33 %	
	Grande	35 %	32 %	45 %	45 %	25 %	25 %	29 %	33 %	
Hétérogénéité des scores à l'entrée en 6 ^e (terciles)	Faible	22 %	38 %	27 %	42 %	37 %	38 %	38 %	33 %	0,01
	Moyenne	36 %	32 %	39 %	25 %	38 %	31 %	29 %	33 %	
	Forte	42 %	30 %	34 %	33 %	25 %	31 %	33 %	33 %	
IPS (terciles)	IPS faible	28 %	28 %	30 %	30 %	22 %	29 %	26 %	28 %	0,01
	IPS moyen	46 %	43 %	38 %	39 %	47 %	39 %	36 %	42 %	
	IPS élevé	25 %	17 %	29 %	27 %	25 %	28 %	38 %	26 %	
Moyenne FR/MA au DNB	Score faible	1 %	12 %	2 %	4 %	5 %	4 %	0 %	4 %	0,1
	Score moyen	24 %	36 %	31 %	40 %	44 %	37 %	25 %	33 %	
	Score élevé	37 %	33 %	27 %	39 %	28 %	34 %	33 %	33 %	
Évaluations de français à l'entrée en 6 ^e (terciles)	Score faible	38 %	31 %	42 %	21 %	27 %	29 %	42 %	33 %	0,05
	Score moyen	26 %	36 %	33 %	36 %	39 %	38 %	23 %	33 %	
	Score élevé	39 %	31 %	29 %	38 %	32 %	32 %	34 %	33 %	
Évaluations de mathématiques à l'entrée en 6 ^e (terciles)	Score faible	35 %	32 %	38 %	27 %	29 %	30 %	43 %	33 %	0,04
	Score moyen	26 %	33 %	29 %	37 %	41 %	41 %	30 %	33 %	
	Score élevé	38 %	33 %	33 %	31 %	38 %	29 %	30 %	33 %	
E/S (terciles)	E/S 2023 faible	37 %	34 %	38 %	33 %	20 %	30 %	40 %	33 %	0,42
	E/S 2023 moyen	26 %	34 %	33 %	40 %	40 %	38 %	23 %	33 %	
	E/S 2023 élevé	36 %	32 %	31 %	34 %	41 %	30 %	38 %	33 %	

Lecture : Parmi les établissements mettant en place des groupes hétérogènes, 86 % sont situés en dehors de l'éducation prioritaire, 9 % sont en REP et 5 % sont en REP+. La relation entre l'appartenance au réseau de l'éducation prioritaire et la modalité unique de regroupement mise en place dans l'établissement n'est pas significative (p-value de 0,32).

Note : †Les établissements qui mettent en place des groupes constitués des élèves les plus performants ne sont pas affichés car ils représentent une catégorie trop petite pour être analysée. En revanche, ils sont pris en compte dans la catégorie « Ensemble ».

Champ : Collèges publics de France hors Mayotte, ouverts à la rentrée 2024 et accueillant au moins 20 élèves de niveau 6^e et 20 élèves de niveau 5^e. Effectif pondéré : 5 250. Effectif non pondéré : 1 315.

Sources : DEPP, enquête EEC, questionnaires Chefs d'établissement, printemps 2025, données finales repondérées ; Évaluations nationales à l'entrée en 6^e 2023, CYCLADES-DNB 2023, Constat de rentrée 2023.

À l'inverse, les établissements qui ont mis en œuvre des groupes homogènes se caractérisent par une homogénéité plus forte des résultats scolaires à l'entrée en 6^e et sont sur-représentés dans les DROM. Leur IPS est inférieur en moyenne à celui des autres établissements, de même que les notes obtenues par leurs élèves aux évaluations de début de 6^e en français et en mathématiques.

Les établissements ayant instauré des groupes selon le niveau des élèves pour une compétence ou un domaine spécifique se caractérisent également par des IPS faibles et des notes au DNB ou aux évaluations de français en 6^e moins élevées que dans les autres établissements.

Les collèges qui ont défini des groupes selon le niveau des élèves dans les deux disciplines sont sous-représentés dans les grands centres urbains. Il s'agit plus fréquemment de petits établissements avec des scores moyens relativement bas aux évaluations de début de 6^e. Ils se retrouvent également davantage parmi les établissements avec des IPS bas, bien que cette relation ne soit significative qu'au seuil de significativité de 10 %.

Les établissements ayant mis en place des groupes à destination des élèves les plus en difficulté sont rarement des établissements ruraux de petite taille et plus fréquemment des établissements de grande taille situés dans les grands centres urbains. Il n'y a pas assez d'établissements ayant instauré des groupes à destination des élèves les plus performants pour que certaines caractéristiques ressortent de manière significative.

La caractéristique principale des collèges ayant mis en place plusieurs types de regroupement est leur sur-représentation parmi les collèges avec les scores moyens les plus bas aux évaluations nationales de français en début de 6^e. Cependant, cette relation n'est que marginalement significative.

Enfin, les collèges qui n'ont pas mis en place de groupes se retrouvent davantage parmi les petits établissements et ceux situés en zone rurale : ces deux caractéristiques se retrouvent d'ailleurs dans les commentaires des chefs d'établissement concernés, qui les associent généralement à des moyens insuffisants pour mettre en place la réforme.

Détail des modalités organisationnelles

Les nouvelles organisations pédagogiques mises en place à partir de la rentrée 2024 ne se caractérisent pas uniquement par le regroupement d'élèves de niveaux homogènes ou hétérogènes : elles peuvent également s'accompagner d'une réduction des effectifs dans les groupes, ciblée ou non sur les élèves les plus en difficulté, mais également par des retours en classe de référence ou par des réallocations d'élèves entre les groupes en cours d'année. Les choix organisationnels renseignés par les chefs d'établissement témoignent d'une appropriation de la réforme au-delà des simples regroupements d'élèves selon leurs besoins. Il s'agit désormais d'identifier comment les différentes organisations pédagogiques définies au niveau des établissements se traduisent par ces différentes modalités organisationnelles.

Cette section s'appuie sur la partie du questionnaire destinée aux établissements dans lesquels au moins une forme de regroupement d'élèves a été mise en place. Les établissements qui ne déclarent aucun regroupement d'élèves sont donc exclus de cette analyse. Par ailleurs, les établissements qui déclarent des regroupements mais qui ne se retrouvent pas dans les catégories proposées dans le questionnaire ainsi que les établissements qui déclarent uniquement des groupes à destination des élèves les plus performants sont également exclus car ils représentent des catégories trop petites pour pouvoir être efficacement mobilisées dans les analyses statistiques.

Constitution des regroupements

Bien que, dans la grande majorité des établissements, toutes les classes de 6^e et de 5^e aient été concernées par les nouvelles organisations pédagogiques pour le français et les mathématiques, la proportion de chefs d'établissements qui déclarent une mise en place de regroupements pour l'ensemble des classes de 6^e et de 5^e est relativement plus importante pour les collèges avec des groupes composés selon le niveau dans les deux disciplines, des groupes homogènes ou des regroupements constitués selon le niveau pour une compétence ou un domaine particulier (neuf établissements sur dix). Au contraire, seule la moitié des établissements qui mettent en place des modalités multiples de regroupement instaurent des groupes pour l'ensemble des classes de 6^e et de 5^e en français et en mathématiques.

Les groupes organisés selon le niveau dans les deux disciplines ou selon le niveau pour une compétence spécifique se distinguent par ailleurs par la plus forte intégration des élèves issus des Segpa ou des Ulis (pour les regroupements selon une compétence particulière) dans les nouvelles organisations pédagogiques. En termes de volume horaire, les collèges avec des groupes constitués selon une compétence ou un domaine spécifique ou des modalités multiples indiquent moins souvent que la totalité des heures sont concernées par ces nouvelles organisations pédagogiques. À l'inverse, les groupes définis selon le niveau des élèves dans les deux disciplines et les groupes hétérogènes sont plus souvent associés à un plus fort volume horaire réalisé sous forme de groupes.

Réduction des effectifs

La quasi-totalité des établissements a accompagné ces nouvelles organisations pédagogiques d'une réduction d'effectif. 81 % des établissements avec des groupes hétérogènes ont choisi de réduire les effectifs pour tous les élèves, indépendamment de leur niveau scolaire, et seuls 7 % d'entre eux ont spécifiquement ciblé les élèves les plus en difficulté. En revanche, les établissements ayant opté pour des regroupements d'élèves de même niveau, à l'exception des groupes ciblant uniquement les élèves les plus en difficulté, ont majoritairement réduit les effectifs dans les groupes composés des élèves rencontrant le plus de difficultés. À nouveau, les commentaires des chefs d'établissement et des coordonnateurs soulignent que le facteur principal ayant permis la réduction d'effectif est l'attribution de moyens supplémentaires, davantage observée au sein des établissements avec des regroupements hétérogènes.

Dans la plupart des cas, les regroupements d'élèves sont distincts selon les disciplines : les élèves ne sont ainsi pas dans les mêmes groupes en français et en mathématiques. Cependant, ce n'est pas le cas pour 62 % des collèges avec des regroupements hétérogènes, qui fonctionnent avec des groupes à effectif réduit pour les deux disciplines, ni pour 54 % des collèges avec des groupes de niveau, qui se rapprochent de la définition traditionnelle des classes de niveau. Ces fonctionnements en groupes hétérogènes ou de niveau figés sont souvent mis en place dès la rentrée, et rarement après. En revanche, les autres formes de regroupement sont relativement plus souvent mises en place après la rentrée, en particulier pour les groupes à destination des élèves les plus fragiles ou pour les groupes constitués selon le niveau pour une compétence ou un domaine spécifique, bien que la majorité des regroupements restent définis dès la rentrée. Les mises en œuvre plus tardives des regroupements s'observent davantage en 6^e qu'en 5^e, ce qui peut traduire le temps nécessaire pour connaître les élèves, non seulement en termes de résultats scolaires mais également de comportement, afin d'en tenir compte dans la constitution des groupes, comme cela a pu être mentionné dans les commentaires des chefs d'établissement ou des coordonnateurs.

Retour en classe de référence et flexibilité des regroupements

Les établissements avec des groupes hétérogènes se distinguent par une absence plus marquée des retours en classe de référence, pour plus de 80 % d'entre eux, et par les rares réallocations d'élèves entre les groupes : dans 90 % des cas, il n'y en a eu aucun ou très peu au cours de l'année. Les établissements avec des groupes homogènes ou des groupes constitués selon le niveau dans les deux disciplines ont également opéré peu de retours en classe de référence, mais néanmoins davantage que pour les groupes hétérogènes. Les réallocations d'élèves qui y ont été observées, bien qu'impliquant un nombre relativement faible d'élèves (moins d'un tiers des élèves dans neuf établissements sur dix où des réallocations ont été observées), ont concerné plus de 90 % des établissements, contre moins de 50 % pour les établissements avec des groupes hétérogènes. Par rapport aux établissements avec des groupes hétérogènes, ces établissements ont davantage nommé les groupes d'après des libellés explicites.

Les groupes organisés selon le niveau des élèves pour une compétence spécifique ou un domaine particulier représentent la modalité d'organisation la plus flexible : dans les établissements qui les ont mis en place, les retours en classe de référence y ont été davantage observés, et pour des durées plus longues. De même, les réallocations d'élèves y ont été plus fréquentes, et ont concerné davantage d'élèves, jusqu'à plus d'un tiers des élèves dans près de 40 % de ces établissements. Enfin, les établissements qui ont mis en place cette modalité de regroupement sont relativement plus nombreux à mentionner des changements de critères de regroupement en cours d'année (22 % des collèges, contre moins de 15 % pour les autres) et à expérimenter des pratiques de co-intervention (38 %, contre 33 % pour les groupes constitués selon le niveau dans les deux disciplines, 28 % pour les groupes à destination des élèves les plus en difficulté et autour de 20 % pour les autres établissements).

Les établissements avec des groupes à destination des élèves les plus en difficulté ont effectué davantage de retours en classe de référence que les établissements ayant mis en place des groupes homogènes (36 % contre 29 %). Les réallocations d'élèves y ont été légèrement plus rares mais elles ont plus souvent concerné davantage d'élèves lorsqu'elles ont eu lieu. D'après les réponses au questionnaire, il n'est pas possible de distinguer le temps passé en classe de référence après la mise en place des organisations pédagogiques du temps qui a été pris en classe de référence en amont de leur constitution.

Enfin, les établissements qui ont mis en place des modalités multiples ont davantage réalisé des courts temps en classe de référence, et ont aussi procédé à davantage de réallocations d'élèves que la plupart des autres établissements. Un tiers d'entre eux déclarent également avoir changé les critères de regroupement en cours d'année, contre moins de 20 % globalement pour les autres établissements, ce qui peut témoigner d'une volonté d'expérimentation ou d'un abandon de certaines organisations en cours d'année. Ces établissements se démarquent par ailleurs par la mise en place d'autres dispositifs en plus des regroupements d'élèves, légèrement plus fréquente que pour les autres établissements, en particulier en ce qui concerne la généralisation du dispositif « Devoirs faits » ou la proposition d'heures de soutien supplémentaires.

Le **Tableau 6** ci-après propose une synthèse des typologies de regroupement observées à l'échelle nationale.

TABEAU 6 • Récapitulatif des typologies de regroupement selon plusieurs dimensions de mise en œuvre

Type de regroupement	Réduction des effectifs	Ampleur des réallocations d'élèves	Degré de temps passé en classe de référence	Moment de constitution des groupes	Composition des groupes
Hétérogène	Oui (89 %)	Très faible - aucune réallocation d'élèves (54 %) ou uniquement à la marge (36 %).	Aucune semaine en classe de référence (84 %)	Précoce - dès la rentrée ou avant (90 %)	Souvent les mêmes groupes dans les deux disciplines (62 %)
Homogène	Oui (86 %) - plutôt ciblée sur les élèves en difficulté (dans 78 % des cas)	Faible - quelques réallocations à la marge (54 %) ou environ 10 % des élèves (28 %)	Faible – aucune semaine (71 %) ou moins de dix semaines (25 %)	Précoce - dès la rentrée (90 %)	Des groupes différents dans les deux disciplines (87 %)
Niveau dans les deux matières	Oui (84 %) - plutôt ciblée sur les élèves en difficulté (dans 74 % des cas)	Faible - quelques réallocations à la marge (57 %) ou environ 10 % des élèves (23 %)	Faible – aucune semaine (73 %) ou moins de dix semaines (25 %)	Précoce - dès la rentrée (95 %)	Les mêmes groupes dans les deux disciplines (54 %)
Par compétence	Oui (86 %) - plutôt ciblée sur les élèves en difficulté (dans 66 % des cas)	Réallocations fréquentes (98 %), également partagées entre quelques élèves à la marge, 10 % de l'effectif et plus d'un tiers des élèves	Élevé – Au moins une semaine (47 %)	Parfois plus tardif (18 %)	Des groupes différents dans les deux disciplines (88 %)
Élèves en difficulté	Oui (89 %) - plutôt ciblée sur les élèves en difficulté (dans 64 % des cas)	Faible - quelques réallocations à la marge (49 %) ou environ 10 % des élèves (25 %), présence de collègues avec réallocations pour plus d'un tiers des élèves (14 %).	Intermédiaire – aucune semaine (64 %) ou moins de dix semaines (31 %)	Parfois plus tardif (17 %)	Des groupes différents dans les deux disciplines (86 %)

Champ : Collèges publics de France hors Mayotte, ouverts à la rentrée 2024 et accueillant au moins 20 élèves de niveau 6^e et 20 élèves de niveau 5^e, qui ont mis en place des regroupements hétérogènes, homogènes, constitués selon le niveau dans les deux disciplines, constitués selon le niveau pour une compétence ou un domaine particulier ou constitués de manière à rassembler les élèves les plus en difficulté. Effectif pondéré : 3 666. Effectif non pondéré : 915.

Source : DEPP, enquête EEC, questionnaires Chefs d'établissement, printemps 2025, données finales repondérées.

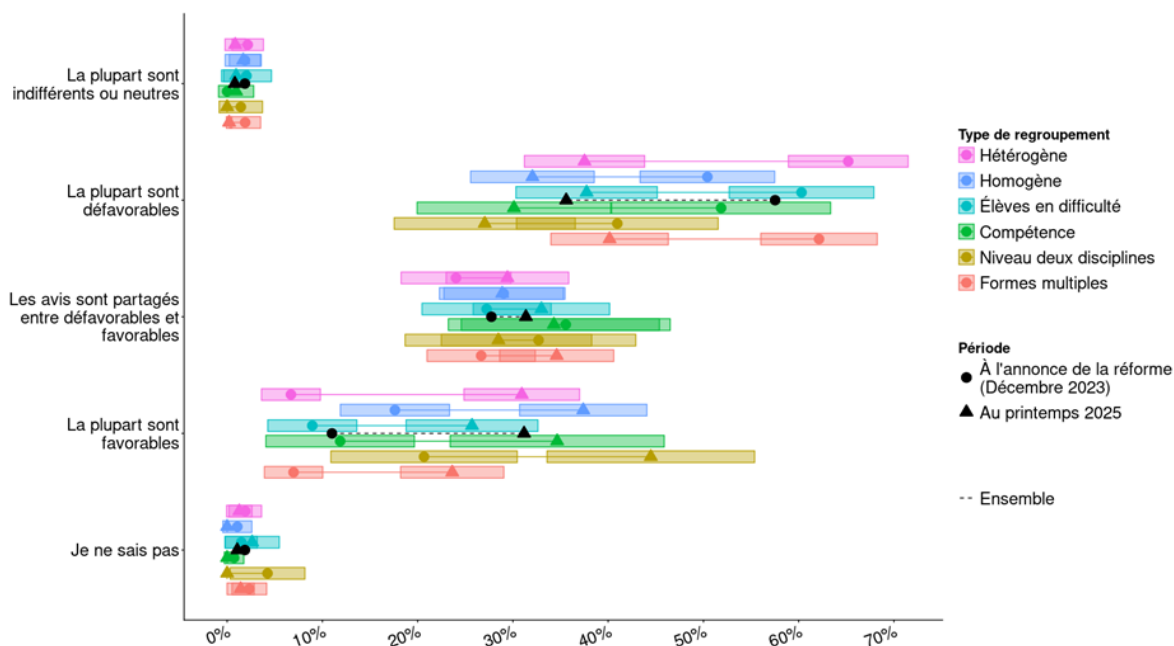
Analyse de l'hétérogénéité des réponses aux questionnaires « chefs d'établissement » selon la nature des regroupements mis en place

Un premier usage de la typologie décrite dans la section précédente s'attache à étudier les réponses aux questionnaires renseignés par les chefs d'établissement selon la nature des organisations pédagogiques retenues par l'établissement.

Une amélioration de la perception de la réforme quel que soit le type de regroupement, davantage marquée en cas de groupes hétérogènes

Du point de vue des chefs d'établissement, la perception de la réforme par les enseignants de français ou de mathématiques s'est améliorée entre son annonce et la réponse au questionnaire de l'enquête (**Figures 5 & 6**). Cela concerne notamment les établissements qui ont mis en place des groupes hétérogènes : l'adhésion y était la plus faible lors de l'annonce de la réforme (bien que la différence avec les groupes constitués des élèves les plus en difficulté ou les formes de regroupement multiples ne soit pas statistiquement significative), et elle a fortement progressé jusqu'au moment de l'enquête, pour rejoindre un niveau comparable à celui observé dans l'ensemble des établissements. Dans leurs commentaires, les chefs d'établissement et les coordonnateurs justifient cette plus grande satisfaction par la possibilité d'enseigner à des groupes de plus petite taille sans avoir à séparer les élèves selon leur niveau initial.

FIGURE 5 • Évolution de la perception de la réforme par les enseignants de français, du point de vue des chefs d'établissement selon le type de regroupement mis en œuvre



Lecture : D'après les chefs d'établissement, 65 % des enseignants de français étaient défavorables à la mesure lors de son annonce dans les établissements ayant mis en place des groupes hétérogènes. Lors de la passation du questionnaire, cette proportion est descendue à 38 %. Pour tous les établissements, ces proportions sont respectivement de 58 % et de 36 % en moyenne.

Note : Les zones rectangulaires autour des points représentent les intervalles de confiance à 95 %.

Champ : Chefs d'établissement de l'échantillon EEC s'étant prononcés sur les perceptions des enseignants de français lors de l'annonce de la mesure puis lors de la réponse à l'enquête. La question de la perception n'étant posée que pour les établissements déclarant avoir mis en place des regroupements d'élèves, seuls ces établissements sont conservés pour ce graphique. Les établissements ayant mis en place uniquement des groupes à destination des élèves les plus performants ne sont pas assez nombreux pour être étudiés statistiquement et ne sont pas affichés. Ils sont néanmoins pris en compte pour le calcul de la catégorie « Ensemble ». Taille de l'échantillon retenu : 1 240 établissements hors pondération, représentant 4 947 établissements avec pondération.

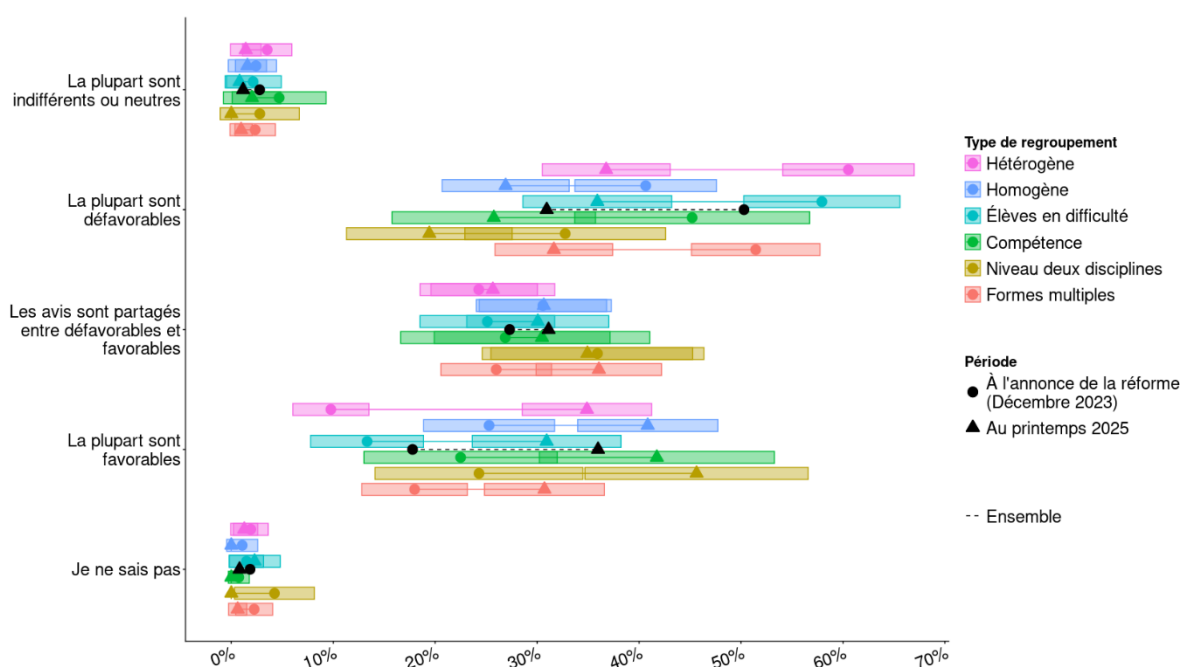
Source : DEPP, enquête EEC, printemps 2025, données finales repondérées.

Les établissements qui ont mis en place des groupes homogènes ou des groupes de niveau accueillent la plus grande proportion d'enseignants de français favorables avant la réforme, et cela est toujours le cas au moment de l'enquête, soit à l'issue d'une année scolaire de mise en œuvre des nouvelles organisations pédagogiques (**Figure 5**). La plupart des différences avec les autres établissements ne sont cependant que faiblement significatives. Les groupes de besoins présentent des niveaux

d'adhésion assez élevés également, mais la faible précision de ces estimations ne permet pas de les distinguer des établissements avec des modalités de regroupement multiples ou des groupes à destination des élèves les plus fragiles. Bien que l'adhésion ait progressé partout, elle reste minoritaire et varie autour de 30 % selon les modes de regroupement. Globalement, il est possible d'observer le passage d'une situation dans laquelle une forte majorité d'enseignants étaient défavorables à la réforme à une situation plus équilibrée dans laquelle près d'un tiers d'enseignants y sont toujours défavorables, un tiers d'enseignants sont partagés, et un tiers d'enseignants y sont favorables.

En mathématiques, la part d'enseignants favorables lors de l'annonce de la réforme était légèrement supérieure à celle observée pour les enseignants de français (Figure 6). L'amélioration des perceptions est également marquée pour tous les types de regroupement, en particulier pour les collèges avec des groupes hétérogènes, mais la précision de ces estimations varie beaucoup selon le nombre d'établissements concernés par chacun des regroupements. Cependant, le constat d'une distribution tripolaire lors de la passation de l'enquête reste valide, avec même peut-être une faible avance des enseignants favorables par rapport à ceux qui restent partagés ou défavorables.

FIGURE 6 • Évolution de la perception de la réforme par les enseignants de mathématiques, du point de vue des chefs d'établissement selon le type de regroupement mis en œuvre



Lecture : D'après les chefs d'établissement, 61 % des enseignants de mathématiques étaient défavorables à la mesure lors de son annonce dans les établissements ayant mis en place des groupes hétérogènes. Lors de la passation du questionnaire, cette proportion est descendue à 37 %. Pour tous les établissements, ces proportions sont respectivement de 50 % et de 31 % en moyenne.

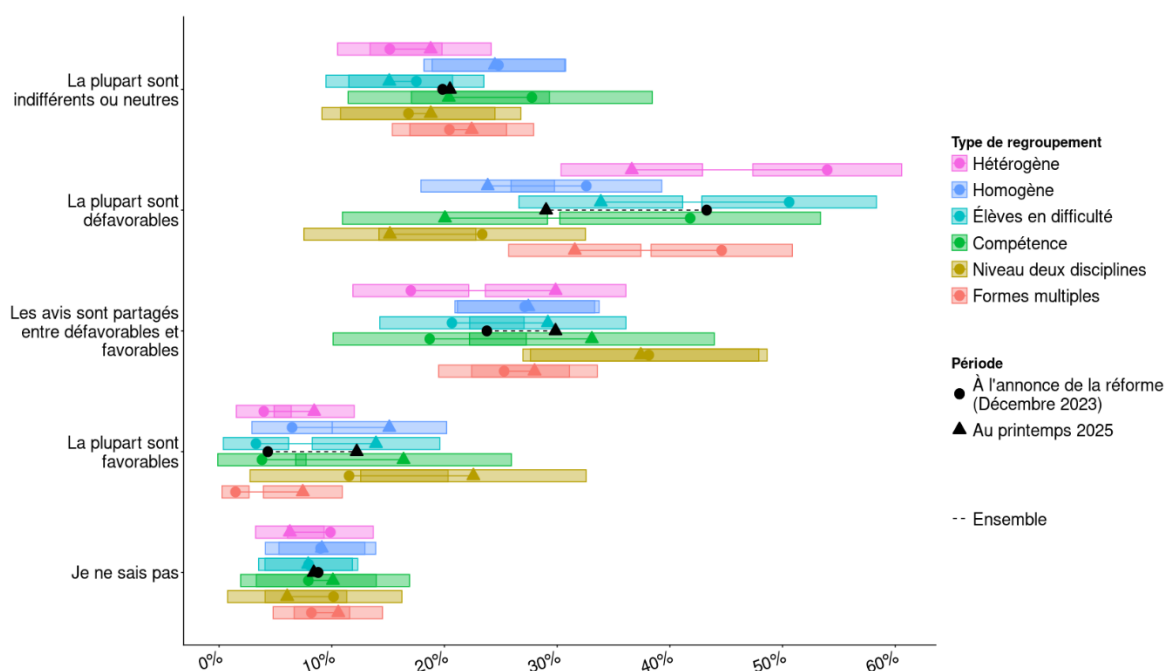
Note : Les zones rectangulaires autour des points représentent les intervalles de confiance à 95 %.

Champ : Chefs d'établissement de l'échantillon EEC s'étant prononcés sur les perceptions des enseignants de mathématiques lors de l'annonce de la mesure puis lors de la réponse à l'enquête. La question de la perception n'étant posée que pour les établissements déclarant avoir mis en place des regroupements d'élèves, seuls ces établissements sont conservés pour ce graphique. Les établissements ayant mis en place uniquement des groupes à destination des élèves les plus performants ne sont pas assez nombreux pour être étudiés statistiquement et ne sont pas affichés. Ils sont néanmoins pris en compte pour le calcul de la catégorie « Ensemble ». Taille de l'échantillon retenu : 1 240 établissements hors pondération, représentant 4 952 établissements avec pondération.

Source : DEPP, enquête EEC, printemps 2025, données finales repondérées.

Selon les chefs d'établissement, la perception de la réforme par les enseignants des autres niveaux ou disciplines s'est également améliorée entre l'annonce de la réforme et la passation des questionnaires, bien qu'elle reste à des niveaux d'approbation nettement inférieurs à ceux observés en français ou en mathématiques (Figure 7). D'après les chefs d'établissement, entre 7 % et 23 % d'entre eux sont favorables à cette réforme tandis qu'entre 15 % et 37 % y sont toujours défavorables. Par rapport aux mathématiques ou au français, le mouvement le plus visible pour les enseignants des autres niveaux ou disciplines est l'augmentation de la proportion d'enseignants dont les avis sont partagés, d'après les chefs d'établissement.

FIGURE 7 • Évolution de la perception de la réforme par les enseignants des autres niveaux ou disciplines, du point de vue des chefs d'établissement selon le type de regroupement mis en œuvre



Lecture : D'après les chefs d'établissement, 54 % des enseignants des autres niveaux et disciplines étaient défavorables à la mesure lors de son annonce dans les établissements ayant mis en place des groupes hétérogènes. Lors de la passation du questionnaire, cette proportion est descendue à 37 %. Pour tous les établissements, ces proportions sont respectivement de 43 % et de 29 % en moyenne.

Note : Les zones rectangulaires autour des points représentent les intervalles de confiance à 95 %.

Champ : Chefs d'établissement de l'échantillon EEC s'étant prononcés sur les perceptions des enseignants de mathématiques lors de l'annonce de la mesure puis lors de la réponse à l'enquête. La question de la perception n'étant posée que pour les établissements déclarant avoir mis en place des regroupements d'élèves, seuls ces établissements sont conservés pour ce graphique. Les établissements ayant mis en place uniquement des groupes à destination des élèves les plus performants ne sont pas assez nombreux pour être étudiés statistiquement et ne sont pas affichés. Ils sont néanmoins pris en compte pour le calcul de la catégorie « Ensemble ». Taille de l'échantillon retenu : 1 216 établissements hors pondération, représentant 4 855 établissements avec pondération.

Lecture : DEPP, enquête EEC, printemps 2025, données finales repondérées.

L'avis des chefs d'établissement sur les effets de la réforme dépend de la nature des groupes et du niveau des élèves

Il a été demandé aux chefs d'établissement s'ils pensaient que les dispositifs mis en place ont eu des effets sur les acquis des élèves. Cette appréciation subjective ne peut pas se substituer à une mesure d'impact, telle que réalisée plus bas, mais elle permet de préciser la réception de la politique dans les établissements.

De manière générale, les chefs d'établissement pensent que les nouvelles organisations pédagogiques ont eu un impact positif sur les acquis des *élèves les plus en difficulté* : c'est ainsi le cas de 57 % d'entre eux (Figure 8). Ils sont également 58 % à faire état d'effets positifs sur la confiance des élèves les plus en difficulté en leur capacité à réussir, et 49 % à ressentir une amélioration de leur investissement dans le travail scolaire.

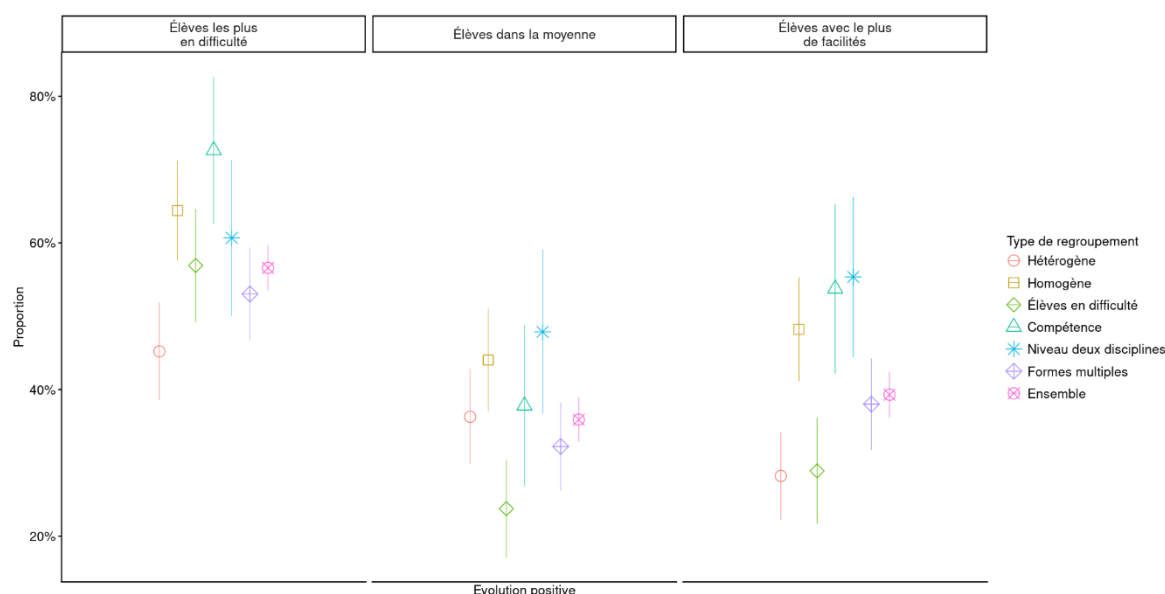
La proportion des chefs d'établissement qui identifient un impact positif des nouvelles organisations pédagogiques sur ces différents éléments varie peu selon les types de regroupements mis en place : la principale différence se situe entre les chefs d'établissements des collèges avec des regroupements hétérogènes et les autres : alors que les premiers ne sont que 45 % à déclarer que les regroupements d'élèves ont eu un impact positif sur les acquis scolaires des élèves les plus en difficulté, cette proportion est de l'ordre de 60 % pour les seconds.

Concernant les autres élèves, les positions des chefs d'établissement sont plus mitigées. Pour les élèves situés dans la moyenne, un peu plus d'un tiers des chefs d'établissement estiment que les regroupements d'élèves ont eu un effet positif sur leurs acquis, quel que soit le type de regroupement

mis en œuvre. Cette proportion d'ensemble est légèrement supérieure pour les élèves avec le plus de facilités.

Lorsque les chefs d'établissement n'identifient pas d'effet positif de la réforme, c'est qu'ils estiment globalement qu'elle n'a pas eu d'impact sur les acquis des élèves, ou qu'ils ne sont pas en mesure de se prononcer sur cette question : moins de 10 % de chefs d'établissement pensent qu'il y a des effets négatifs des nouvelles organisations pédagogiques sur les acquis des élèves. Ces constats se retrouvent au sujet de la confiance des élèves en leur capacité à réussir ou dans leur investissement dans le travail scolaire.

FIGURE 8 • Pourcentage de chefs d'établissement percevant un impact positif de la mesure sur les acquis des élèves, selon le type de regroupement mis en œuvre



Lecture : Parmi les établissements qui ont constitué des groupes selon le niveau pour une compétence ou un domaine spécifique, 73 % estiment que la mesure a eu un impact positif sur les acquis des élèves les plus en difficulté. Pour l'ensemble des établissements, cette proportion est de 57 % en moyenne.

Champ : Chefs d'établissement de l'échantillon EEC s'étant prononcés sur l'impact perçu de la mesure. Les questions relatives aux impacts perçus n'étant posées que pour les établissements déclarant avoir mis en place des regroupements d'élèves, seuls ces établissements sont conservés pour ce graphique. Les établissements ayant mis en place uniquement des groupes à destination des élèves les plus performants ne sont pas assez nombreux pour être étudiés statistiquement et ne sont pas affichés. Ils sont néanmoins pris en compte pour le calcul de la catégorie « Ensemble ». Taille de l'échantillon retenu : 1 223 établissements pour la question sur les élèves les plus en difficulté (4 881 avec pondération), 1 220 pour la question sur les élèves dans la moyenne (4 860 avec pondération) et 1 220 pour la question sur les élèves avec le plus de facilités (4 863 avec pondération).

Source : DEPP, enquête EEC, printemps 2025, données finales repondérées.

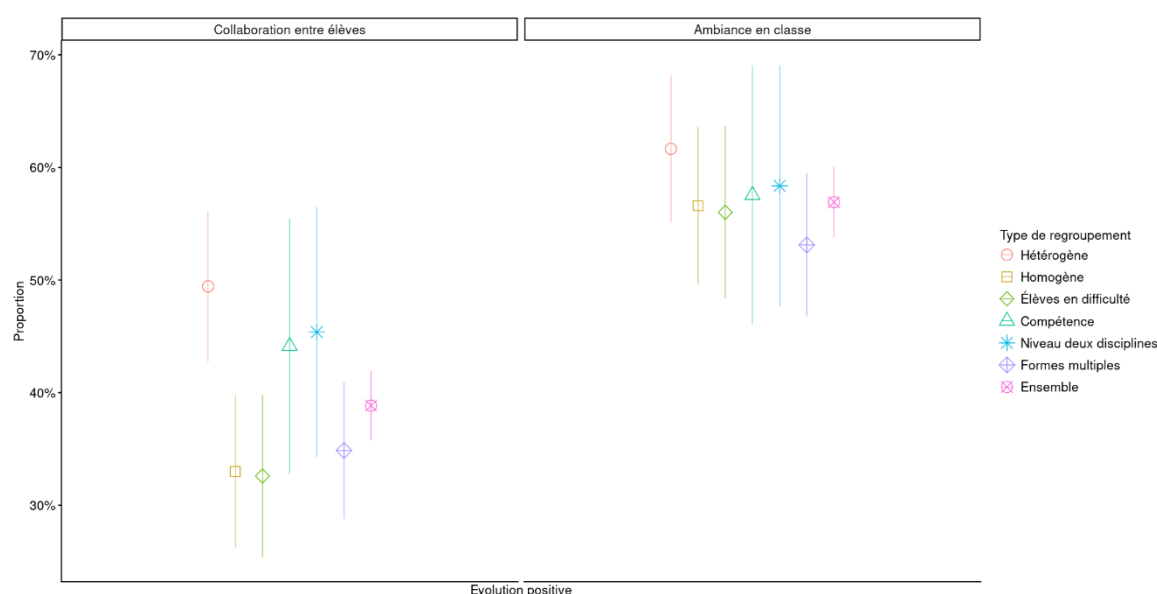
Par rapport aux chefs d'établissement, les coordonnateurs se montrent moins optimistes quant à l'absence d'effets négatifs, en particulier pour les élèves les plus en difficulté. D'après leurs réponses aux questionnaires, les coordonnateurs sont trois fois plus nombreux que les chefs d'établissement à estimer que les regroupements d'élèves ont pu avoir un effet négatif sur les élèves rencontrant le plus de difficulté. Plus précisément, sans tenir compte des formes de regroupement mises en œuvre¹⁷, 15 % des coordonnateurs estiment que les regroupements d'élèves ont eu un impact négatif sur les acquis des élèves les plus en difficulté, 12 % estiment qu'il y a eu un impact négatif sur leur perception de leur capacité à réussir et 18 % qu'il y a eu un impact négatif sur leur investissement scolaire.

¹⁷ Les réponses des coordonnateurs n'ont pas été déclinées par type de regroupement comme l'ont été celles des chefs d'établissement. Il s'agit donc ici de l'ensemble des réponses des coordonnateurs qui déclarent la présence d'au moins une forme de regroupement pour au moins un niveau dans leur discipline.

Les chefs d'établissement perçoivent que la collaboration entre élèves est favorisée par la mise en place de groupes hétérogènes

À la différence des autres dimensions relatives aux effets perçus sur les élèves, la question de la collaboration entre élèves fait apparaître une hiérarchie inverse des réponses : les chefs des établissements qui ont instauré des regroupements hétérogènes sont ceux qui identifient le plus des effets positifs sur la collaboration entre élèves (49 %) (Figure 9). D'après les commentaires des chefs d'établissement et des coordonnateurs, les effectifs réduits ont facilité la collaboration entre élèves, celle-ci pouvant être d'autant plus riche que les élèves au sein d'un même groupe ont des niveaux différents. Dans les établissements avec des groupes homogènes ou des groupes ciblant les élèves les plus fragiles, la part de chefs d'établissement identifiant des effets positifs est significativement plus faible (moins de 35 %).

FIGURE 9 • Impact perçu de la mesure sur la collaboration entre élèves et l'ambiance de travail du point de vue des chefs d'établissement, selon le type de regroupement mis en œuvre



Lecture : 49 % des chefs d'établissement des collèges ayant mis en place des groupes hétérogènes estiment que la mesure a eu un impact positif sur la collaboration entre élèves. Pour l'ensemble des établissements, cette proportion est de 39 % en moyenne.

Champ : Chefs d'établissement de l'échantillon EEC s'étant prononcés sur l'impact perçu de la mesure. Les questions relatives aux impacts perçus n'étant posées que pour les établissements déclarant avoir mis en place des regroupements d'élèves, seuls ces établissements sont conservés pour ce graphique. Les établissements ayant mis en place uniquement des groupes à destination des élèves les plus performants ne sont pas assez nombreux pour être étudiés statistiquement et ne sont pas affichés. Ils sont néanmoins pris en compte pour le calcul de la catégorie « Ensemble ». Taille de l'échantillon retenu : 1 225 établissements pour la question sur la collaboration (4 883 avec pondération) et 1 226 pour la question sur l'ambiance en classe (4 890 avec pondération).

Source : DEPP, enquête EEC, printemps 2025, données finales repondérées.

Pour une majorité de chefs d'établissement, les regroupements d'élèves ont contribué à améliorer l'ambiance en classe : 57 % d'entre eux en font le constat, et cette proportion varie peu selon la forme de regroupement mise en place. Cependant, dans leurs commentaires, les chefs d'établissement et les coordonnateurs soulignent les effets ambigus que peuvent avoir les regroupements d'élèves, quelle que soit leur forme. Plusieurs d'entre eux affirment que ces nouvelles organisations pédagogiques, à travers la diminution du nombre d'élèves dans les groupes, ont permis d'établir une ambiance de classe plus « sereine » ou « apaisée », et plus propice aux enseignements. Mais, dans une moindre mesure, plusieurs coordonnateurs signalent une ambiance de classe « cassée » par la mise en place des groupes, par rapport à une situation en classe de référence, et certains coordonnateurs mettent en avant la dégradation du climat de classe dans les groupes d'élèves les plus en difficulté.

Des effets perçus comme défavorables sur la vie de l'établissement, sans différences marquées selon les dimensions considérées

Dans la plupart des collèges ayant mis en œuvre des regroupements d'élèves, des substitutions d'heures entre différents dispositifs ont eu lieu pour permettre la mise en place des nouvelles organisations pédagogiques en français et en mathématiques pour les classes de 6^e et de 5^e. En particulier, près des deux tiers des chefs d'établissement indiquent avoir diminué le nombre d'heures dédiées à l'accompagnement personnalisé. Les proportions de chefs d'établissement qui déclarent des substitutions d'heures varient peu selon la forme de regroupement mise en œuvre. Il apparaît toutefois que les établissements avec des groupes hétérogènes sont légèrement plus nombreux à signaler la suppression des demi-groupes en sciences ou technologie et la suppression d'options. Cela est possiblement lié au choix de réduire les effectifs pour l'ensemble des élèves, qui implique de consacrer davantage de moyens sur les seuls enseignements de français et de mathématiques en 6^e et en 5^e. Les établissements qui ont mis en place des groupes à destination des élèves les plus en difficulté déclarent pour leur part davantage de diminution des heures dédiées à l'accompagnement personnalisé par rapport aux autres établissements.

De même, pour les impacts négatifs sur la vie de l'établissement comme l'organisation des conseils de classe ou le rôle de professeur principal, les déclarations des chefs d'établissement varient peu selon les formes de regroupements mises en œuvre.

Une exception notable concerne les établissements qui ont mis en œuvre des groupes hétérogènes : ils sont significativement plus nombreux que les autres à déclarer des effets négatifs sur les sorties ou les voyages scolaires ainsi que sur la réalisation de projets interdisciplinaires. Inversement, les établissements avec des groupes homogènes sont significativement moins nombreux à identifier un impact négatif sur les sorties scolaires, mais ils sont plus nombreux à en faire état pour l'organisation des conseils de classe. Enfin, les établissements avec des groupes composés selon le niveau dans les deux disciplines rapportent significativement moins d'effets négatifs à la fois pour la réalisation de projets interdisciplinaires et pour les relations entre les parents et les enseignants.

La question des sorties scolaires revient beaucoup plus fréquemment dans les commentaires des coordonnateurs que dans ceux des chefs d'établissement, quel que soit le type de regroupement. Les coordonnateurs mentionnent ainsi la difficulté d'organiser des sorties, essentiellement réalisées à l'échelle de la classe de référence, car cela « *empiète* » sur plusieurs groupes et donc « *affecte* » davantage de collègues enseignants. De nombreux coordonnateurs insistent d'ailleurs davantage sur l'effet négatif des sorties scolaires sur la progression des élèves que sur l'effet négatif des regroupements sur les sorties : lorsque seuls quelques élèves du groupe sont absents car de sortie avec leur classe, ils peuvent prendre du retard et avoir du mal à le rattraper auprès des autres élèves du groupe qui ne font pas partie de leur classe de référence et qu'ils connaissent donc moins. Pour pallier ce problème, des sorties avec l'ensemble des classes pourraient être mises en place, mais certains coordonnateurs font état de difficultés matérielles pour les organiser. D'autres expriment la démotivation des enseignants et le moindre investissement dans des projets qui concernent une grande partie d'élèves qu'ils ne connaissent finalement pas.

➤ Les effets sur les apprentissages des élèves

Analyses corrélationnelles : évolution des résultats aux évaluations nationales

Avant de fournir de premiers résultats sur l'impact des organisations pédagogiques sur les compétences en français et en mathématiques, cette section débute par la présentation de résultats d'analyses corrélationnelles. Celles-ci comparent l'évolution des résultats aux évaluations nationales de rentrée de 4^e en 2025, en français et en mathématiques selon les organisations pédagogiques mises en place dans le cadre de la réforme. La rentrée de septembre 2025 correspond à la première vague d'évaluations nationales de 4^e passées par des élèves concernés par la réforme : il s'agit de collégiens qui étaient en classe de 5^e lors de sa mise en œuvre à la rentrée 2024.

Dans la suite du document de travail, et sauf mention contraire, les analyses sont réalisées à partir des données non pondérées et non redressées de l'échantillon des établissements répondants à l'enquête administrée par la DEPP au printemps 2025 et ayant mis en place un seul type de regroupement des élèves de 5^e en mathématiques et en français¹⁸. Ainsi, contrairement à la description de la politique mise en œuvre et des modalités organisationnelles retenues par les établissements, l'analyse causale de l'effet de cette politique sur les résultats des élèves ne cherche pas à décrire une population et n'impose pas l'utilisation des poids de sondage pour être pertinente. Cela s'explique par la recherche dans l'échantillon d'une relation entre une cause et son effet, à travers l'exposition à un traitement, indépendamment du nombre d'établissements représentés par une observation. En présence d'effets hétérogènes, l'analyse présentée dans ce document de travail n'est donc pas forcément représentative de l'éventuel effet agrégé sur la population cible. Dans le cadre de cette étude, cependant, en raison des taux de réponse élevés et de l'absence de facteur observable fort influençant la non-réponse, l'introduction de la pondération dans les analyses causales ne pourrait que marginalement modifier les estimations : les scores moyens de 4^e et leur évolution entre 2023 et 2025 sont semblables pour la base de sondage, pour l'échantillon pondéré et pour l'échantillon non-pondéré (**Encadré 4**).

Les types de regroupements mis en œuvre sont construits à partir des réponses des chefs d'établissement à l'enquête, et sont rassemblés en trois catégories :

1. Groupes hétérogènes : les établissements qui mettent en œuvre des groupes composés d'élèves présentant des niveaux scolaires ou des besoins différents ;
2. Groupes selon le profil scolaire : les établissements pour lesquels les groupes sont constitués selon le profil scolaire des élèves. Cette catégorie inclut les types de regroupement désignés jusqu'ici comme groupes homogènes, groupes par compétence, groupes de même niveau scolaire en français et mathématiques, groupes constitués d'élèves rencontrant le plus de difficulté et groupes constitués des élèves les plus performants ;
3. Pas de groupe : les établissements qui ne mettent en œuvre aucun regroupement d'élèves au cours de l'année 2024-2025.

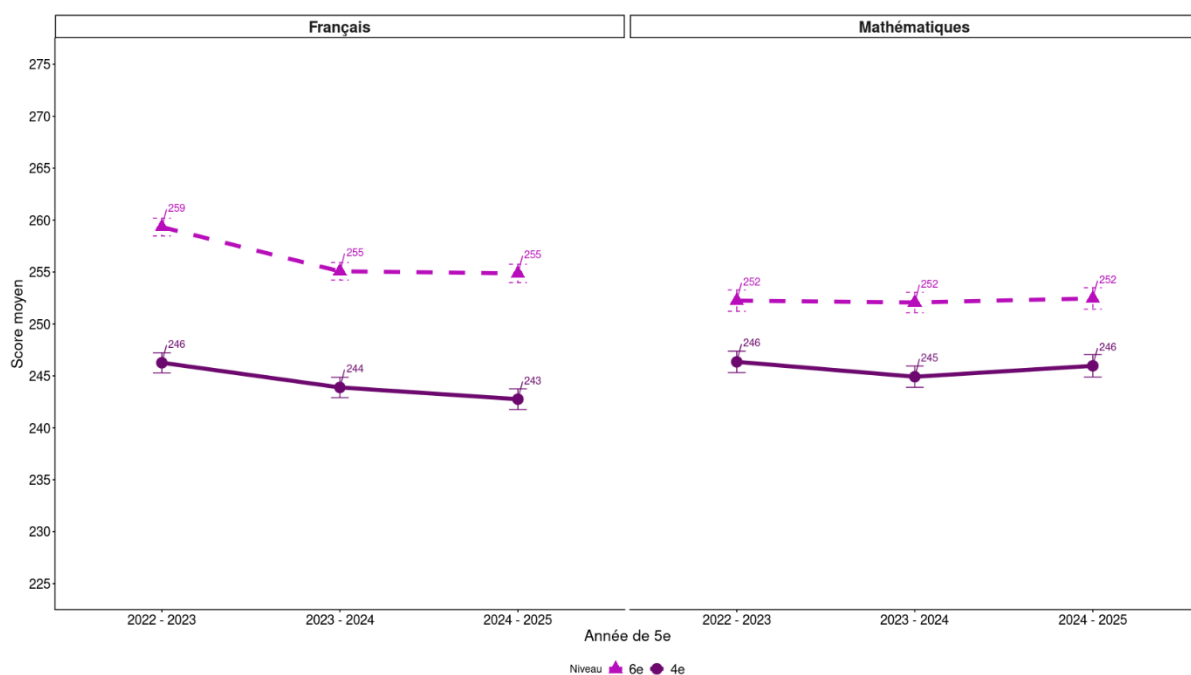
Des résultats moyens globalement stables entre 2024 et 2025 dans les deux disciplines pour l'ensemble des établissements

On étudie ici le score moyen aux évaluations nationales de début de 4^e des élèves ayant effectué leur année de 5^e au sein des collèges de l'échantillon au cours de différentes années. En comparant les cohortes d'élèves ayant effectué leur année de 5^e en 2022-2023 et 2023-2024 (les deux dernières années avant la réforme) et en 2024-2025 (première année de mise en œuvre de la réforme) dans ces mêmes établissements, on constate que le score moyen en français a très légèrement diminué au cours du temps : baisse de deux points entre les cohortes 2022-2023 et 2023-2024, et d'un point

¹⁸ Les établissements ayant mis en place des groupes mixtes (autour de 2 % en 5^e dans chaque discipline) ne sont pas retenus dans l'analyse et ne sont donc pas pris en compte pour le calcul des scores moyens pour l'ensemble des élèves, qui apparaissent également sur les graphiques.

entre les cohortes 2023-2024 et 2024-2025, passant ainsi de 246 à 244 puis 243 (**Figure 10**)¹⁹. De manière intéressante, cette évolution semble assez similaire à celle observée pour les scores de français des élèves des mêmes cohortes, lorsqu'ils étaient en 6^e. D'un point de vue macroscopique, la réforme ne semble donc pas avoir entraîné de changement net entre le début de 6^e et le début de la 4^e. Qui plus est, les scores aux évaluations nationales présentés ici sont des scores standardisés avec un écart-type de 50 : une évolution d'un point représente donc une évolution mineure, de l'ordre de 2 % d'un écart-type. De même, l'évolution observée en mathématiques en début de 4^e pour les trois mêmes cohortes d'élèves apparaît très proche de celle observée pour ces cohortes lorsqu'elles étaient en 6^e : une baisse d'un point entre les cohortes 2022-2023 et 2023-2024, suivie d'une hausse d'un point entre les cohortes 2023-2024 et 2024-2025. Globalement, à ce niveau d'analyse très agrégé, la réforme mise en œuvre en 2024-2025 pour les élèves de 5^e ne semble donc pas avoir eu d'effet notable, dans un sens ou dans l'autre, à l'aune des résultats des élèves en français et en mathématiques en début de 4^e. Ces écarts de scores moyens entre les années sont cohérents avec ceux observés pour l'ensemble des établissements de France, en particulier au sein de la population de référence (**Encadré 4**). Cela suggère que l'échantillon conserve une bonne représentativité de la population cible, grâce au plan de sondage initial, à un taux de réponse à l'enquête statistique auprès des chefs d'établissement de l'ordre de 80 %, et au comportement de non-réponse peu différencié selon les caractéristiques observables (voir **Annexe 2**).

FIGURE 10 • Évolution des scores moyens en français et en mathématiques aux évaluations nationales de 6^e et de 4^e selon l'année de scolarisation en 5^e, pour les élèves scolarisés en 5^e au sein des établissements de l'échantillon



Lecture : Le score moyen aux évaluations nationales de français de début de 4^e des élèves scolarisés en 5^e au cours de l'année scolaire 2022-2023 dans les collèges de l'échantillon qui adoptent une forme de regroupement unique en 5^e au cours de l'année 2024-2025 est de 246. Le score moyen de français de ces mêmes élèves à l'entrée en 6^e est de 259.

Note : Les intervalles de confiance à 95 % sont affichés sur le graphique.

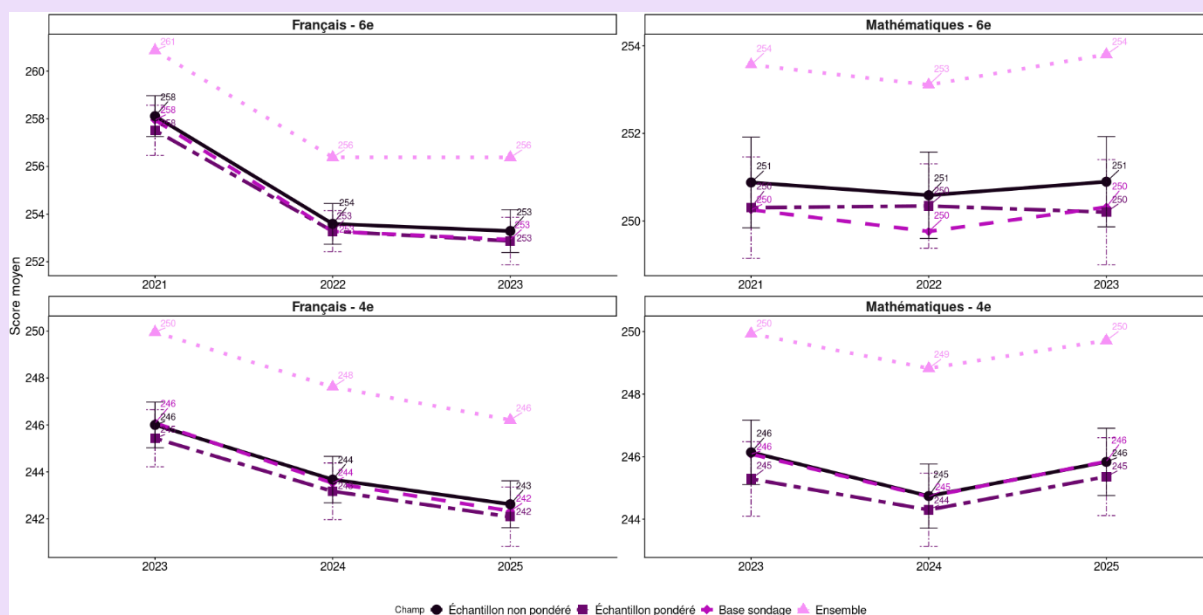
Champ : Établissements de l'enquête EEC ayant adopté une forme unique de regroupement ou aucune en 5^e pour la discipline considérée. Cela correspond à 1 289 collèges pour le français et à 1 291 collèges pour les mathématiques. Il n'y a pas de restriction ici sur la nécessité pour les élèves d'être retrouvés à la fois dans les évaluations nationales de début de 6^e et dans celles de début de 4^e.

Sources : DEPP, Évaluations nationales en français et mathématiques de 6^e (2021, 2022, 2023) et de 4^e (2023, 2024, 2025) ; enquête EEC, questionnaires Chefs d'établissement, printemps 2025 ; Constat de rentrée 2025.

¹⁹ La population d'intérêt est ici constituée des élèves qui ont effectué leur classe de 5^e au cours de l'année 2024-2025, et qui ont donc pu être exposés aux nouvelles organisations pédagogiques pendant un an avant de passer leurs évaluations nationales de 4^e en septembre 2025. L'échantillon de travail exclut donc les élèves qui ont redoublé leur année de 4^e. Ce choix est appliqué à l'ensemble des années, afin de garantir la cohérence des résultats. En revanche, les élèves qui ont redoublé leur classe de 5^e au maximum une fois sont pris en compte dans l'échantillon de travail : l'année de l'évaluation de 6^e n'est donc pas toujours de deux ans antérieure à l'année de l'évaluation de 4^e. C'est pourquoi l'analyse raisonne par rapport à la dernière année scolaire de 5^e précédant le passage en 4^e.

ENCADRE 4 Représentativité de l'échantillon au regard des scores aux évaluations nationales

FIGURE 11 • Évolution des scores moyens de français et de mathématiques aux évaluations nationales de 6^e entre 2021 et 2023 et aux évaluations nationales de 4^e entre 2023 et 2025



Lecture : À la rentrée de septembre 2023, le score moyen aux évaluations nationales de français de 4^e de l'ensemble des établissements est égal à 250. Ce score moyen est de 246 dans la base de sondage, qui correspond aux collèges publics de France hors Mayotte, ouverts à la rentrée 2024 et comptant au moins 20 élèves en 6^e et 20 élèves en 5^e. Dans les 1 315 collèges de l'échantillon retenus pour l'analyse, le score moyen est égal à 246 sans pondération et à 245 avec la prise en compte des poids d'échantillonnage des établissements, corrigés pour tenir compte de la non-réponse.

Note : Pour les collèges de l'échantillon, les scores pris en compte dans ces calculs sont ceux des élèves qui passent les évaluations nationales en début de 4^e, et non les élèves qui étaient scolarisés en 5^e au sein de ces établissements. Pour les données de l'échantillon, les intervalles de confiance à 95 % sont affichés sur le graphique.

Champ : Établissements du ministère de l'Éducation nationale, France + COM (hors Nouvelle-Calédonie et Wallis-et-Futuna), public et privé sous contrat.

Source : DEPP, évaluations exhaustives de début de sixième et de début de quatrième ; enquête EEC, questionnaires Chefs d'établissement, printemps 2025, données finales repondérées.

Que la pondération soit prise en compte ou non, les scores moyens en français et en mathématiques aux évaluations nationales de 6^e et de 4^e obtenus à partir de l'échantillon sont similaires à ceux des collèges de la population de référence, c'est-à-dire des collèges publics de France hors Mayotte ouverts à la rentrée 2024 et comptant au moins 20 élèves en classe de 6^e et 20 élèves en classe de 5^e (Figure 11). Il s'agit ici des scores moyens aux évaluations nationales passées en 6^e et en 4^e dans les établissements de la population de référence ou de l'échantillon, sans tenir compte de l'établissement de scolarisation en 5^e. Les différences de scores entre les années sont ainsi les mêmes dans l'échantillon et dans la population de référence. Lorsque des différences sont visibles, comme c'est notamment le cas pour les scores de mathématiques en 6^e, elles sont très faibles et non statistiquement significatives.

Les scores moyens obtenus à partir des collèges de l'échantillon ou de la population de référence sont inférieurs à ceux de l'ensemble des collèges. Cela est lié au choix d'exclure les collèges privés de la population de référence de cette étude.

D'un point de vue strictement corrélationnel, la mise en place des nouvelles organisations pédagogiques en 5^e au cours de l'année scolaire 2024-2025 s'accompagne ainsi d'une stabilité des scores moyens en français et en mathématiques (Figure 10). À partir de ces seules observations, il n'est cependant pas possible d'identifier un effet causal de la réforme. En effet, si cette dernière n'avait pas été mise en place, les scores en français et en mathématiques auraient également pu s'améliorer ou diminuer. Les évolutions observées ne peuvent donc pas être attribuées directement à la réforme sans disposer d'un point de comparaison permettant d'estimer ce qui se serait produit en son absence.

Les écarts observés entre ces mêmes cohortes à leur entrée en 6^e, avant toute exposition aux nouvelles organisations pédagogiques mises en place au collège, invitent à interpréter avec prudence le rôle de ces dernières dans les évolutions observées des scores de 4^e. En effet, l'atténuation de la dynamique de diminution des scores moyens en français entre les cohortes ayant effectué leur 5^e en 2023-2024 et 2024-2025 par rapport à la cohorte précédente s'observe déjà dans les résultats de ces mêmes élèves à leur entrée en 6^e. Si l'amplitude est différente, le profil des évolutions est semblable, ce qui souligne des différences entre cohortes préexistantes à la réforme. En mathématiques, en revanche, les évolutions des scores des élèves à leur entrée en 6^e sont moins marquées.

Outre la dimension temporelle, **l'analyse des progressions des élèves peut être menée en comparant les établissements selon qu'ils ont ou non mis en place de nouvelles organisations pédagogiques en 2024-2025** et, le cas échéant, selon la ou les formes de regroupement retenues. Le choix de la forme de regroupement mise en place est partiellement lié aux caractéristiques des établissements : en particulier, ceux qui ont opté pour des regroupements hétérogènes se caractérisent par des niveaux d'IPS plus élevés, tandis que plusieurs formes de regroupements selon le profil scolaire des élèves ont été davantage adoptées par les établissements présentant des niveaux d'IPS plus faibles.

Une évolution globale des scores aux évaluations nationales de 4^e qui ne semble pas dépendre du type de regroupement mis en place

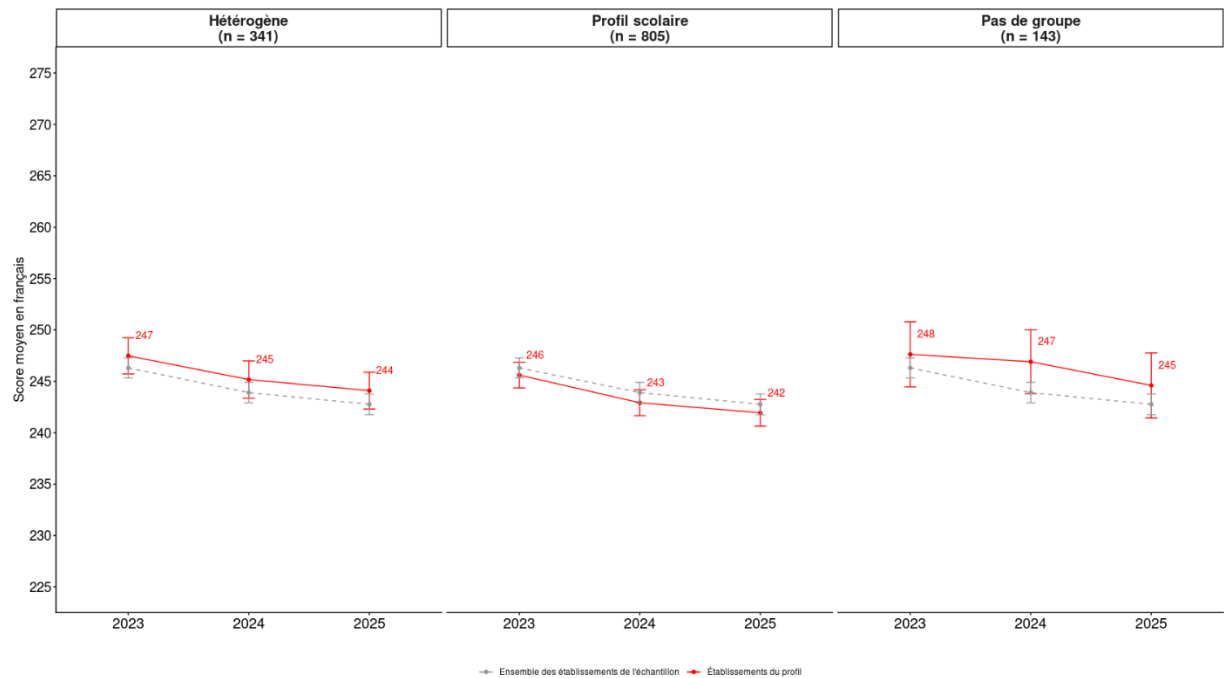
Les **Figures 12 et 13** présentent les résultats moyens des élèves aux évaluations nationales de début de 4^e selon le type d'établissement où ils ont été scolarisés au cours de leur année de 5^e. Les nouvelles organisations pédagogiques ont été mises en place à partir de la rentrée de septembre 2024 : les premiers élèves à y avoir été exposés sont donc ceux qui ont passé les évaluations nationales de 4^e en 2025. Sur les **Figures 12 et 13**, les points correspondant aux années 2023 et 2024 renvoient à des observations antérieures à la réforme (les établissements sont classés en fonction des regroupements mis en place en 2024-2025) : l'objectif est ici de comparer la réussite aux évaluations de 4^e des élèves scolarisés en 5^e en 2024-2025 dans les établissements de l'échantillon par rapport à celles des élèves scolarisés en 5^e dans ces mêmes établissements au cours des deux années précédant la mise en œuvre des regroupements.

Entre septembre 2024 et 2025, les évolutions des scores moyens aux évaluations nationales de 4^e ne diffèrent pas sensiblement selon les types de regroupement auxquels les élèves ont été exposés. En français, une diminution d'un à deux points s'observe pour toutes les organisations pédagogiques mises en place (**Figure 12**). Pour les trois profils d'établissements, ces évolutions dans le temps sont significatives au moins au seuil de 5 %.

Ces graphiques permettent également de comparer les niveaux moyens de réussite entre établissements avant la mise en place des nouvelles organisations pédagogiques. On constate ainsi que les trois groupes d'établissements ne sont pas directement comparables en 2025, dans la mesure où de légères différences de niveau moyen préexistaient à la réforme.

En mathématiques, l'évolution des scores moyens entre les rentrées 2023 et 2024 semble suivre des tendances similaires quelles que soient les organisations pédagogiques mises en place dans les établissements, avec une baisse d'un à deux points entre septembre 2023 et 2024 puis une hausse d'ampleur similaire entre 2024 et 2025 (**Figure 13**).

FIGURE 12 • Évolution des scores moyens en français aux évaluations nationales de 4^e entre 2023 et 2025



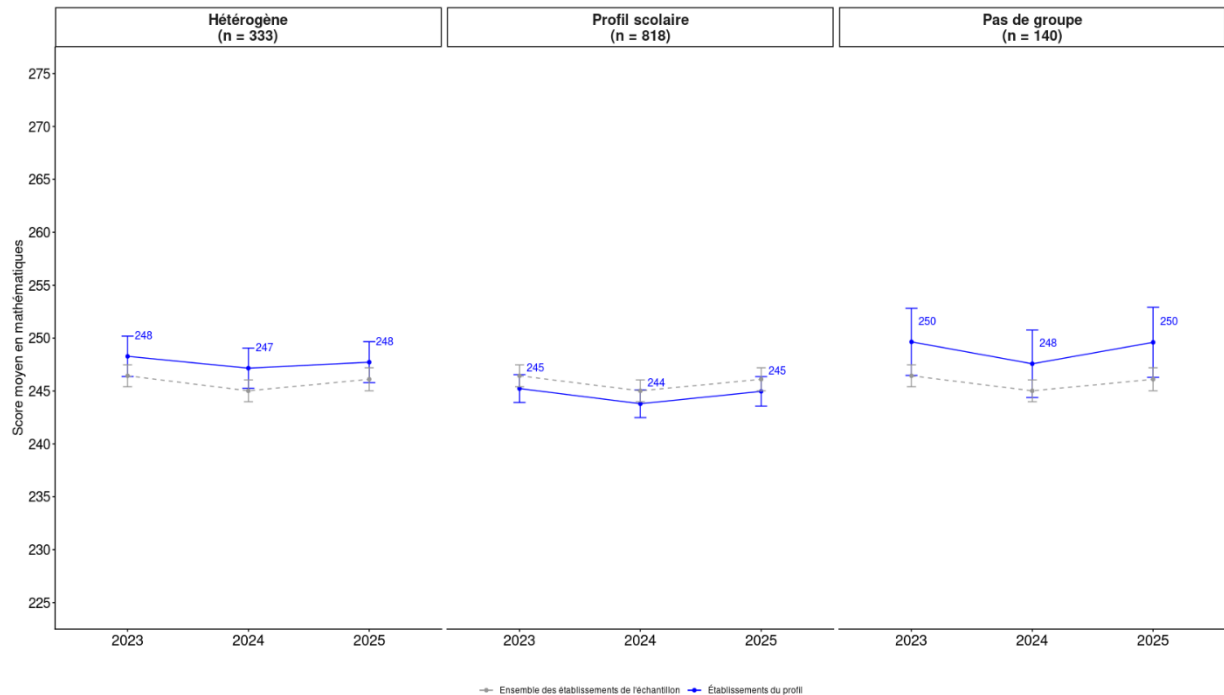
Lecture : Dans les collèges ayant mis en place des groupes hétérogènes, le score moyen aux évaluations de début de 4^e en français était de 244 points en 2025. Le score moyen pour l'ensemble des établissements de l'échantillon qui ont adopté une modalité de regroupement unique en 5^e pour le français était égal à 243 en 2025 (en pointillé).

Note : Les intervalles de confiance à 95 % sont affichés sur le graphique.

Champ : Élèves scolarisés en 5^e au sein des établissements de l'enquête EEC ayant adopté une forme unique de regroupement ou aucune en 5^e pour le français. Cela correspond à 1 289 collèges.

Sources : DEPP, Évaluations nationales en français de 4^e (2023, 2024, 2025) ; enquête EEC, questionnaires Chefs d'établissement, printemps 2025 ; Constat de rentrée 2025.

FIGURE 13 • Évolution des scores moyens en mathématiques aux évaluations nationales de 4^e entre 2023 et 2025



Lecture : Dans les collèges ayant mis en place des groupes hétérogènes, le score moyen aux évaluations de début de 4^e en mathématiques était de 248 points en 2025. Le score moyen pour l'ensemble des établissements de l'échantillon qui ont adopté une modalité de regroupement unique en 5^e pour les mathématiques était égal à 246 en 2025 (en pointillé).

Note : Les intervalles de confiance à 95 % sont affichés sur le graphique.

Champ : Élèves scolarisés en 5^e au sein des établissements de l'enquête EEC ayant adopté une forme unique de regroupement ou aucune en 5^e pour les mathématiques. Cela correspond à 1 291 collèges.

Sources : DEPP, Évaluations nationales en mathématiques de 4^e (2023, 2024, 2025) ; enquête EEC, questionnaires Chefs d'établissement, printemps 2025 ; Constat de rentrée 2025.

Enfin, en mathématiques, pour les trois années observées (septembre 2023, 2024 et 2025), les scores moyens des élèves scolarisés dans les établissements ayant choisi des regroupements selon le profil scolaire des élèves au cours de l'année 2024-2025 sont plus faibles que ceux scolarisés dans les établissements ayant opté pour des regroupements hétérogènes ou pour l'absence de regroupement, et cette différence est toujours significative au moins au seuil de 5 %. En revanche, l'écart entre les élèves scolarisés dans des établissements ayant mis en place des groupes hétérogènes et ceux scolarisés dans des établissements sans regroupements n'est jamais statistiquement significatif. Globalement, l'évolution des scores moyens en français et mathématiques suit des tendances très similaires au cours du temps quelles que soient les organisations pédagogiques mises en place.

L'évolution de la répartition des élèves dans les groupes de performance aux évaluations nationales est cohérente avec l'évolution des scores observée selon la nature des organisations pédagogiques mises en œuvre en 2024-2025 (**Annexe 4**).

Conclusion relative aux premières analyses corrélationnelles

Globalement, l'analyse corrélationnelle des résultats aux évaluations nationales de 4^e ne laisse pas entrevoir de rupture majeure entre les années qui précèdent la réforme et l'année 2025, qui voit arriver en 4^e les élèves qui ont pu être exposés aux nouvelles organisations pédagogiques au cours de leur année de 5^e. Les élèves scolarisés dans les établissements qui constituent des groupes selon le profil scolaire ont en moyenne des scores inférieurs à ceux des autres établissements, mais **les évolutions dans le temps sont similaires et ces écarts de niveaux, déjà présents avant la réforme, semblent rester globalement stables**. Seuls les élèves scolarisés au sein d'établissements sans regroupements d'élèves semblent connaître une évolution légèrement différente entre 2023 et 2025, notamment une progression en français.

Étude corrélationnelle en fonction du niveau initial de performance des élèves

En conservant une approche corrélationnelle, il est possible d'observer la progression des élèves entre la 6^e et la 4^e selon les établissements qu'ils ont fréquentés en 5^e et la nature des organisations pédagogiques mises en œuvre. Cette approche reste toutefois descriptive : elle ne permet pas de distinguer l'impact de la réforme de tous les autres facteurs susceptibles d'influencer l'évolution de la réussite des élèves. En particulier, les établissements qui optent pour une certaine forme de regroupement plutôt qu'une autre peuvent présenter des caractéristiques propres influençant les résultats des élèves, indépendamment de la réforme. De même, certaines formes de regroupement peuvent être mises en place précisément pour répondre aux besoins spécifiques des élèves, dont les parcours scolaires ont des chances d'être différents de celui des élèves pour lesquels de tels regroupements ne sont pas jugés nécessaires, y compris en l'absence de réforme.

Les évaluations nationales réalisées par la DEPP permettent la comparaison rigoureuse des scores moyens de toute une cohorte d'élèves d'une année à l'autre pour un même niveau. En revanche, ces scores n'ont pas été construits pour analyser la progression d'un même élève d'un niveau à un autre. Plutôt que de mesurer la progression individuelle des élèves entre la 6^e et la 4^e, l'analyse mobilise ici les groupes de performance, qui sont construits à partir des scores obtenus pour chaque niveau (voir **Encadré 5** de l'**Annexe 4**).

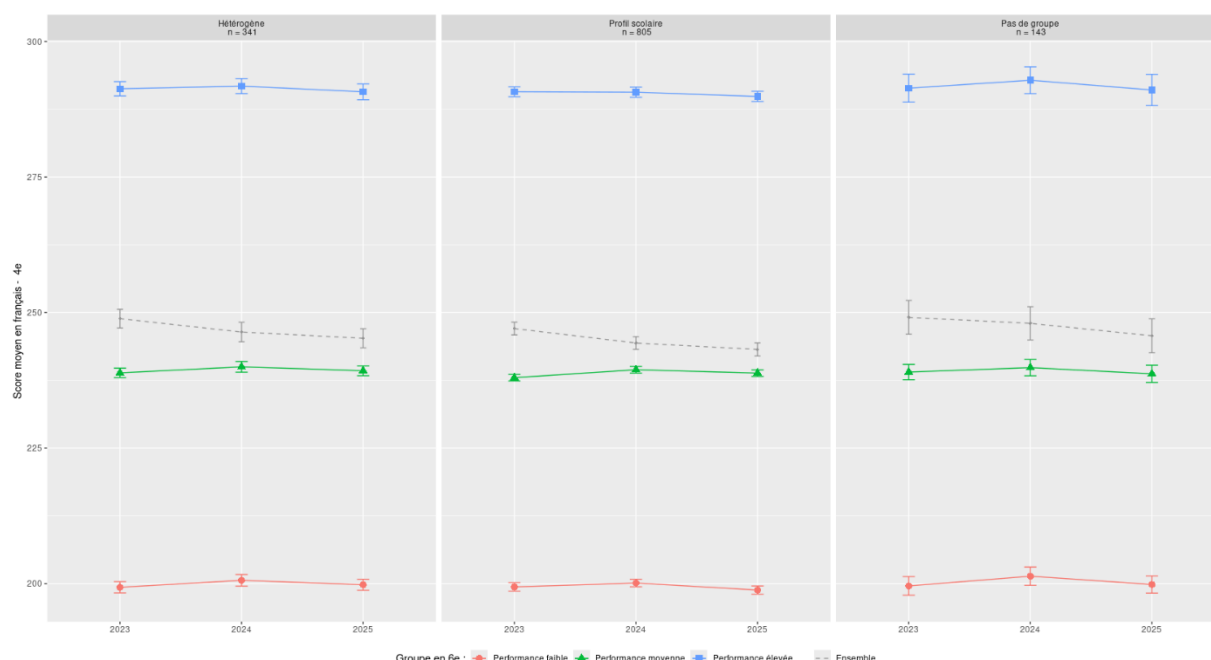
Évolution des scores moyens en 4^e selon le niveau initial de performance en 6^e et le type de regroupement

Dans un premier temps, il est possible de comparer les scores moyens des élèves aux évaluations de 4^e selon le type d'organisation pédagogique auquel ils ont été exposés au cours de leur année de 5^e, tout en tenant compte de leur niveau initial dans la discipline aux évaluations standardisées de 6^e (**Figures 14 & 15**). Ces comparaisons, de nature corrélationnelle, permettent de documenter

d'éventuelles différences de progression jusqu'en 4^e entre élèves de niveau initial comparable, selon qu'ils ont été scolarisés dans des établissements ayant mis en œuvre différentes organisations pédagogiques

En français, les scores varient peu selon le niveau initial et le type de regroupement auquel les élèves ont été exposés (**Figure 14**). Les élèves qui appartenaient aux groupes de performance faible en 6^e ont des scores inférieurs en septembre 2025 aux élèves de la même catégorie l'année précédente, quelles que soient les organisations pédagogiques auxquelles ils ont été exposés au cours de leur année de 5^e. Cependant, cette modeste diminution n'est significative au seuil de 5 % que pour les élèves scolarisés dans les établissements ayant mis en place des groupes selon le profil scolaire des élèves. Les scores moyens aux évaluations nationales de français de début de 4^e des élèves dont le niveau initial de performance en 6^e était moyen ou élevé diminuent également légèrement entre septembre 2024 et 2025, mais cette diminution n'est jamais significative au seuil de 5 %.

FIGURE 14 • Évolution des scores moyens en français aux évaluations nationales de 4^e entre 2023 et 2025 selon le niveau initial (groupe de performance en 6^e) et le type de regroupement mis en œuvre



Lecture : Les élèves qui appartiennent au groupe de performance élevée en 6^e pour le français, qui effectuent leur 5^e en 2022-2023 au sein des établissements qui mettent en place des regroupements hétérogènes en 5^e en français au cours de l'année 2024-2025 et qui passent leurs évaluations nationales de début de 4^e en 2023 obtiennent un score moyen de 291 en français.

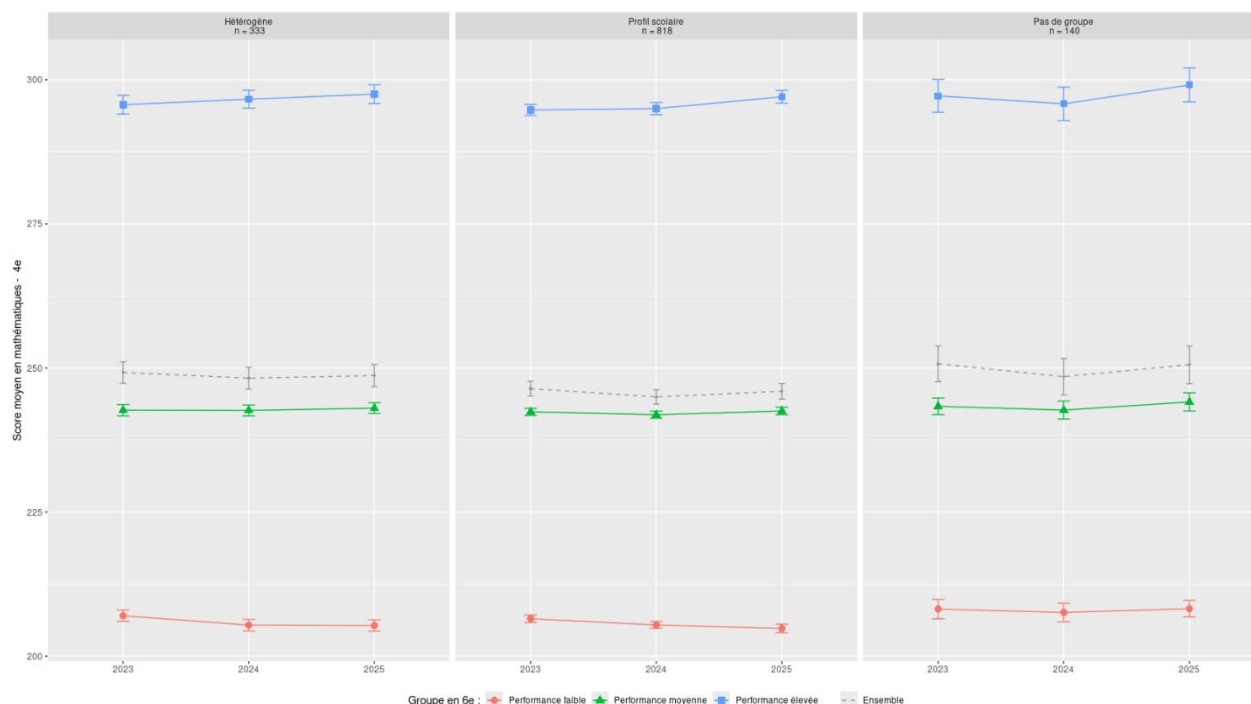
Note : Les scores moyens selon le type de regroupements sont rappelés en pointillés. Les intervalles de confiance à 95 % sont affichés sur le graphique.

Champ : Élèves scolarisés en 5^e au sein des établissements de l'enquête EEC ayant adopté une forme unique de regroupement ou aucune en 5^e pour le français. Cela correspond à 1 289 collèges. Seuls les élèves retrouvés à la fois dans les évaluations nationales de français de début de 6^e et dans celles de début de 4^e ont été pris en compte dans l'analyse.

Sources : DEPP, Évaluations nationales en français de 6^e (2021, 2022, 2023) et de 4^e (2023, 2024, 2025) ; enquête EEC, questionnaires Chefs d'établissement, printemps 2025 ; Constat de rentrée 2025.

En mathématiques, aucune différence statistiquement significative au seuil de 5 % n'est observée entre septembre 2024 et 2025 pour les scores moyens des élèves dont le niveau initial (mesuré par le groupe de performance aux évaluations à l'entrée en 6^e) était faible, quel que soit le type de regroupement mis en place au cours de l'année scolaire 2024-2025 (**Figure 15**). En revanche, les élèves dont le niveau initial de performance était moyen ou élevé en 6^e et qui ont fréquenté des établissements avec des groupes constitués selon le profil scolaire des élèves en 5^e présentent en moyenne de meilleurs scores aux évaluations nationales en 2025 qu'en 2024. Il faut toutefois noter que cette évolution positive s'observe également pour les élèves dont le niveau initial de performance en 6^e était élevé et qui ont effectué leur 5^e au sein d'établissements n'ayant pas mis en place de regroupements. Elle ne semble donc pas être particulièrement liée à la mise en place de la réforme.

FIGURE 15 • Évolution des scores moyens en mathématiques aux évaluations nationales de 4^e entre 2023 et 2025 selon le niveau initial (groupe de performance en 6^e) et le type de regroupement mis en œuvre



Lecture : Les élèves qui appartiennent au groupe de performance élevée en 6^e pour les mathématiques, qui effectuent leur 5^e en 2022-2023 au sein des établissements qui mettent en place des regroupements hétérogènes en 5^e en mathématiques au cours de l'année 2024-2025 et qui passent leurs évaluations nationales de début de 4^e en 2023 obtiennent un score moyen de 296 en mathématiques.

Note : Les scores moyens selon le type de regroupement sont rappelés en pointillés. Les intervalles de confiance à 95 % sont affichés sur le graphique.

Champ : Élèves scolarisés en 5^e au sein des établissements de l'enquête EEC ayant adopté une forme unique de regroupement ou aucune en 5^e pour les mathématiques. Cela correspond à 1 291 collèges. Seuls les élèves retrouvés à la fois dans les évaluations nationales de mathématiques de début de 6^e et dans celles de début de 4^e ont été pris en compte dans l'analyse.

Sources : DEPP, Évaluations nationales en mathématiques de 6^e (2021, 2022, 2023) et de 4^e (2023, 2024, 2025) ; enquête EEC, questionnaires Chefs d'établissement, printemps 2025 ; Constat de rentrée 2025.

À ce stade de l'analyse, les légères variations des résultats aux évaluations nationales observées entre 2024 et 2025 selon les différents types de regroupement ne peuvent pas être interprétées comme des effets causaux des organisations pédagogiques. En effet, le choix de mettre en place une modalité de regroupement plutôt qu'une autre, à l'échelle d'une classe, d'un niveau, d'une discipline ou d'un établissement, a pu dépendre des perspectives de réussite des élèves concernés par ces décisions organisationnelles. Par exemple, les établissements ont pu choisir une forme de regroupement en réponse aux caractéristiques de leurs élèves, qui elles-mêmes influencent leur progression d'un niveau à l'autre ou d'une année à l'autre pour un même niveau, indépendamment de l'organisation pédagogique retenue. Il est donc tout à fait possible que les résultats des élèves des trois types d'établissements aient été différents même en l'absence de la réforme étudiée ici. La question cruciale est donc de déterminer quelle aurait été leur situation en l'absence de réorganisation pédagogique, exercice difficile pour lequel il est nécessaire de mobiliser des méthodologies statistiques rigoureuses comme les différences de différences afin d'isoler et de mesurer les éventuels effets causaux de ces organisations pédagogiques. C'est l'objet des prochaines parties de ce document de travail.

Analyses causales

Méthodologie : Différences de différences

Cadre général

L'approche expérimentale constitue la meilleure solution pour mesurer l'effet causal des regroupements d'élèves en neutralisant l'éventuelle influence d'autres facteurs explicatifs. Cette méthode a notamment été mise en œuvre par Duflo et al. (2011), dans le cadre d'une étude de l'impact des classes de niveau au Kenya. Leur expérimentation consistait à sélectionner aléatoirement les écoles amenées à mettre en place des classes de niveau, mais également d'assigner aléatoirement les enseignants à chacune de ces classes afin de neutraliser l'effet des ressources allouées aux différents regroupements d'élèves. La comparaison des scores en fin d'année permet ainsi de mesurer l'effet d'avoir été assigné à une école qui met en place des classes de niveau plutôt que dans une école sans. Cette mesure de l'effet peut être déclinée selon le niveau initial des élèves. Cependant, il n'est pas toujours possible de modifier en profondeur le fonctionnement des établissements ou des systèmes scolaires pour expérimenter différentes modalités organisationnelles.

Pour se rapprocher du cadre expérimental et par analogie avec les expérimentations avec assignation aléatoire, les groupes exposés à une réforme ou à une politique sont désignés comme « traités » ou bénéficiaires du traitement, tandis que ceux qui ne sont pas concernés sont identifiés comme « groupe de contrôle » ou « groupe témoin ». Une des particularités de la réforme des groupes de besoin réside dans la pluralité du traitement considéré, puisque chaque établissement a été libre de choisir ses propres modalités d'organisation au sein du cadre proposé par l'arrêté du 15 mars 2024. À partir de la typologie des organisations pédagogiques réalisée à l'échelle du niveau et de la discipline, il est possible de définir un groupe témoin, qui n'a pas instauré de nouvelle organisation pédagogique à la rentrée 2024, ainsi que deux branches de traitement : les établissements qui ont regroupé les élèves indépendamment de leur niveau, en créant ainsi des groupes hétérogènes (potentiellement de taille réduite), et les établissements ayant regroupé les élèves selon leur profil scolaire.

Parmi les approches méthodologiques non-expérimentales les plus rigoureuses figurent les « différences de différences » (ou « doubles différences »), qui comparent l'évolution des performances d'un groupe ayant bénéficié de la réforme à celle d'un groupe comparable n'en ayant pas bénéficié. On parle de « double » différence parce que le calcul prend en compte deux différences : d'abord la variation des performances avant et après la mise en œuvre de l'intervention, d'une part pour le groupe traité et d'autre part pour le groupe témoin, puis la différence entre ces deux différences, qui s'interprète, sous certaines hypothèses, comme un effet causal estimé de la politique. Cette méthode a déjà été utilisée pour évaluer l'effet de la mise en place de groupes de niveau dans différents pays (Van de Werfhorst, 2019 ; Terrin & Triventi, 2023).

Dans cette approche, les différences de caractéristiques entre les établissements ne représentent pas une menace pour l'identification des effets de la politique publique si elles sont stables dans le temps et si elles influencent toujours les performances des élèves de la même manière. Cette approche permet ainsi d'estimer un effet causal sous réserve que certaines conditions soient respectées. Premièrement, la mise en œuvre des regroupements doit être le seul changement majeur qui intervient dans la période de temps considérée. Dans le cas contraire, il ne sera pas possible de distinguer l'effet propre de cette réforme des effets des autres changements simultanés. La seconde condition impose que les établissements qui mettent en place des regroupements, et ceux qui ne le font pas auraient suivi des trajectoires similaires au cours du temps, si aucun des deux n'avait introduit cette nouvelle organisation. En d'autres termes, des caractéristiques du collège telles que les performances des élèves ou l'ancienneté des enseignants peuvent être structurellement différentes entre les établissements traités et témoins. Toutefois, il faut faire l'hypothèse que cette

différence aurait été stable au cours du temps en l'absence de réforme. Si cette hypothèse, dite de « tendances parallèles », n'est pas vérifiée, il n'est pas possible d'interpréter les écarts qui suivent la mise en place des organisations en groupes comme les effets causaux de cette réforme.

Par définition, il est impossible de tester directement la validité de l'hypothèse des tendances parallèles, car on ne peut pas observer simultanément la situation d'un établissement traité après la mise en place d'une réforme et la situation de ce même établissement en l'absence de réforme. Cette dernière situation, inobservée, est appelée la situation « contrefactuelle ». La validité de l'hypothèse doit donc être défendue par des arguments reposant sur la connaissance du contexte ou sur certains éléments empiriques comme le parallélisme des trajectoires avant la mise en place de la réforme.

Les élèves qui effectuent leur 5^e au sein des établissements des deux branches de traitement décrites précédemment (regroupements hétérogènes donc indépendants du niveau des élèves et regroupements selon le profil scolaire) sont comparés aux élèves des établissements n'ayant pas mis en place de regroupements (groupe témoin). Dans un premier temps, il faut déterminer si l'hypothèse de tendances parallèles se vérifie avant la mise en œuvre des regroupements et, le cas échéant, cette comparaison permet de mesurer l'effet de la mise en place de chaque type de regroupement au cours de l'année 2024-2025 pour les enseignements de français et de mathématiques en 6^e et en 5^e. Dans un second temps, les élèves des établissements ayant mis en place des regroupements selon le profil des élèves sont comparés aux élèves des établissements ayant opté pour une organisation en groupes hétérogènes.

Deux approches complémentaires ont été adoptées pour mesurer l'effet de ces nouvelles organisations pédagogiques sur la réussite scolaire des élèves :

- la première consiste à comparer les résultats des élèves aux évaluations de 4^e à l'échelle des établissements, entre les dernières cohortes d'élèves ayant précédé la réforme (2023 et 2024) et la première cohorte concernée par les nouvelles organisations pédagogiques (2025) ;
- la seconde approche consiste à comparer la progression année après année des élèves d'une même cohorte, selon le type de regroupements mis en place dans leur collège en classe de 5^e en 2024-2025, à partir des évaluations de début de CP, mi-CP, CE1, 6^e, et 4^e.

Ces deux méthodes ont chacune leurs avantages et leurs inconvénients. L'avantage de la première est qu'elle permet d'analyser des effets sur l'ensemble des élèves, alors que la seconde impose de se limiter aux élèves que l'on arrive à suivre dans les données administratives entre le début de CP et la 4^e, ce qui n'est pas le cas pour une part importante d'entre eux. Inversement, la seconde méthode présente l'avantage d'offrir un recul temporel plus important, permettant d'observer les évolutions des différents groupes sur une période plus longue avant la mise en œuvre des regroupements : quatre périodes avant pour la seconde méthode (début de CP, mi-CP, CE1 et 6^e), contre seulement deux périodes pour la première (2023 et 2024). Par ailleurs, la seconde approche présente l'avantage de comparer des progressions scolaires d'élèves provenant des mêmes cohortes, et à ce titre affectés de la même manière par les chocs de différentes natures pouvant influencer sur la réussite scolaire (crise sanitaire, réduction de la taille des classes en éducation prioritaire, etc.).

Première approche : comparaison des résultats des élèves à l'échelle des établissements, avant et après la mise en œuvre de la réforme

La première approche consiste à comparer les résultats des élèves de la dernière cohorte avant la mise en place de la réforme à ceux de la première cohorte concernée par les nouvelles organisations pédagogiques. L'estimateur des doubles différences mesure ici l'effet de l'introduction des nouvelles organisations pédagogiques au cours de l'année 2024-2025. Concrètement, il s'agit de mesurer l'évolution, entre 2024 et 2025, des scores moyens aux évaluations nationales de 4^e des élèves scolarisés dans les établissements ayant mis en place des regroupements au cours de l'année 2024-2025, et de comparer cette évolution à celle observée dans les établissements n'ayant pas mis en

œuvre de regroupements spécifiques. La notion d'établissement désigne ici le collège fréquenté au cours de l'année de 5^e, qui n'est pas nécessairement celui où l'élève a passé l'évaluation de début de 4^e. Seules les évaluations nationales de début de 4^e sont mobilisées, car celles conduites en 5^e n'avaient pas encore été généralisées à tous les collèges à la rentrée 2024. La validité de l'hypothèse des tendances parallèles implique ici que les écarts de performance entre les élèves selon la nature des regroupements mis en place sont stables avant 2024, et n'auraient pas évolué différemment entre 2024 et 2025 si la réforme n'avait pas été mise en place. Cet estimateur a déjà été mobilisé par la DEPP pour l'analyse des effets à long-terme de la taille des classes (El Bah, 2026).

Le modèle de différence-de-différences (*Difference in differences*, ou DiD, en anglais) peut s'écrire de la manière suivante :

$$Y_{ist} = \alpha + \beta \cdot \text{Traitement}_s + \gamma \cdot \text{Post}_t + \tau \cdot \text{Traitement}_s \times \text{Post}_t + \epsilon_{ist} \quad (1)$$

Où :

- Y_{ist} représente le score aux évaluations nationales de 4^e de l'élève i , scolarisé en 5^e dans le collège s au cours de l'année t .
- Traitement_s est une variable binaire, ou indicatrice, définie au niveau du collège fréquenté en 5^e. Elle prend la valeur 1 pour les élèves scolarisés dans un établissement ayant mis en place une forme de regroupement, et la valeur 0 pour les élèves scolarisés dans un établissement n'ayant pas mis en place de regroupements. Cette équation est estimée séparément pour les groupes hétérogènes et pour les regroupements selon le profil scolaire des élèves. Dans le premier cas, la variable Traitement_s prend la valeur 1 pour les établissements ayant mis en place des groupes hétérogènes, et 0 pour les établissements sans regroupements. Les établissements ayant mis en place des groupes constitués selon le profil scolaire des élèves sont alors exclus de l'estimation. Dans le second cas, la variable Traitement_s prend la valeur 1 pour les établissements ayant mis en place des groupes selon le profil scolaire des élèves, et 0 pour les établissements sans regroupements. Les établissements ayant mis en place des groupes hétérogènes sont alors exclus de l'estimation.
- Post_t est une variable indicatrice qui prend la valeur 1 pour les élèves de la cohorte exposée à la réforme, c'est-à-dire scolarisée en 5^e en 2024-2025 qui a passé les évaluations nationales de 4^e en septembre 2025, et 0 pour les élèves de la cohorte précédente²⁰. Les scores de 4^e de la rentrée 2023 sont utilisés pour la recherche de tendances parallèles en amont de la réforme, mais ne sont pas pris en compte dans l'estimation de l'effet de la réforme.
- τ est le paramètre d'intérêt : il mesure l'effet moyen de la mise en place des regroupements d'élèves lors de l'année scolaire 2024-2025 dans les établissements traités par rapport à l'évolution simulée dans la situation contrefactuelle.
- ϵ_{ist} est le terme d'erreur du modèle.

Ce modèle correspond à une définition classique de la double différence, qui se concentre sur les deux périodes qui encadrent la mise en place de la réforme en mobilisant les évaluations de 4^e de 2024 et celles de 2025. Il est estimé séparément pour les scores en français et en mathématiques. L'avantage de cette approche est sa transparence et sa simplicité d'interprétation : le coefficient estimé peut être calculé manuellement à partir des scores moyens des établissements traités et témoins en 2024 et en 2025 à partir de la formule suivante :

$$\hat{\tau} = (\bar{Y}_{\{\text{traitement}=1, \text{Post}=1\}} - \bar{Y}_{\{\text{traitement}=1, \text{Post}=0\}}) - (\bar{Y}_{\{\text{traitement}=0, \text{Post}=1\}} - \bar{Y}_{\{\text{traitement}=0, \text{Post}=0\}}) \quad (2)$$

Où :

- $\hat{\tau}$ est la mesure de l'effet moyen de la mise en place des regroupements d'élèves lors de l'année scolaire 2024-2025 dans les établissements traités par rapport à l'évolution simulée dans la situation contrefactuelle ;

²⁰ Les analyses portent ici uniquement sur les élèves exposés à la réforme au cours de leur année de 5^e, en excluant donc les élèves qui redoublaient leur 4^e à la rentrée 2025.

- $\bar{Y}_{\{traitement=k, Post=l\}}$ est le résultat moyen aux évaluations nationales des élèves dont le statut de traitement est k (1 pour les élèves du groupe traité et 0 pour les élèves du groupe témoin) et dont la cohorte a le statut d'exposition l (1 pour la cohorte exposée à la réforme et 0 pour les autres).

L'une des limites de ce modèle est qu'il ne mobilise pas l'ensemble des informations contenues dans les données, en particulier l'évolution des trajectoires entre 2023 et 2024, qui est pourtant observée. Une seconde méthode est proposée pour tirer profit de cette information : il s'agit d'une autre modélisation des différences de différences avec des effets fixes pour les années et les établissements (*Two-way fixed effects*, TWFE). L'intégration des effets fixes par établissement permet de retirer de l'estimation les différences de niveau des élèves qui sont attribuables aux caractéristiques permanentes de l'établissement. Les effets fixes temporels permettent quant à eux de contrôler les spécificités de chacune des cohortes d'élèves et les chocs intervenant au cours d'une année particulière, communs à tous les établissements. Contrairement au premier modèle de doubles différences, qui repose sur les évolutions moyennes entre les groupes témoin et traité entre 2024 et 2025 exclusivement, ce second modèle repose sur l'analyse des trajectoires individuelles observées sur l'ensemble de la période, après prise en compte des caractéristiques permanentes propres à chaque établissement et des spécificités de chaque année.

Le modèle avec les effets fixes individuels et temporels s'écrit de la manière suivante :

$$Y_{ist} = \mu_s + \gamma_t + \delta \cdot Traitement_s \times Post_t + \eta_{ist} \quad (3)$$

Où :

- Y_{ist} représente le score aux évaluations nationales de 4^e de l'élève i , scolarisé en 5^e dans l'établissement s , pour l'année t .
- $Traitement_s$ est la variable binaire qui prend la valeur 1 si l'établissement est traité et 0 sinon, telle que définie dans l'équation précédente.
- $Post_t$ est une variable indicatrice qui prend la valeur 1 pour les élèves de la cohorte exposée à la réforme, c'est-à-dire pour la cohorte scolarisée en 5^e en 2024-2025 qui a passé les évaluations nationales de 4^e en septembre 2025, et 0 pour les élèves des deux cohortes précédentes. Dans ce modèle, toutes les années sont mobilisées.
- δ est le paramètre d'intérêt : il mesure l'effet moyen de la mise en place des regroupements d'élèves lors de l'année scolaire 2024-2025 dans les établissements traités par rapport à l'évolution simulée dans la situation contrefactuelle.
- μ_s est l'effet fixe « établissement », qui est propre à chaque collège fréquenté en 5^e.
- γ_t est l'effet fixe temporel qui correspond à l'année de passation des évaluations nationales de début de 4^e.
- η_{ist} est le terme d'erreur du modèle.

Ce dernier modèle est également estimé séparément pour chaque discipline. Il peut également être décomposé afin d'estimer séparément les écarts entre les groupes (traités et témoins) pour chaque année. Le coefficient mesuré pour l'année 2025 correspond alors à l'effet du traitement mesuré uniquement entre 2024 et 2025 et l'estimation pour 2023 correspond en évaluation à un test de pré-tendance (*pre-trend*) : il s'agit de rechercher des dynamiques différentes entre les établissements traités et les établissements témoins avant la réforme. Comme la réforme n'était pas encore entrée en vigueur, aucun effet significatif ne devrait être observé pour cette période. Dans le cas contraire, cela indiquerait que les différents groupes d'établissement ne suivaient pas des trajectoires similaires avant la réforme du point de vue de la réussite des élèves aux évaluations nationales de 4^e, ce qui remettrait en cause la crédibilité de l'hypothèse de tendances parallèles. Cette analyse est connue sous le nom d'analyse événementielle (ou *event study* en anglais), et correspond au modèle suivant :

$$Y_{ist} = \mu_s + \gamma_t + \sum_{j \in \{2023, 2025\}} \delta_j \cdot Traitement_s \times 1[t = j] + \eta_{ist} \quad (4)$$

Avec $1[t = j]$ une variable binaire qui vaut 1 lorsque l'année prend la valeur j , c'est-à-dire successivement 2023 et 2025, l'année 2024 étant toujours l'année de référence par rapport à laquelle sont mesurés les écarts.

Dans tous les modèles, les erreurs-types sont « clustérisées » au niveau de l'établissement dans lequel les élèves sont scolarisés en 5^e, qui correspond à l'unité de traitement. Cette modélisation permet de tenir compte de la corrélation des scores des élèves au sein d'un même établissement et conduit ainsi à une mesure plus robuste de la précision des estimations.

Cette première approche par doubles différences permet de limiter certains biais, en particulier ceux liés à « l'endogénéité » du choix des établissements d'opter pour une forme de regroupement plutôt qu'une autre en 2024 : ce choix n'est pas aléatoire et peut dépendre d'autres caractéristiques (observables ou non) potentiellement corrélées aux perspectives de progression des élèves dans ces établissements. En revanche, cette méthode d'estimation ne permet pas de neutraliser d'éventuelles différences entre cohortes qui pourraient menacer la validité de l'hypothèse des tendances parallèles. En particulier, la cohorte qui est entrée en 5^e en 2024 et qui a passé les évaluations de 4^e en 2025 est la première cohorte à avoir connu le dédoublement des classes de CP en REP (les classes de CP en REP+ étaient déjà concernées par la mesure depuis la rentrée précédente)²¹. Par ailleurs, cette cohorte a connu les confinements liés à la pandémie de Covid-19 entre le CE1 et le CE2, tandis que la cohorte précédente les a connus entre le CE2 et le CM1, ce qui pourrait avoir eu des effets différenciés sur les apprentissages.

Une autre limite susceptible de fragiliser l'analyse tiendrait au fait que les élèves ou leurs familles puissent choisir leur collège en fonction des organisations pédagogiques qui y sont mises en place. Néanmoins, 2024 étant la première année de mise en œuvre de la réforme, il est peu probable que les élèves et leurs familles aient disposé, au moment du choix de l'établissement, d'informations précises sur les regroupements mis en place à la rentrée. De telles stratégies ont toutefois pu jouer à la marge pour certaines familles particulièrement informées. En revanche, l'annonce de la réforme des groupes de besoins, dans le cadre du « Choc des savoirs », pourrait avoir influencé les choix des parents entre l'enseignement public et l'enseignement privé. En effet, bien que concernés par la réforme, les établissements privés ont très tôt annoncé leur choix de l'appliquer selon leurs propres modalités et à hauteur des moyens disponibles²². Comme les analyses réalisées dans ce document de travail s'appuient sur l'établissement fréquenté en 5^e, celles-ci ne tiennent pas compte d'éventuels effets de composition induits par les mouvements d'élèves entre les collèges du secteur public et du secteur privé.

Seconde approche : comparaison des progressions scolaires d'une même cohorte d'élèves dans le temps

La seconde approche mobilisée consiste à comparer la progression d'une année sur l'autre des élèves d'une même cohorte entre eux, selon le type de regroupements mis en place dans le collège fréquenté. La double différence repose ici sur la comparaison des résultats d'un même élève entre deux niveaux (progression de la 6^e à la 4^e, par exemple), et entre les élèves d'établissements ayant opté pour différentes organisations pédagogiques en 5^e. Deux cohortes pourraient ainsi être étudiées séparément : celle des élèves entrés en 6^e en 2024, exposés à la réforme au cours de leur première année au collège, et celle des élèves entrés en 5^e en 2024, concernés par les organisations pédagogiques en classe de 5^e. En raison des particularités des évaluations de 5^e, qui ne reposent que sur l'évaluation de trois domaines spécifiques, seule cette seconde cohorte est retenue dans l'analyse (**Annexe 6**). Il s'agit donc, comme pour la première approche, des élèves exposés à la réforme au cours de leur année de 5^e en 2024-2025, qui ont passé les évaluations nationales de 4^e à la rentrée

²¹ Une étude récente de la DEPP nuance cependant l'importance des effets de long terme de cette politique, en particulier sur les résultats aux évaluations nationales de 6^e (El Bah, 2026).

²² Conférence de presse du 27 mars 2024 du secrétaire général de l'enseignement catholique (dépêche AEF n°709811).

2025. Cette cohorte d'élèves a passé les évaluations nationales de CP et de mi-CP en 2018-2019, de CE1 en 2019-2020, de 6^e en 2023-2024. Les scores obtenus à ces évaluations sont mobilisés dans cette seconde approche.

L'estimateur mesure moins ici l'effet de l'introduction des regroupements d'élèves que l'effet de l'exposition différentielle à différentes modalités d'organisation pédagogique au cours de la scolarité. Il s'agit ainsi de comparer les progressions d'élèves au sein d'une même cohorte, selon qu'ils aient été regroupés de manière hétérogène, selon leur profil scolaire, ou bien non regroupés en classe de 5^e. Le modèle de régression pour les doubles différences (DiD) permettant d'estimer l'effet de l'exposition à ces différentes formes de regroupement peut être écrit de la manière suivante :

$$Y_{ist} = \alpha + \beta \cdot \text{Traitement}_s + \gamma \cdot \text{Niveau}_t + \tau \cdot \text{Traitement}_s \times \text{Niveau}_t + \epsilon_{ist} \quad (5)$$

Où :

- Y_{ist} représente le score aux évaluations de l'élève i , scolarisé en 5^e dans le collège s , pour le niveau t .
- Traitement_s est une variable indicatrice définie au niveau du collège fréquenté en 5^e, qui prend la valeur 1 pour les élèves des établissements ayant mis en place une forme de regroupement, et 0 pour les élèves issus des établissements qui n'ont pas mis en place de regroupements d'élèves. Cette équation est estimée séparément pour les groupes hétérogènes et pour les regroupements selon le profil des élèves. Dans le premier cas, Traitement_s prend la valeur 1 pour les établissements avec des groupes hétérogènes et 0 pour les établissements sans regroupements. Les établissements avec des regroupements constitués selon le profil scolaire des élèves ne sont pas pris en compte dans l'estimation de ce modèle. C'est le contraire dans le second cas, pour l'estimation de l'effet d'avoir été dans un établissement avec des regroupements constitués selon le profil scolaire des élèves. Pour la comparaison entre les groupes selon le profil scolaire des élèves et les groupes hétérogènes, Traitement_s prend la valeur 1 si l'établissement a instauré des groupes selon le profil scolaire et 0 sinon, les établissements n'ayant pas mis en place de regroupements d'élèves étant retirés de l'échantillon pour cette estimation.
- Niveau_t est une variable indicatrice qui prend la valeur 1 pour le niveau 4^e et 0 pour le niveau 6^e.
- τ est le paramètre d'intérêt : il mesure l'effet moyen de l'exposition aux regroupements d'élèves par rapport à l'évolution simulée dans la situation contrefactuelle. C'est ce paramètre qui est estimé à l'aide de l'estimateur des différences de différences.
- ϵ_{ist} est le terme d'erreur du modèle économétrique.

Ce modèle est estimé séparément pour chaque domaine ou discipline. Cette seconde approche permet de neutraliser l'effet de sélection à l'échelle des établissements tout en évitant d'introduire des biais liés à la comparaison de cohortes différentes. Les facteurs qui interviennent au sein d'une cohorte ne sont cependant pas éliminés par cette approche : c'est notamment le cas du choix de l'établissement en fonction du type de regroupement susceptible d'y être mis en œuvre, ou encore des mouvements entre enseignement public et enseignement privé.

Comme pour la première approche, un second modèle de doubles différences avec des effets fixes individuels et temporels (*Two-way fixed effects*, ou TWFE) a été estimé, pour chaque domaine et discipline :

$$Y_{ist} = \mu_i + \gamma_t + \delta \cdot \text{Traitement}_s \times \text{Post}_t + \eta_{ist} \quad (6)$$

Où :

- Y_{ist} représente le score aux évaluations de l'élève i , scolarisé en 5^e dans le collège s , pour le niveau t .

- $Traitement_s$ est la variable binaire qui prend la valeur 1 si l'établissement est traité et 0 sinon, telle que définie dans l'équation précédente.
- $Post_t$ est une variable indicatrice qui prend la valeur 1 si le score correspond au score en 4^e, et 0 sinon. Dans ce modèle, tous les niveaux sont mobilisés, du CP à la 4^e.
- δ est le paramètre d'intérêt : il mesure l'effet moyen de l'exposition aux regroupements d'élèves par rapport à l'évolution simulée dans la situation contrefactuelle.
- μ_i est l'effet fixe individuel qui correspond à l'identifiant unique de l'élève.
- γ_t est l'effet fixe temporel qui correspond au niveau de l'évaluation (CP, mi-CP, CE1, 6^e ou 4^e).
- η_{ist} est le terme d'erreur du modèle.

Aussi, une analyse événementielle est également réalisée afin de discuter de la pertinence de l'hypothèse des tendances parallèles :

$$Y_{ist} = \mu_i + \gamma_t + \sum_{j \in \{CP, miCP, CE1, 4^e\}} \delta_j \cdot Traitement_s \times 1[t = j] + \eta_{ist} \quad (7)$$

Avec $1[t = j]$ une variable binaire qui vaut 1 lorsque le niveau prend la valeur j , c'est-à-dire successivement CP, mi-CP, CE1 et 4^e, la 6^e étant toujours le niveau de référence par rapport auquel sont mesurés les écarts.

Cette seconde approche nécessite un travail d'harmonisation des scores des élèves afin de pouvoir mesurer leur progression d'une année sur l'autre. En effet, les scores standardisés des évaluations nationales sont construits de manière à être comparables pour un même niveau, mais pas d'un niveau à l'autre. Il est ainsi possible de comparer directement les résultats aux évaluations nationales des élèves entrés en 4^e en septembre 2024 et ceux des élèves entrés en 4^e en 2025, mais il n'est pas possible de comparer les résultats des élèves entrés en 2023 en 6^e avec les résultats de ces mêmes élèves à leur entrée en 4^e en 2025. De nouveaux scores standardisés sont donc calculés pour les différentes compétences évaluées chaque année du CP à la 4^e afin de les rendre comparables et de mesurer leur progression (voir [Annexe 6](#)).

Résultats de la première approche : comparaison des élèves de la première cohorte exposée aux nouvelles organisations pédagogiques avec ceux de la cohorte précédente

La première méthode d'estimation de l'effet des regroupements d'élèves sur leurs performances scolaires repose sur la comparaison des élèves de la première cohorte exposée à la réforme avec ceux de la dernière cohorte l'ayant précédée. Les données mobilisées sont les évaluations nationales de début de 4^e : les élèves scolarisés en 4^e en 2025 sont les premiers à avoir été potentiellement concernés par les nouveaux regroupements d'élèves, pendant un an au cours de leur année de 5^e, en 2024-2025. Les élèves qui ont passé les évaluations nationales de 4^e en 2024 n'ont jamais été confrontés aux nouvelles organisations pédagogiques de la rentrée 2024. Il s'agit donc ici de comparer d'une part les élèves de ces deux cohortes (première différence), et d'autre part les élèves d'une même cohorte mais scolarisés en 5^e au sein d'établissements ayant mis en place différentes organisations pédagogiques au cours de l'année scolaire 2024-2025 (seconde différence). En particulier, les élèves des établissements qui organisent des regroupements d'élèves, hétérogènes ou selon le profil scolaire (établissements traités), sont comparés à ceux des établissements qui n'ont pas organisé de regroupement d'élèves (établissements témoins). Le statut d'un établissement est défini pour chacune des années (2023, 2024 et 2025) selon les organisations pédagogiques mises en œuvre en classe de 5^e au cours de l'année 2024-2025 au sein de cet établissement.

Validité de l'hypothèse des tendances parallèles

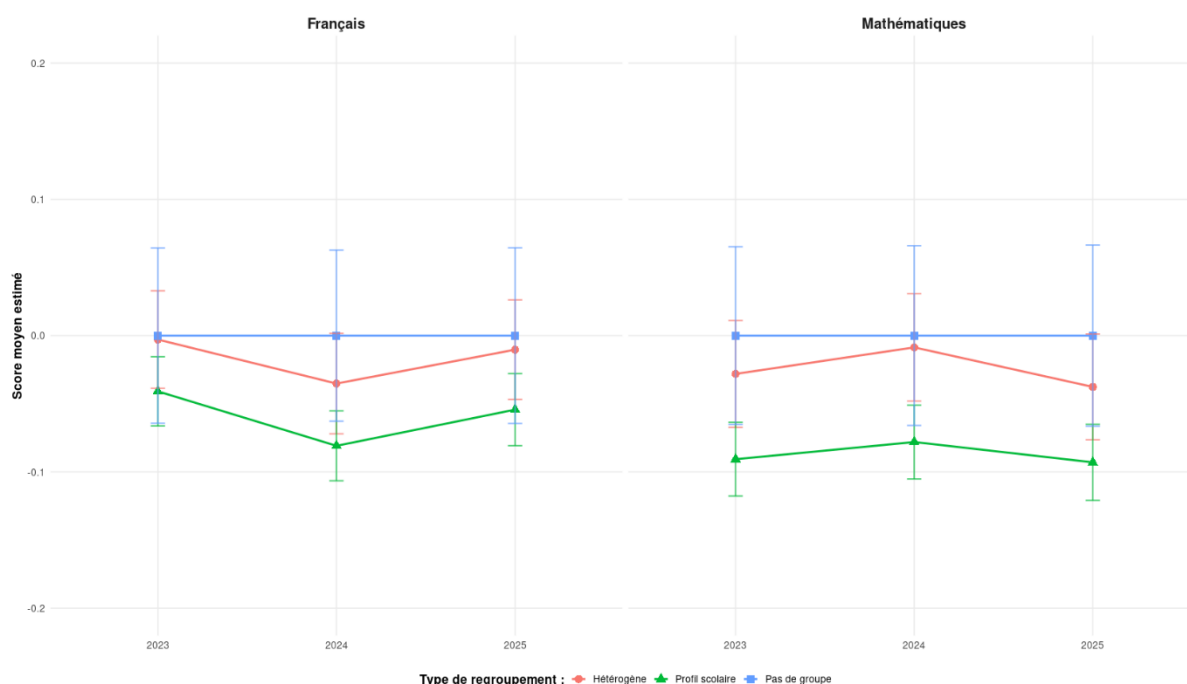
Les établissements qui ont choisi une forme de regroupement plutôt qu'une autre peuvent présenter certaines caractéristiques observables et inobservables qui les distinguent des collèges ayant choisi d'autres organisations pédagogiques ou ayant choisi de ne pas regrouper les élèves en 2024-2025. Malgré ces différences inhérentes, il est possible d'identifier l'effet des regroupements en se reposant

sur la dimension temporelle : si l'écart de performance moyenne observé entre ces différents collèges est stable au cours des années qui précèdent la réforme, alors on peut supposer que l'influence éventuelle des caractéristiques des établissements sur le niveau des élèves est stable chaque année. Une variation significative de l'écart entre les établissements au cours du temps signalerait alors un choc exogène, comme par exemple la mise en place d'une nouvelle politique éducative, ou une tendance à la divergence qui aurait pu se poursuivre en l'absence d'intervention. Il s'agit ici d'une condition qui n'est pas suffisante pour garantir la validité de l'approche, mais qui permet de renforcer sa crédibilité.

Les évaluations nationales de début de 4^e n'ont été généralisées qu'à partir de la rentrée 2023. Ainsi, la variation des scores des élèves au cours du temps selon les organisations pédagogiques retenues par leur établissement de 5^e ne peuvent être observées qu'entre deux années avant l'entrée en vigueur de la mesure. Pour l'analyse causale, les scores standardisés de 4^e présentés précédemment ont été centrés et réduits selon les scores obtenus par les élèves du groupe témoin ([Annexe 6](#)). Cela permet de lire les évolutions sur une échelle qui est couramment utilisée dans la recherche en éducation, et qui permet de comparer les effets à ceux d'autres interventions.

Visuellement, entre 2023 et 2024, l'écart entre les élèves scolarisés en 5^e dans les établissements qui ont mis en place des regroupements d'élèves au cours de l'année 2024-2025 (appelés « traités » par la suite) et les élèves des établissements qui n'ont pas organisé de regroupement d'élèves au cours de l'année 2024-2025 (dits « témoins ») s'accroît légèrement en français et se réduit légèrement en mathématiques ([Figure 16](#)). Mais ces mouvements sont de faible ampleur et statistiquement peu significatifs.

FIGURE 16 • Évolution des scores moyens de 4^e standardisés entre 2023 et 2025 selon le type de regroupements mis en place



Lecture : En 2023, l'écart des scores moyens aux évaluations nationales de début de 4^e en français était presque nul (-0,003) entre les élèves ayant effectué leur 5^e au sein de collèges qui mettent en place des regroupements hétérogènes en français en 5^e au cours de l'année 2024-2025 et ceux ayant été scolarisés en 5^e au sein d'établissements qui ne mettent en place aucune forme de regroupement. En 2024, cet écart s'est creusé et atteint -0,04.

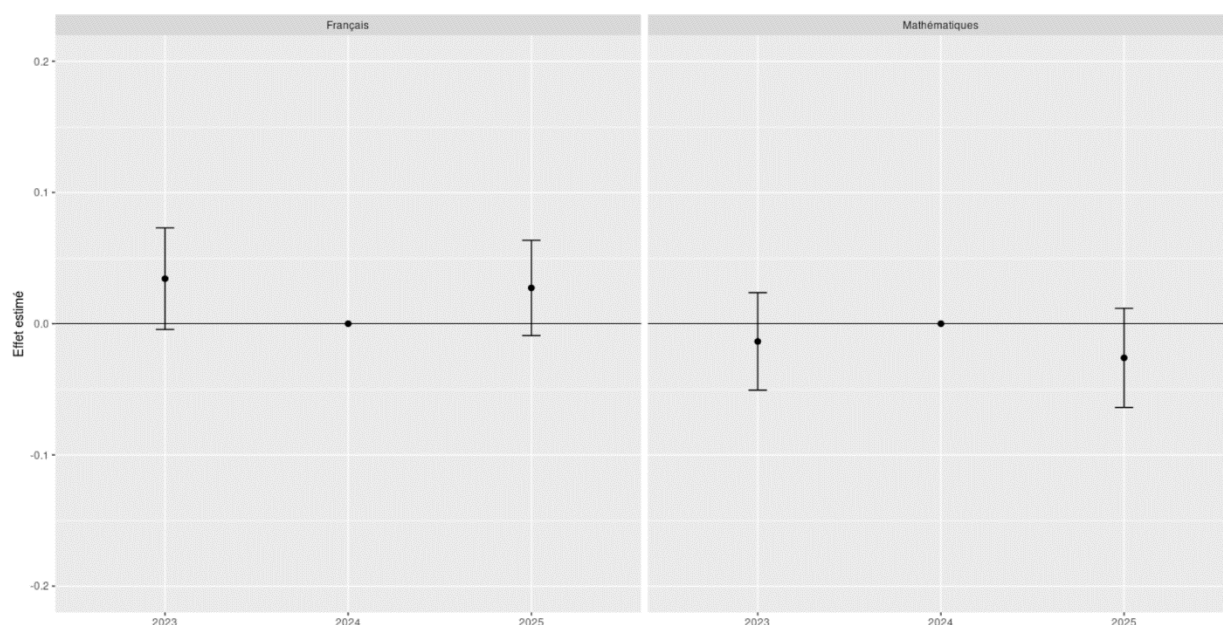
Note : Pour chaque année, les scores sont standardisés par rapport aux scores des élèves scolarisés en 5^e dans les établissements qui ne mettent en place aucune forme de regroupement au cours de l'année scolaire 2024-2025. Les intervalles de confiance à 95 % sont affichés sur le graphique.

Champ : Élèves scolarisés en 5^e au sein des établissements de l'enquête EEC ayant adopté une forme unique de regroupement ou aucune en 5^e pour la discipline. Cela correspond à 1 289 collèges pour le français et 1 291 collèges pour les mathématiques.

Sources : DEPP, Évaluations nationales en français et mathématiques de 4^e (2023, 2024, 2025) ; enquête EEC, questionnaires Chefs d'établissement, printemps 2025 ; Constat de rentrée 2025.

Les modèles d'analyse événementielle permettent de mesurer plus formellement la crédibilité de l'hypothèse des tendances parallèles tout en tenant compte, à travers les effets fixes par établissement fréquenté en 5^e et par année, des différences structurelles observées entre les établissements et des chocs susceptibles d'affecter l'ensemble des élèves d'une cohorte en particulier (Équation 4). En neutralisant ces différences, le constat est le même que celui observé à partir de la Figure 16 : les scores moyens de 4^e en français suivent des évolutions qui sont peu différentes entre les élèves des établissements avec des groupes hétérogènes et ceux des établissements sans regroupements (Figure 17). Ainsi, avant la mise en place des nouvelles organisations pédagogiques, la différence observée en septembre 2023 n'est pas statistiquement différente du même écart en septembre 2024 (cet écart est normalisé à zéro dans la Figure 17). On trouve une variation légèrement significative entre ces deux années quand on compare les collèges qui mettent en place des regroupements selon le profil scolaire des élèves et les collèges témoins (Figure 18) : mais cet écart est faible et marginalement significatif. En mathématiques, il n'y a pas de différences significatives dans les trajectoires définies selon le même principe, quel que soit le traitement considéré (Figures 17, 18 & 19). Ainsi, l'hypothèse des tendances parallèles semble crédible dans toutes les situations, avec une nuance pour la comparaison des scores de français entre les élèves des établissements avec des regroupements selon le profil scolaire des élèves et ceux des établissements sans regroupements (Figure 18).

FIGURE 17 • Différence de scores aux évaluations nationales à l'entrée en 4^e avant et après la mise en place de la réforme (comparaison : groupes hétérogènes/pas de regroupement d'élèves)



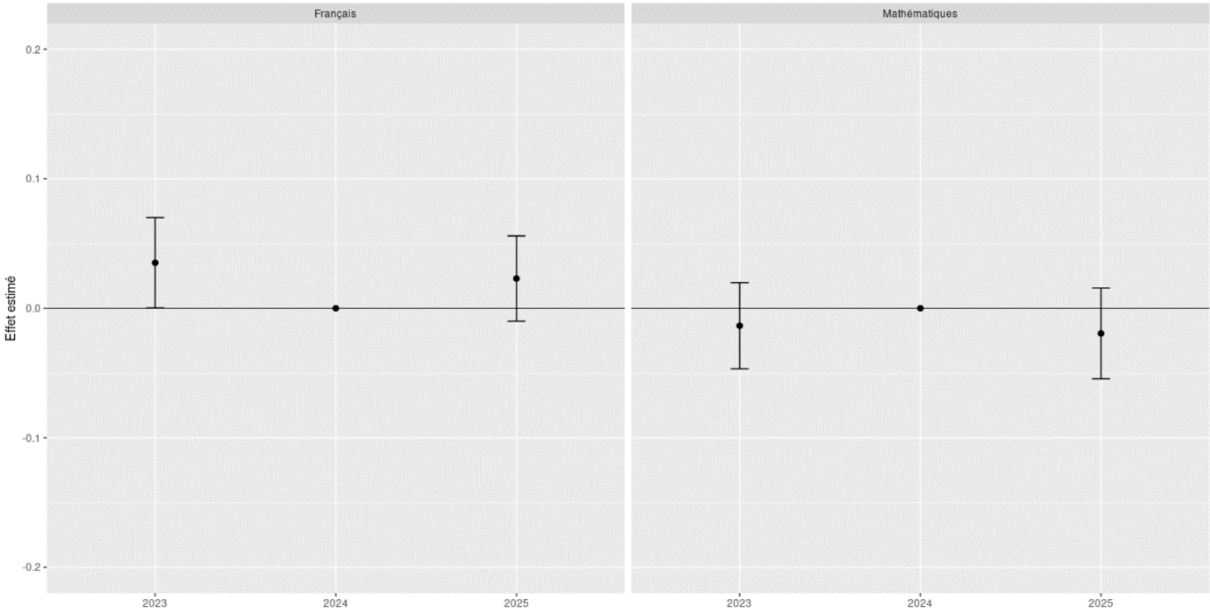
Lecture : Entre septembre 2023 et septembre 2024, le score moyen de français des élèves ayant effectué leur 5^e au sein de collèges qui mettent en place des regroupements hétérogènes en français en 5^e au cours de l'année 2024-2025 a diminué de manière plus marquée (ou a moins progressé) que celui des élèves ayant été scolarisés en 5^e au sein d'établissements qui ne mettent en place aucune forme de regroupement : le coefficient de l'analyse événementielle est égal à 0,03, avec l'année 2024 comme référence. Ce coefficient n'est pas significatif au seuil de 5 %. Entre 2024 et 2025, le score moyen de français a davantage progressé (ou moins diminué) pour les élèves scolarisés dans des établissements qui mettent en place des groupes hétérogènes que pour ceux des établissements sans aucune forme de regroupement : le coefficient est égal à 0,03. Cependant, ce coefficient n'est pas non plus significatif au seuil de 5 %.

Note : Pour chaque année, les scores sont standardisés par rapport aux scores des élèves scolarisés en 5^e dans les établissements qui ne mettent en place aucune forme de regroupement au cours de l'année scolaire 2024-2025. Les erreurs-types sont clusterisées au niveau de l'établissement fréquenté en 5^e. Les intervalles de confiance à 95 % sont affichés sur le graphique. L'année 2024 est l'année de référence.

Champ : Élèves scolarisés en 5^e au sein des établissements de l'enquête EEC ayant mis en place des regroupements hétérogènes ou aucune forme de regroupement en 5^e pour la discipline. Cela correspond à 484 collèges pour le français et 473 collèges pour les mathématiques.

Sources : DEPP, Évaluations nationales en français et mathématiques de 4^e (2023, 2024, 2025) ; enquête EEC, questionnaires Chefs d'établissement, printemps 2025 ; Constat de rentrée 2025.

FIGURE 18 • Différence de scores aux évaluations nationales à l’entrée en 4^e avant et après la mise en place de la réforme (comparaison : groupes composés selon le profil scolaire des élèves/pas de regroupement d’élèves)



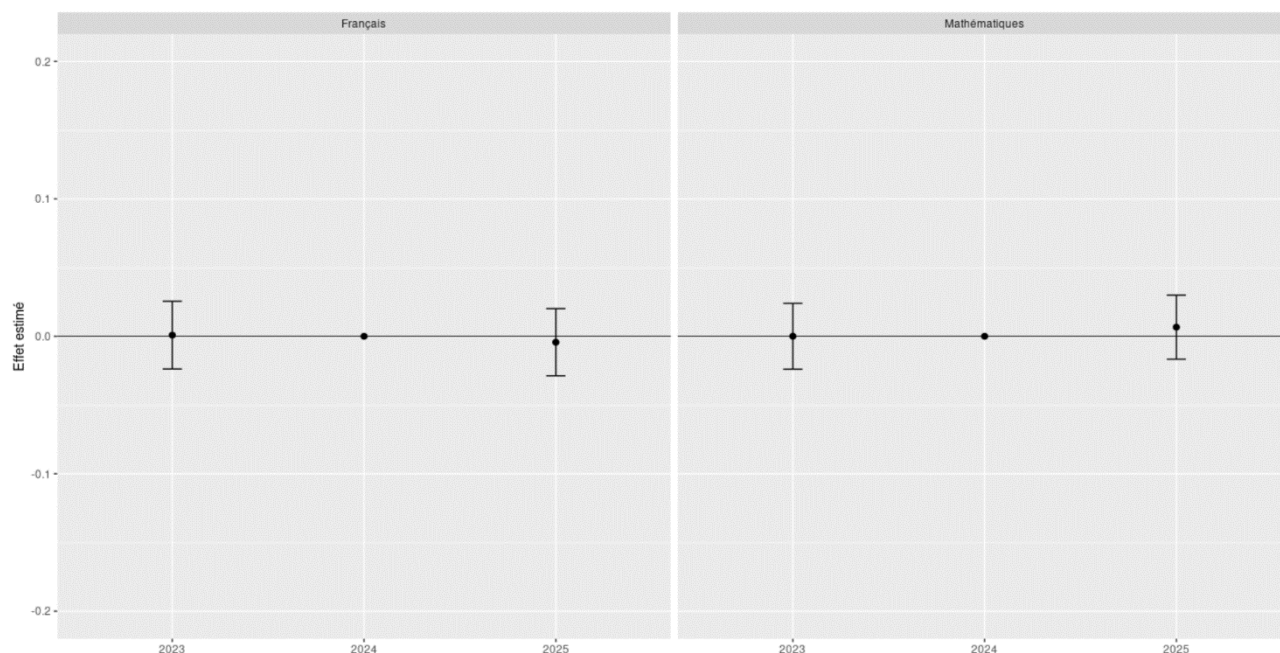
Lecture : Entre septembre 2023 et septembre 2024, le score moyen de français des élèves ayant effectué leur 5^e au sein de collèges qui mettent en place des regroupements selon le profil scolaire des élèves en français en 5^e au cours de l’année 2024-2025 a diminué de manière plus marquée (ou a moins progressé) que celui des élèves ayant été scolarisés en 5^e au sein d’établissements qui ne mettent en place aucune forme de regroupement : le coefficient de l’analyse événementielle est égal à 0,04, avec l’année 2024 comme référence. Ce coefficient est significatif au seuil de 5 %. Entre 2024 et 2025, le score moyen de français a davantage progressé (ou moins diminué) pour les élèves scolarisés dans des établissements qui mettent en place des groupes selon le profil scolaire des élèves que pour ceux des établissements sans aucune forme de regroupement : le coefficient est égal à 0,02. Cependant, ce coefficient n’est pas significatif au seuil de 5 %.

Note : Pour chaque année, les scores sont standardisés par rapport aux scores des élèves scolarisés en 5^e dans les établissements qui ne mettent en place aucune forme de regroupement au cours de l’année scolaire 2024-2025. Les erreurs-types sont clusterisées au niveau de l’établissement fréquenté en 5^e. Les intervalles de confiance à 95 % sont affichés sur le graphique. L’année 2024 est l’année de référence.

Champ : Élèves scolarisés en 5^e au sein des établissements de l’enquête EEC ayant mis en place des regroupements selon le profil scolaire des élèves ou aucune forme de regroupement en 5^e pour la discipline. Cela correspond à 948 collèges pour le français et 958 collèges pour les mathématiques.

Sources : DEPP, Évaluations nationales en français et mathématiques de 4^e (2023, 2024, 2025) ; enquête EEC, questionnaires Chefs d’établissement, printemps 2025 ; Constat de rentrée 2025.

FIGURE 19 • Différence de scores aux évaluations nationales à l'entrée en 4^e avant et après la mise en place de la réforme (comparaison : groupes composés selon le profil scolaire des élèves/groupes hétérogènes)



Lecture : Entre septembre 2023 et septembre 2024, le score moyen de français des élèves ayant effectué leur 5^e au sein de collèges qui mettent en place des regroupements selon le profil scolaire des élèves en français en 5^e au cours de l'année 2024-2025 a évolué comme celui des élèves scolarisés en 5^e au sein des établissements qui mettent en place des regroupements hétérogènes au cours de l'année scolaire 2024-2025 : le coefficient de l'analyse événementielle est égal à 0,0009, avec l'année 2024 comme référence, et n'est pas significatif au seuil de 5 %. Entre 2024 et 2025, le score moyen de français a relativement moins progressé (ou davantage diminué) pour les élèves scolarisés dans des établissements qui mettent en place des groupes selon le profil scolaire des élèves que pour ceux des établissements avec des regroupements hétérogènes : le coefficient est égal à -0,004. Cependant, ce coefficient n'est pas significatif au seuil de 5 %.

Note : Pour chaque année, les scores sont standardisés par rapport aux scores des élèves scolarisés en 5^e dans les établissements qui ne mettent en place aucune forme de regroupement au cours de l'année scolaire 2024-2025. Les erreurs-types sont clusterisées au niveau de l'établissement fréquenté en 5^e. Les intervalles de confiance à 95 % sont affichés sur le graphique. L'année 2024 est l'année de référence.

Champ : Élèves scolarisés en 5^e au sein des établissements de l'enquête EEC ayant mis en place des regroupements selon le profil scolaire des élèves ou des groupes hétérogènes en 5^e pour la discipline. Cela correspond à 1 146 collèges pour le français et 1 151 collèges pour les mathématiques.

Sources : DEPP, Évaluations nationales en français et mathématiques de 4^e (2023, 2024, 2025) ; enquête EEC, questionnaires Chefs d'établissement, printemps 2025 ; Constat de rentrée 2025.

Une absence d'effet des nouvelles organisations pédagogiques sur la réussite des élèves aux évaluations nationales de 4^e

L'effet estimé des regroupements sur l'évolution des scores aux évaluations nationales est représenté par le coefficient présenté pour septembre 2025 dans les **Figure 17 à 19**. On remarque que les résultats de l'analyse par doubles différences confirment les intuitions obtenues à partir de l'observation des évolutions corrélationnelles des évaluations nationales : une absence d'effet significatif des regroupements sur l'évolution des résultats.

Par rapport à l'absence de regroupements, la présence de regroupements hétérogènes suggère un effet positif sur la performance des élèves en français en 4^e, mais le coefficient associé à l'effet du traitement est très petit (2,5 % d'un écart-type) et n'est pas statistiquement significatif (**Figure 17 ; Tableau 7**). L'effet mesuré est d'une magnitude similaire aux écarts observés avant la mise en place de la réforme, et il reflète donc sans doute simplement des variations normales des résultats aux tests d'une année sur l'autre. Il en va de même de la comparaison entre les élèves des établissements avec des groupes constitués selon le profil scolaire des élèves et ceux des établissements sans regroupements, avec un effet de même ampleur et non-significatif (**Figure 18 ; Tableau 8**). Sur ce dernier effet, cependant, l'analyse des tendances menée dans la section précédente nous conduit à être prudent sur la validité du protocole d'estimation.

En mathématiques, la présence de regroupements, qu'ils soient hétérogènes ou constitués selon le profil scolaire des élèves, semble avoir un effet négatif sur la réussite des élèves (**Figures 17 & 18 ;**

Tableaux 7 & 8). Cependant, ces résultats ne sont pas significatifs et sont, là aussi, de même ampleur que les écarts observés sur la période qui précède l’entrée en vigueur de la réforme.

Sur la base de cette analyse, il est possible de conclure **qu’on ne peut pas détecter d’effet des différentes organisations pédagogiques sur les résultats moyens des élèves**, et ceci malgré le fait que la taille des échantillons pourrait nous permettre de détecter des effets, même très petits (de l’ordre de 4 % d’un écart-type), s’ils existaient. **Ce résultat est cohérent avec la littérature scientifique** (voir Section « Revue de littérature »).

En toute logique, la comparaison entre les élèves des établissements avec des groupes hétérogènes et ceux des établissements avec des groupes constitués selon le profil scolaire ne permet pas non plus de conclure à un effet d’un type de regroupement plutôt qu’un autre sur les performances des élèves en français et en mathématiques (**Figure 19 ; Tableau 9**).

Les deux modélisations de la double différence, l’une comparant uniquement les scores de septembre 2024 à ceux de 2025 (modèle « DiD » dans les tableaux ci-dessous) et la seconde prenant en compte l’ensemble de la période avec des effets fixes par année et par établissement (modèle « TWFE »), conduisent à des conclusions similaires.

La littérature sur les regroupements d’élèves suggère que leurs effets pourraient varier selon le niveau initial des élèves ou selon leur sexe. Des modèles d’analyse événementielle ont donc été estimés afin d’explorer ces dimensions sur des sous-échantillon d’élèves définis selon leur niveau à l’entrée en 6^e (mesurés par tercile), et selon leur sexe (**Annexe 7**). Compte tenu des données disponibles et de la définition du traitement à l’échelle de l’établissement, ces analyses restent toutefois peu précises et ne permettent pas de mettre en évidence d’effets hétérogènes significatifs à partir de ce modèle de doubles différences reposant sur la comparaison des cohortes.

TABLEAU 7 • Coefficients de l’estimation des effets des groupes hétérogènes par rapport à l’absence de regroupements sur les scores aux évaluations nationales de 4^e

Discipline	Français		Mathématiques	
Groupe traité	Établissements ayant mis en place des groupes hétérogènes			
Groupe témoin	Établissements n'ayant mis en place aucun regroupement			
Modèle	DiD	TWFE	DiD	TWFE
Coefficient	0,025 (0,019)	0,01 (0,017)	-0,029 (0,02)	-0,019 (0,017)
Crédibilité des tendances parallèles	Oui	Oui	Oui	Oui
Effets fixes	Non	UAI en 5 ^e + année	Non	UAI en 5 ^e + année
Cluster	UAI en 5 ^e	UAI en 5 ^e	UAI en 5 ^e	UAI en 5 ^e
N obs.	107 857	163 926	104 961	159 256
N établissements	484	484	473	473

Lecture : Pour le français, le modèle de doubles différences sans effet fixe estime un coefficient de 0,025 pour l’effet du traitement : le fait d’appartenir à la cohorte d’élèves qui entre en 5^e en 2024-2025 et d’avoir été scolarisé dans un établissement qui met en place des groupes hétérogènes plutôt que dans un établissement qui ne met en place aucune forme de regroupement augmente en moyenne le score en français de 0,025, soit 2,5 % d’un écart-type. Ce coefficient n’est cependant pas significatif.

Note : Les modèles DiD correspondent aux modèles de doubles différences classiques (**Équation 1**) tandis que les modèles TWFE désignent les modèles avec effets fixes (**Équation 3**). Les erreurs types clusterisées au niveau de l’établissement fréquenté en 5^e sont présentées entre parenthèses. Seuils de significativité : « *** », pour un coefficient significatif à 0,1 %, « ** », pour un coefficient significatif à 1 %, « * », pour un coefficient significatif à 5 % et « . » pour un coefficient significatif à 10 %.

Champ : Élèves scolarisés en 5^e au sein des établissements de l’enquête EEC ayant mis en place des regroupements hétérogènes ou aucune forme de regroupement en 5^e pour la discipline. Cela correspond à 484 collèges pour le français et 473 collèges pour les mathématiques.

Sources : DEPP, Évaluations nationales en français et mathématiques de 4^e (2023, 2024, 2025) ; enquête EEC, questionnaires Chefs d’établissement, printemps 2025 ; Constat de rentrée 2025.

TABLEAU 8 • Coefficients de l'estimation des effets des groupes composés selon le profil scolaire des élèves par rapport à l'absence de regroupements sur les scores aux évaluations nationales de 4^e

Discipline	Français		Mathématiques	
Groupe traité	Établissements ayant mis en place des groupes selon le profil scolaire des élèves			
Groupe témoin	Établissements n'ayant mis en place aucun regroupement			
Modèle	DiD	TWFE	DiD	TWFE
Coefficient	0,027 (0,018)	0,005 (0,016)	-0,015 (0,019)	-0,013 (0,016)
Crédibilité des tendances parallèles	Non	Non	Oui	Oui
Effets fixes	Non	UAI en 5 ^e + année	Non	UAI en 5 ^e + année
Cluster	UAI en 5 ^e	UAI en 5 ^e	UAI en 5 ^e	UAI en 5 ^e
N obs.	220 463	333 058	224 206	339 003
N établissements	948	948	958	958

Lecture : Pour le français, le modèle de doubles différences sans effet fixe estime un coefficient de 0,027 pour l'effet du traitement : le fait d'appartenir à la cohorte d'élèves qui entre en 5^e en 2024-2025 et d'avoir été scolarisé dans un établissement qui met en place des groupes constitués selon le profil scolaire des élèves plutôt que dans un établissement qui ne met en place aucune forme de regroupement augmente en moyenne le score en français de 0,027, soit 2,7 % d'un écart-type. Ce coefficient n'est cependant pas significatif.

Note : Les modèles DiD correspondent aux modèles de doubles différences classiques (Équation 1) tandis que les modèles TWFE désignent les modèles avec effets fixes (Équation 3). Les erreurs types clusterisées au niveau de l'établissement fréquenté en 5^e sont présentées entre parenthèses. Seuils de significativité : « *** », pour un coefficient significatif à 0,1 %, « ** », pour un coefficient significatif à 1 %, « * », pour un coefficient significatif à 5 % et « . » pour un coefficient significatif à 10 %.

Champ : Élèves scolarisés en 5^e au sein des établissements de l'enquête EEC ayant mis en place des regroupements selon le profil scolaire des élèves ou aucune forme de regroupement en 5^e pour la discipline. Cela correspond à 948 collèges pour le français et 958 collèges pour les mathématiques.

Sources : DEPP, Évaluations nationales en français et mathématiques de 4^e (2023, 2024, 2025) ; enquête EEC, questionnaires Chefs d'établissement, printemps 2025 ; Constat de rentrée 2025.

TABLEAU 9 • Coefficients de l'estimation des effets des groupes composés selon le profil scolaire des élèves par rapport aux groupes hétérogènes sur les scores aux évaluations nationales de 4^e

Discipline	Français		Mathématiques	
Groupe traité	Établissements ayant mis en place des groupes selon le profil scolaire des élèves			
Groupe témoin	Établissements ayant mis en place des groupes hétérogènes			
Modèle	DiD	TWFE	DiD	TWFE
Coefficient	0,002 (0,013)	-0,005 (0,011)	0,014 (0,012)	0,007 (0,011)
Validité des tendances parallèles	Oui	Oui	Oui	Oui
Effets fixes	Non	UAI en 5 ^e + année	Non	UAI en 5 ^e + année
Cluster	UAI en 5 ^e	UAI en 5 ^e	UAI en 5 ^e	UAI en 5 ^e
N obs.	269 562	407 748	271 027	410 007
N établissements	1 146	1 146	1 151	1 151

Lecture : Pour le français, le modèle de doubles différences sans effet fixe estime un coefficient de 0,002 pour l'effet du traitement : le fait d'appartenir à la cohorte d'élèves qui entre en 5^e en 2024-2025 et d'avoir été scolarisé dans un établissement qui met en place des groupes constitués selon le profil scolaire des élèves plutôt que dans un établissement qui met en place des regroupements hétérogènes augmente en moyenne le score en français de 0,002, soit 0,2 % d'un écart-type. Ce coefficient n'est cependant pas significatif.

Note : Les modèles DiD correspondent aux modèles de doubles différences classiques (Équation 1) tandis que les modèles TWFE désignent les modèles avec effets fixes (Équation 3). Les erreurs types clusterisées au niveau de l'établissement fréquenté en 5^e sont présentées entre parenthèses. Seuils de significativité : « *** », pour un coefficient significatif à 0,1 %, « ** », pour un coefficient significatif à 1 %, « * », pour un coefficient significatif à 5 % et « . » pour un coefficient significatif à 10 %.

Champ : Élèves scolarisés en 5^e au sein des établissements de l'enquête EEC ayant mis en place des regroupements selon le profil scolaire des élèves ou aucune forme de regroupement en 5^e pour la discipline. Cela correspond à 1 146 collèges pour le français et 1 151 collèges pour les mathématiques.

Sources : DEPP, Évaluations nationales en français et mathématiques de 4^e (2023, 2024, 2025) ; enquête EEC, questionnaires Chefs d'établissement, printemps 2025 ; Constat de rentrée 2025.

Résultats de la seconde approche : comparaison de la progression des élèves selon leur scolarisation dans un établissement ayant introduit ou non de nouvelles organisations pédagogiques en 2024-2025

La seconde approche mobilisée pour évaluer l'effet de la réforme des organisations pédagogiques consiste à comparer la progression des élèves selon leur scolarisation ou non au sein d'un établissement ayant introduit une forme de regroupement d'élèves au cours de l'année 2024-2025 à celle des élèves scolarisés dans des collèges qui n'ont pas mis en place de regroupement d'élève lors de cette même année. À nouveau, il s'agit de comparer des élèves à la fois entre établissements mais également au cours du temps. La différence avec la première approche tient ici au fait que la dimension temporelle se mesure ici à l'échelle d'un même élève, pour chaque niveau, et non entre des élèves de cohortes différentes.

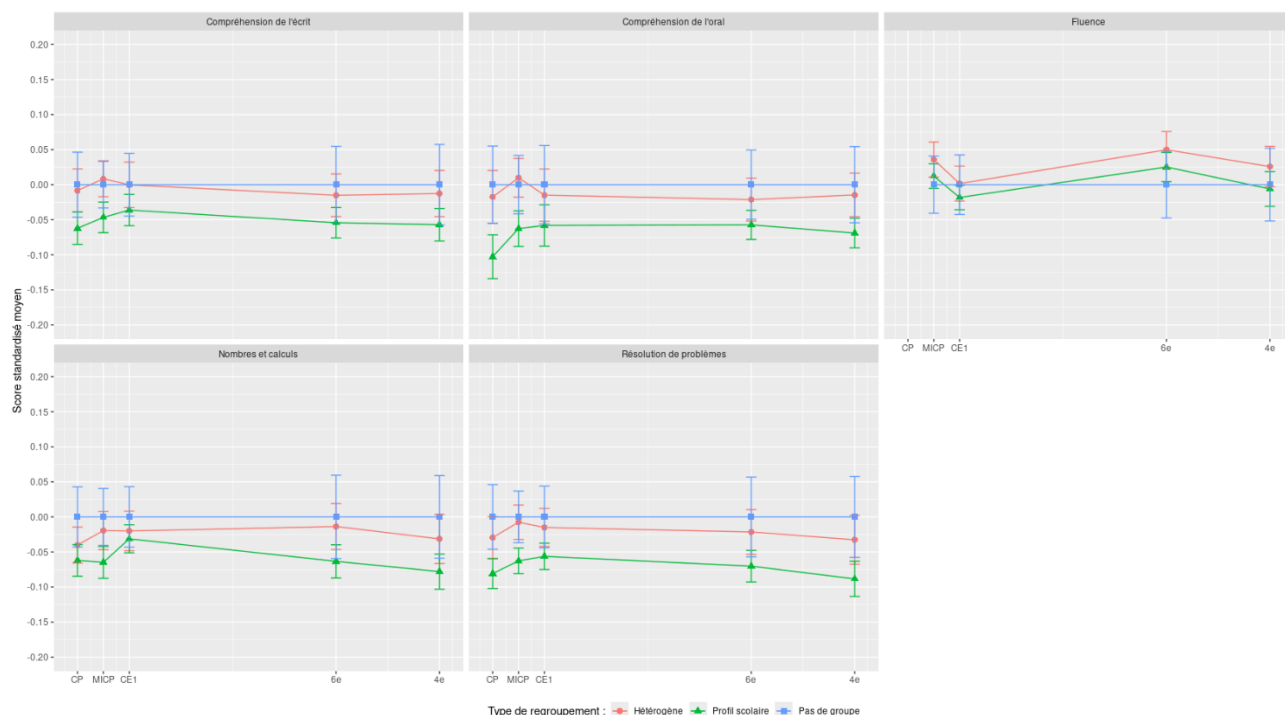
Validité de l'hypothèse des tendances parallèles

L'hypothèse des tendances parallèles se traduit ici par la stabilité de l'écart des scores au cours de la scolarité passée entre les élèves qui auront effectué leur 5^e au sein d'un établissement ayant mis en place au moins une nouvelle forme d'organisation pédagogique (établissements dits « traités ») et ceux qui auront effectué leur 5^e au sein d'un collège n'ayant mis en place aucun regroupement (établissements dits « témoins »). Ainsi, même si le niveau scolaire des élèves des établissements traités diffère structurellement de celui des élèves des établissements témoins, il est toujours possible de mesurer l'effet d'une réforme en examinant si les écarts de performances entre ces différents élèves s'accroissent ou se réduisent, précisément lorsque certains d'entre eux sont exposés à l'intervention. Cette approche est valide si l'influence des différents facteurs observés ou inobservés susceptibles d'affecter les performances des élèves reste stable dans le temps, et si seule l'introduction des regroupements dans certains collèges est susceptible de modifier cette trajectoire.

Par construction, les scores des élèves scolarisés en 5^e dans les établissements qui n'ont pas mis en place de regroupements d'élèves, qui composent le groupe témoin, sont standardisés avec une moyenne de 0 chaque année et un écart-type de 1 (voir [Annexe 6](#)). Pour chaque niveau, la valeur des scores moyens des élèves scolarisés dans les établissements ayant mis en œuvre au moins une forme de regroupement représente l'écart entre les résultats moyens de ces élèves et les résultats des élèves scolarisés dans les établissements témoins. Il s'agit ici d'un écart relatif : par exemple, l'absence de variation du score moyen des élèves d'un établissement traité d'un niveau à l'autre peut cacher une progression de tous les élèves ou au contraire une dégradation de leurs performances scolaires.

Les évolutions des scores moyens dans les différents domaines et disciplines observés depuis le CP jusqu'à la date de début des interventions présentent quelques variations entre les élèves scolarisés dans les établissements traités (groupes hétérogènes et groupes constitués selon le profil scolaire des élèves) et dans les établissements témoins ([Figures 20 & 21](#)). En français, entre le CP et la 6^e, les scores des élèves exposés aux regroupements selon le profil scolaire en 5^e se sont rapprochés des scores de ceux qui ne connaissent aucune forme de regroupement, avec une différence de trajectoire importante entre l'entrée en CP et le point d'étape de mi-CP. Cet écart de trajectoire passée se retrouve également entre les élèves exposés aux groupes hétérogènes et les élèves dont l'établissement de 5^e n'a mis en place aucune forme de regroupement. Enfin, les scores de fluence ne suivent aucunement des trajectoires similaires entre les élèves qui connaissent des regroupements en 5^e et ceux qui n'en connaissent pas. En mathématiques, au niveau de la discipline dans son ensemble, une amélioration globale des scores moyens par rapport à ceux du groupe témoin (ou au contraire une moindre diminution des scores moyens) a pu être observée entre le CP et la 6^e pour les élèves des établissements qui mettent en œuvre des regroupements hétérogènes ou selon le profil scolaire des élèves. À nouveau, les élèves qui effectuent leur 5^e au sein d'établissements avec différentes formes de regroupements, hétérogènes ou constitués selon le profil scolaire des élèves, sont ceux qui ont suivi les trajectoires les plus similaires du CP à la 4^e.

FIGURE 20 • Évolution des scores moyens par domaine aux évaluations nationales de français et mathématiques en CP, mi-CP, CE1, 6^e et 4^e selon le type de regroupements mis en place dans l'établissement fréquenté en classe de 5^e



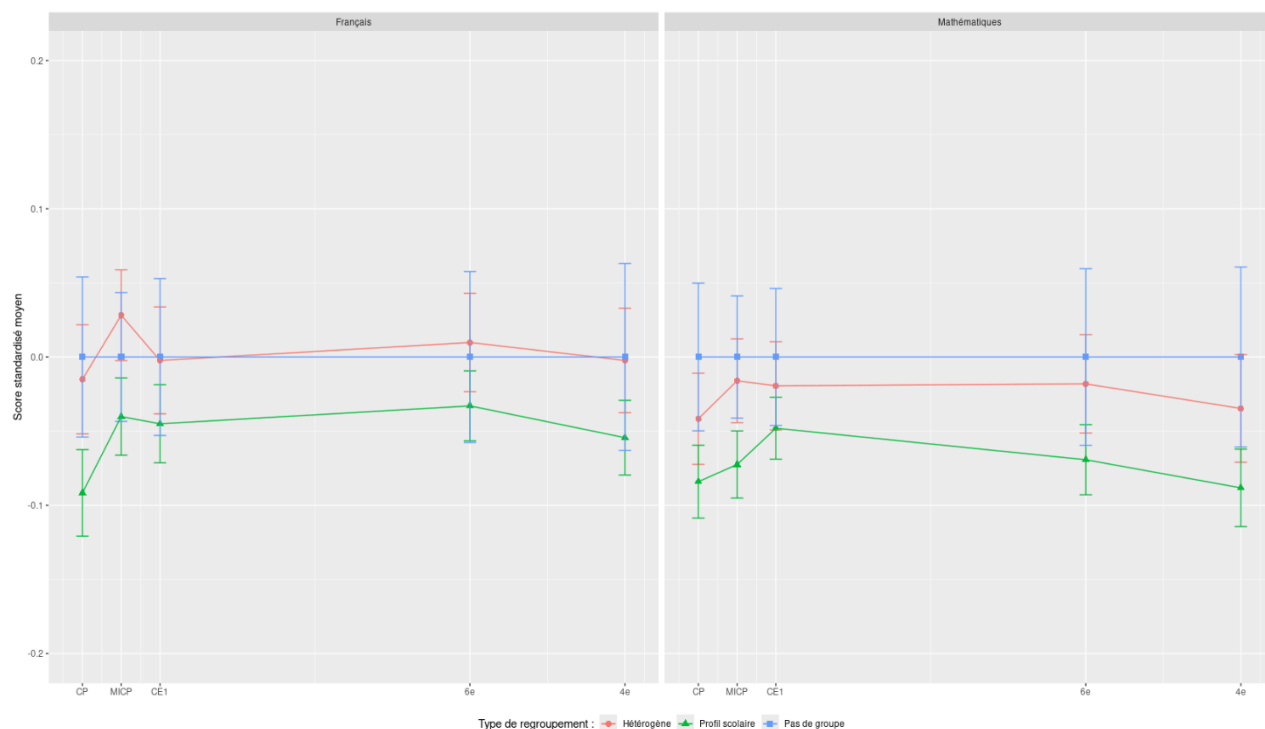
Lecture : En CP, les élèves qui sont scolarisés en 5^e au sein d'un établissement qui met en place des regroupements d'élèves selon leur profil scolaire ont des scores moyens de compréhension de l'écrit inférieurs de 0,06 écart-type à ceux des élèves qui sont scolarisés en 5^e au sein d'un établissement qui ne met en place aucune forme de regroupement.

Note : Pour chaque niveau, les scores sont standardisés par rapport aux scores des élèves scolarisés en 5^e dans les établissements qui ne mettent en place aucune forme de regroupement au cours de l'année scolaire 2024-2025. Les intervalles de confiance à 95 % sont affichés sur le graphique.

Champ : Éléves scolarisés en 5^e au sein des établissements de l'enquête EEC ayant adopté une forme unique de regroupement ou aucune en 5^e pour la discipline. Cela correspond à 1 289 collèges pour le français et 1 291 collèges pour les mathématiques. Pour chaque niveau et pour chaque domaine, le score par domaine est calculé uniquement pour les élèves ayant un score dans chacune des compétences retenues pour ce domaine.

Sources : DEPP, Évaluations nationales en français et mathématiques de CP (2018), de mi-CP (2019), de CE1 (2019), de 6^e (2023) et de 4^e (2025) ; enquête EEC, questionnaires Chefs d'établissement, printemps 2025 ; Constat de rentrée 2025.

FIGURE 21 • Évolution des scores moyens aux évaluations nationales de français et de mathématiques en CP, mi-CP, CE1, 6^e et 4^e selon le type de regroupements mis en place dans l'établissement fréquenté en classe de 5^e



Lecture : En CP, les élèves qui sont scolarisés en 5^e au sein d'un établissement qui met en place des regroupements d'élèves selon leur profil scolaire ont des scores moyens de français inférieurs de 0,09 écart-type à ceux des élèves qui sont scolarisés en 5^e au sein d'un établissement qui ne met en place aucune forme de regroupement.

Note : Pour chaque niveau, les scores sont standardisés par rapport aux scores des élèves scolarisés en 5^e dans les établissements qui ne mettent en place aucune forme de regroupement au cours de l'année scolaire 2024-2025. Les intervalles de confiance à 95 % sont affichés sur le graphique.

Champ : Élèves scolarisés en 5^e au sein des établissements de l'enquête EEC ayant adopté une forme unique de regroupement ou aucune en 5^e pour la discipline. Cela correspond à 1 289 collèges pour le français et 1 291 collèges pour les mathématiques. Pour chaque niveau et pour chaque discipline, le score par discipline est calculé uniquement pour les élèves ayant un score dans chacun des domaines retenus pour cette discipline.

Sources : DEPP, Évaluations nationales en français et mathématiques de CP (2018), de mi-CP (2019), de CE1 (2019), de 6^e (2023) et de 4^e (2025) ; enquête EEC, questionnaires Chefs d'établissement, printemps 2025 ; Constat de rentrée 2025.

À partir de l'évolution des moyennes, il est difficile de mesurer la significativité des différences afin de conclure sur la validité ou non de l'hypothèse des tendances parallèles. Les analyses événementielles réalisées à partir des modèles avec effets fixes permettent de mesurer plus directement la significativité des écarts de trajectoire entre les élèves avant leur année de 5^e, au cours de laquelle ils ont été exposés, ou non, à différentes formes de regroupement.

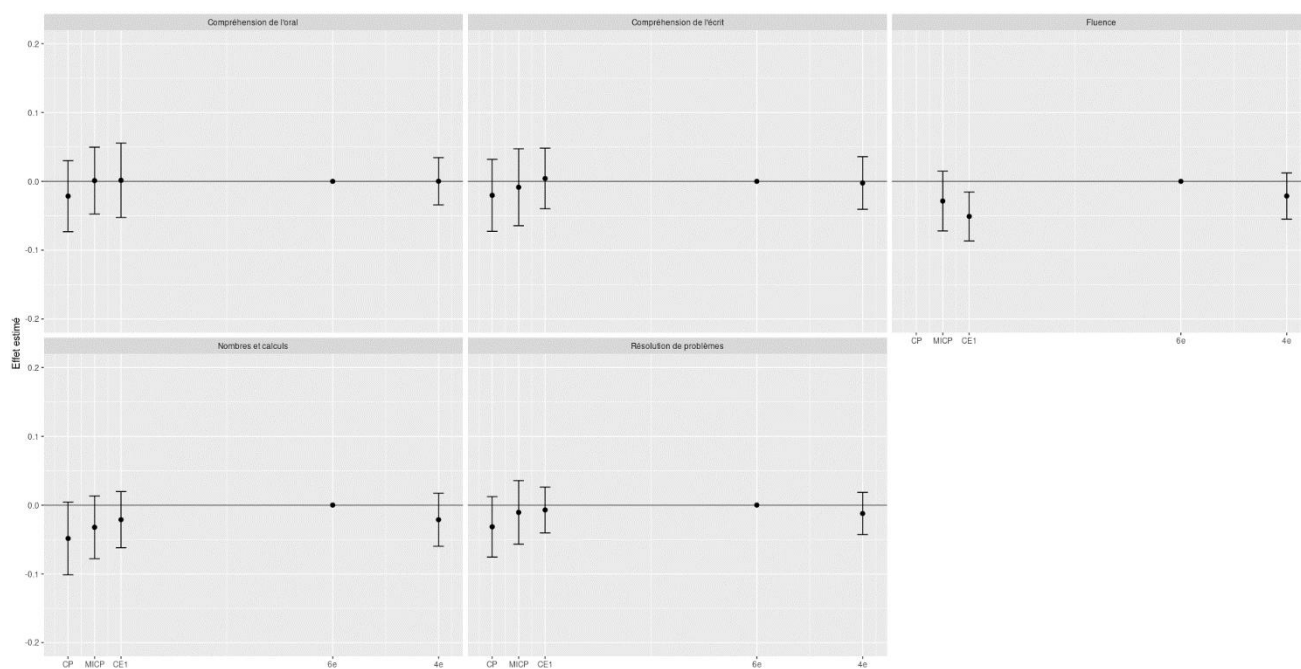
Analyse événementielle

Le modèle de différences de différences incluant des effets fixes permet de contrôler des écarts moyens de performance entre élèves qui restent fixes dans le temps et des chocs communs qui influencent l'ensemble des élèves pour un niveau donné (**Équations 3 & 6**). Cette modélisation permet ainsi de neutraliser certains facteurs qui peuvent exacerber la divergence des trajectoires d'élèves observées dans l'évolution, niveau après niveau, des scores standardisés moyens selon les organisations pédagogiques retenues en classe de 5^e. Par exemple, si les élèves progressent plus vite lorsqu'ils viennent d'un milieu familial plus favorisé, et que ces mêmes élèves se retrouvent davantage dans un type de regroupement que dans un autre, les graphiques précédents vont intégrer cette dynamique dans les évolutions présentées. En revanche, si ces éléments sont pris en compte dans la modélisation, à travers les effets fixes, alors l'analyse se concentre sur les évolutions individuelles indépendamment de toutes les différences initiales, et il est alors possible d'étudier de façon isolée les véritables divergences de trajectoires libérées des effets de structures.

Lorsque les effets fixes des élèves et des niveaux sont inclus dans les équations d'estimation pour l'analyse événementielle (**Équations 4 & 7**, et **Figures 22 à 27**), la plupart des différences observées en

français, en amont du traitement, entre les élèves exposés aux différents types de regroupements et les autres n'apparaissent plus significatives, à part pour la fluence mais il s'agit d'un décalage de faible ampleur et isolé. On détecte aussi une légère amélioration des scores de mathématiques entre le CP et la 6^e pour les élèves qui rejoindront les établissements qui mettent en place des regroupements hétérogènes ou selon le profil scolaire des élèves, relativement aux établissements témoins. Ces écarts sont ponctuellement significatifs. En revanche, on ne perçoit pas de différences dans les dynamiques quand on compare les établissements qui ont mis en œuvre l'un ou l'autre des regroupements.

FIGURE 22 • Différence de scores par domaine aux évaluations nationales avant et après la mise en place de la réforme (comparaison groupes hétérogènes/pas de regroupement)



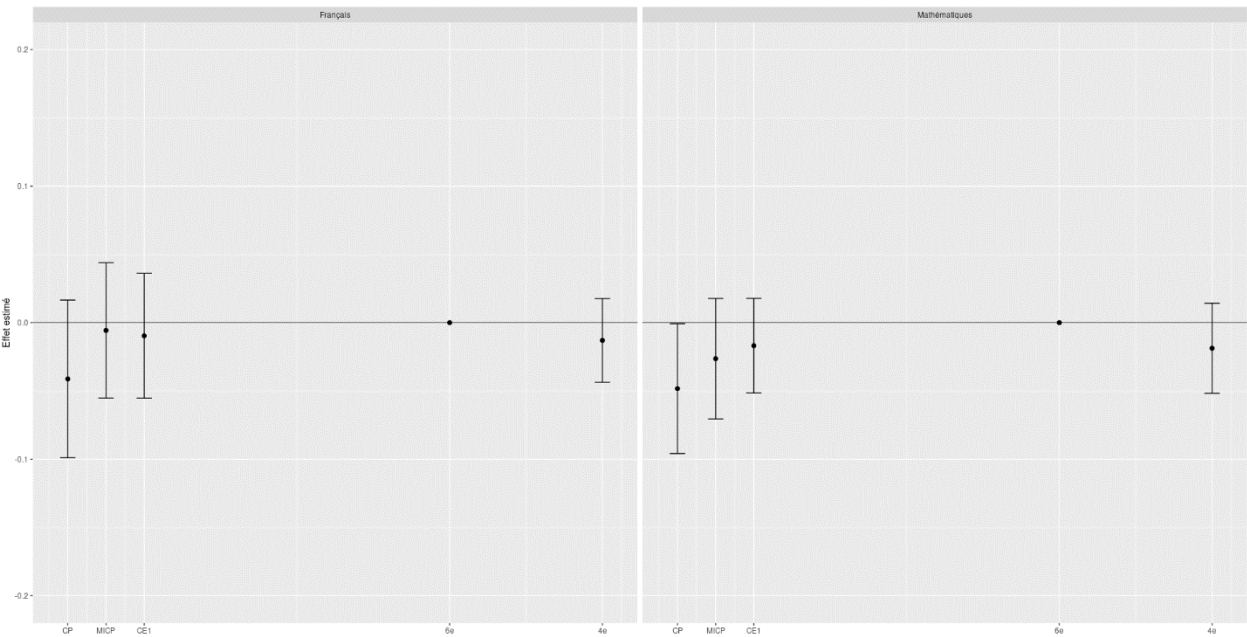
Lecture : Le coefficient de l'analyse événementielle est égal à -0,05 pour le domaine « Nombres et calculs » en CP, par rapport à la 6^e qui est le niveau de référence. Cela signifie qu'entre le CP et la 6^e, le score moyen dans ce domaine a davantage progressé (ou moins diminué) pour les élèves scolarisés en 5^e au sein d'établissements qui mettent en place des groupes hétérogènes que pour les élèves des établissements sans regroupements. En 4^e, le coefficient associé à ce domaine est égal à -0,02 : cela signifie que le score des élèves scolarisés en 5^e au sein d'établissements qui mettent en place des groupes hétérogènes a diminué (ou a moins augmenté) par rapport à celui des élèves des établissements sans regroupements. Aucun de ces deux coefficients n'est cependant significatif au seuil de 5 %.

Note : Pour chaque niveau, les scores sont standardisés par rapport aux scores des élèves scolarisés en 5^e dans les établissements qui ne mettent en place aucune forme de regroupement au cours de l'année scolaire 2024-2025. Les erreurs-types sont clusterisées au niveau de l'établissement fréquenté en 5^e. Les intervalles de confiance à 95 % sont affichés sur le graphique. La 6^e est le niveau de référence.

Champ : Élèves scolarisés en 5^e au sein des établissements de l'enquête EEC ayant mis en place des regroupements hétérogènes ou aucune forme de regroupement en 5^e pour la discipline. Cela correspond à 484 collèges pour le français et 473 collèges pour les mathématiques. Pour chaque niveau et pour chaque domaine, le score par domaine est calculé uniquement pour les élèves ayant un score dans chacune des compétences retenues pour ce domaine.

Sources : DEPP, Évaluations nationales en français et mathématiques de CP (2018), de mi-CP (2019), de CE1 (2019), de 6^e (2023) et de 4^e (2025) ; enquête EEC, questionnaires Chefs d'établissement, printemps 2025 ; Constat de rentrée 2025.

FIGURE 23 • Différence de scores aux évaluations nationales avant et après la mise en place de la réforme (comparaison groupes hétérogènes/pas de regroupement)



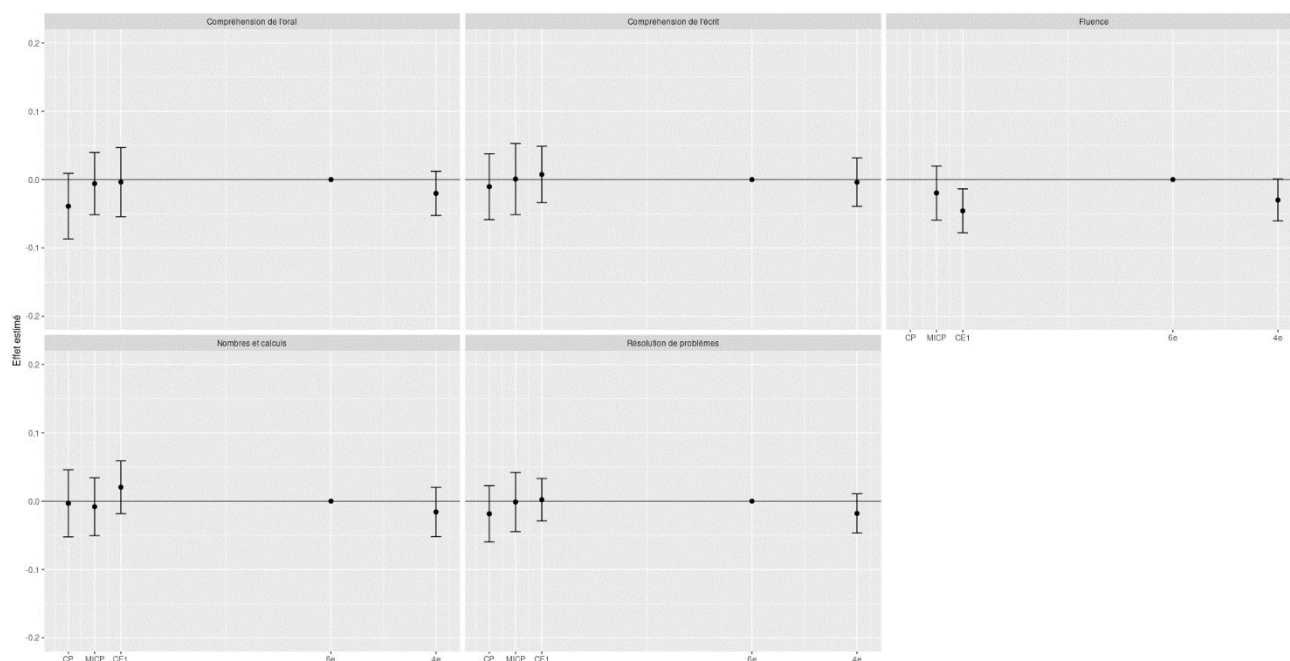
Lecture : Le coefficient de l’analyse événementielle est égal à -0,04 pour le français en CP, par rapport à la 6^e qui est le niveau de référence. Cela signifie qu’entre le CP et la 6^e, le score moyen dans cette discipline a davantage progressé (ou moins diminué) pour les élèves scolarisés en 5^e au sein d’établissements qui mettent en place des groupes hétérogènes que pour les élèves des établissements sans regroupements. En 4^e, le coefficient associé à cette discipline est égal à -0,02 : cela signifie que le score des élèves scolarisés en 5^e au sein d’établissements qui mettent en place des groupes hétérogènes a diminué (ou a moins augmenté) par rapport à celui des élèves des établissements sans regroupements. Aucun de ces deux coefficients n’est cependant significatif au seuil de 5 %.

Note : Pour chaque niveau, les scores sont standardisés par rapport aux scores des élèves scolarisés en 5^e dans les établissements qui ne mettent en place aucune forme de regroupement au cours de l’année scolaire 2024-2025. Les erreurs-types sont clusterisées au niveau de l’établissement fréquenté en 5^e. Les intervalles de confiance à 95 % sont affichés sur le graphique. La 6^e est le niveau de référence.

Champ : Élèves scolarisés en 5^e au sein des établissements de l’enquête EEC ayant mis en place des regroupements hétérogènes ou aucune forme de regroupement en 5^e pour la discipline. Cela correspond à 484 collèges pour le français et 473 collèges pour les mathématiques. Pour chaque niveau et pour chaque discipline, le score par discipline est calculé uniquement pour les élèves ayant un score dans chacun des domaines retenus pour cette discipline.

Sources : DEPP, Évaluations nationales en français et mathématiques de CP (2018), de mi-CP (2019), de CE1 (2019), de 6^e (2023) et de 4^e (2025) ; enquête EEC, questionnaires Chefs d’établissement, printemps 2025 ; Constat de rentrée 2025.

FIGURE 24 • Différence de scores par domaine aux évaluations nationales avant et après la mise en place de la réforme (comparaison groupes composés selon le profil scolaire des élèves/pas de regroupement)



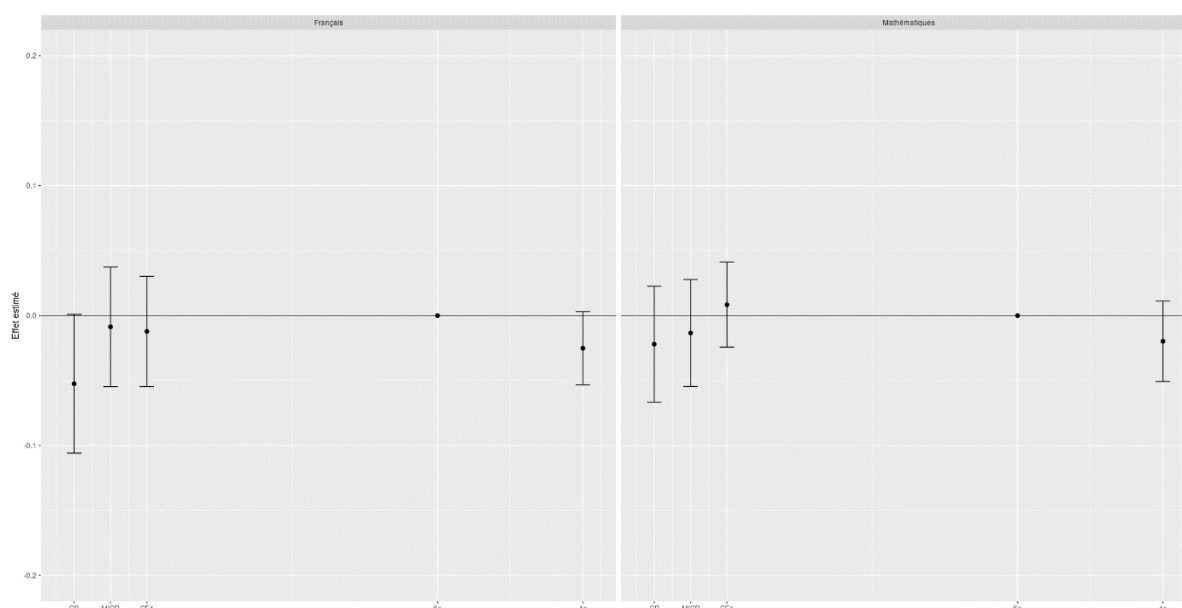
Lecture : Le coefficient de l'analyse événementielle est égal à -0,04 pour le domaine « Compréhension de l'oral » en CP, par rapport à la 6^e qui est le niveau de référence. Cela signifie qu'entre le CP et la 6^e, le score moyen dans ce domaine a davantage progressé (ou moins diminué) pour les élèves scolarisés en 5^e au sein d'établissements qui mettent en place des groupes constitués selon le profil scolaire des élèves que pour les élèves des établissements sans regroupements. En 4^e, le coefficient associé à ce domaine est égal à -0,02 : cela signifie que le score des élèves scolarisés en 5^e au sein d'établissements qui mettent en place des groupes constitués selon le profil scolaire des élèves a diminué (ou a moins augmenté) par rapport à celui des élèves des établissements sans regroupements. Aucun de ces deux coefficients n'est cependant significatif au seuil de 5 %.

Note : Pour chaque niveau, les scores sont standardisés par rapport aux scores des élèves scolarisés en 5^e dans les établissements qui ne mettent en place aucune forme de regroupement au cours de l'année scolaire 2024-2025. Les erreurs-types sont clusterisées au niveau de l'établissement fréquenté en 5^e. Les intervalles de confiance à 95 % sont affichés sur le graphique. La 6^e est le niveau de référence.

Champ : Éléves scolarisés en 5^e au sein des établissements de l'enquête EEC ayant mis en place des groupes constitués selon le profil scolaire des élèves ou aucune forme de regroupement en 5^e pour la discipline. Cela correspond à 948 collèges pour le français et 958 collèges pour les mathématiques. Pour chaque niveau et pour chaque domaine, le score par domaine est calculé uniquement pour les élèves ayant un score dans chacune des compétences retenues pour ce domaine.

Sources : DEPP, Évaluations nationales en français et mathématiques de CP (2018), de mi-CP (2019), de CE1 (2019), de 6^e (2023) et de 4^e (2025) ; enquête EEC, questionnaires Chefs d'établissement, printemps 2025 ; Constat de rentrée 2025.

FIGURE 25 • Différence de scores aux évaluations nationales avant et après la mise en place de la réforme (comparaison groupes composés selon le profil scolaire des élèves/pas de regroupement)



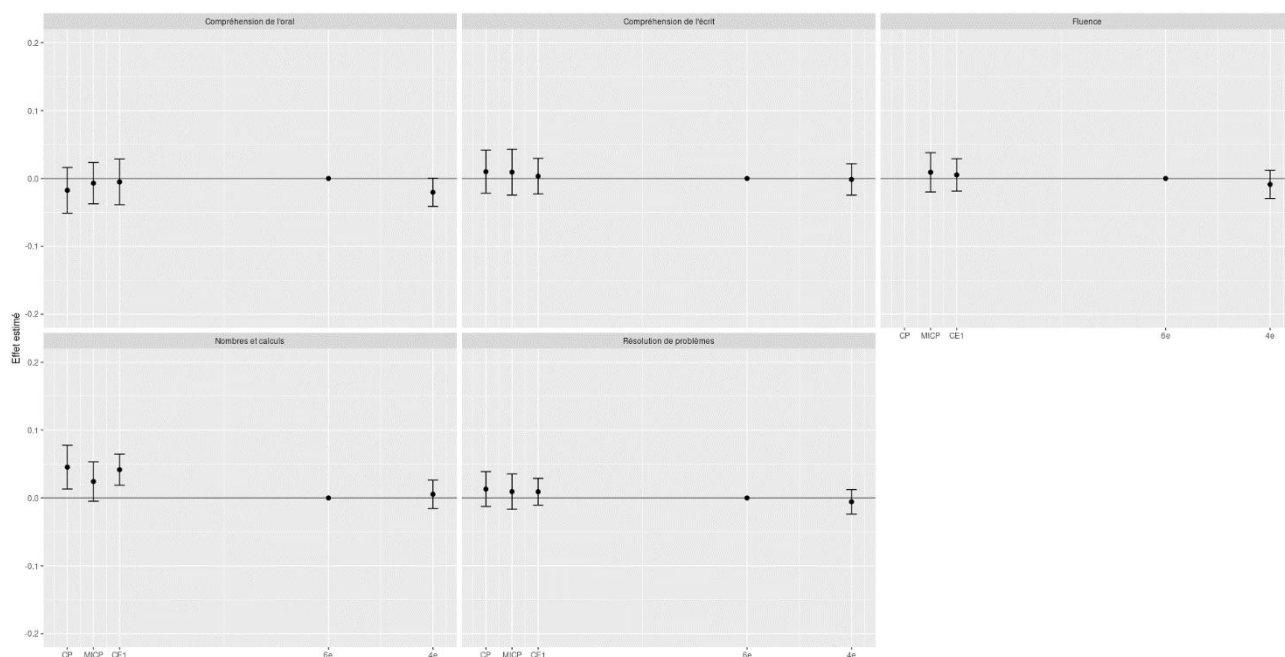
Lecture : Le coefficient de l'analyse événementielle est égal à -0,05 pour le français en CP, par rapport à la 6^e qui est le niveau de référence. Cela signifie qu'entre le CP et la 6^e, le score moyen dans cette discipline a davantage progressé (ou moins diminué) pour les élèves scolarisés en 5^e au sein d'établissements qui mettent en place des groupes constitués selon le profil scolaire des élèves que pour les élèves des établissements sans regroupements. En 4^e, le coefficient associé à cette discipline est égal à -0,03 : cela signifie que le score des élèves scolarisés en 5^e au sein d'établissements qui mettent en place des groupes constitués selon le profil scolaire des élèves a diminué (ou a moins augmenté) par rapport à celui des élèves des établissements sans regroupements. Aucun de ces deux coefficients n'est cependant significatif au seuil de 5 %.

Note : Pour chaque niveau, les scores sont standardisés par rapport aux scores des élèves scolarisés en 5^e dans les établissements qui ne mettent en place aucune forme de regroupement au cours de l'année scolaire 2024-2025. Les erreurs-types sont clusterisées au niveau de l'établissement fréquenté en 5^e. Les intervalles de confiance à 95 % sont affichés sur le graphique. La 6^e est le niveau de référence.

Champ : Élèves scolarisés en 5^e au sein des établissements de l'enquête EEC ayant mis en place des groupes constitués selon le profil scolaire des élèves ou aucune forme de regroupement en 5^e pour la discipline. Cela correspond à 948 collèges pour le français et 958 collèges pour les mathématiques. Pour chaque niveau et pour chaque discipline, le score par discipline est calculé uniquement pour les élèves ayant un score dans chacun des domaines retenus pour cette discipline.

Sources : DEPP, Évaluations nationales en français et mathématiques de CP (2018), de mi-CP (2019), de CE1 (2019), de 6^e (2023) et de 4^e (2025) ; enquête EEC, questionnaires Chefs d'établissement, printemps 2025 ; Constat de rentrée 2025.

FIGURE 26 • Différence de scores par domaine aux évaluations nationales avant et après la mise en place de la réforme (comparaison groupes composés selon le profil scolaire des élèves/groupes hétérogènes)



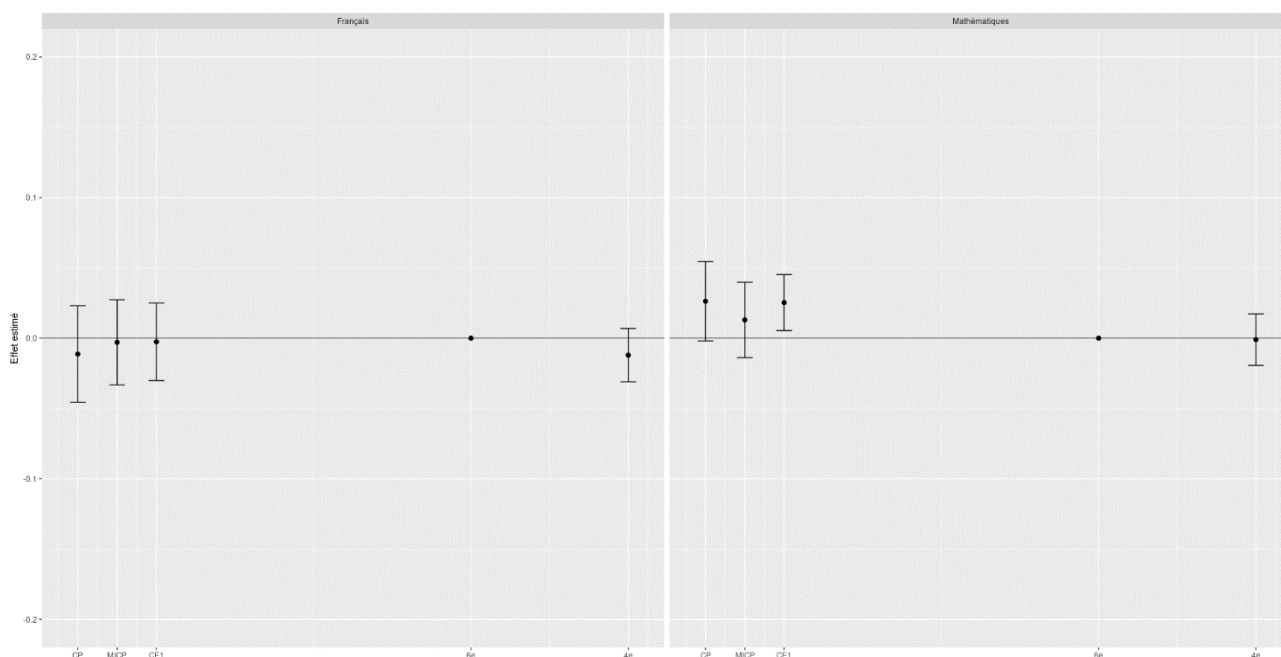
Lecture : Le coefficient de l'analyse événementielle est égal à -0,02 pour le domaine « Compréhension de l'oral » en CP, par rapport à la 6^e qui est le niveau de référence. Cela signifie qu'entre le CP et la 6^e, le score moyen dans ce domaine a davantage progressé (ou moins diminué) pour les élèves scolarisés en 5^e au sein d'établissements qui mettent en place des groupes constitués selon le profil scolaire des élèves que pour les élèves des établissements qui instaurent des regroupements hétérogènes. En 4^e, le coefficient associé à ce domaine est égal à -0,02 : cela signifie que le score des élèves scolarisés en 5^e au sein d'établissements qui mettent en place des groupes constitués selon le profil scolaire des élèves a diminué (ou a moins augmenté) par rapport à celui des élèves des établissements avec des regroupements hétérogènes. Aucun de ces deux coefficients n'est cependant significatif au seuil de 5 %.

Note : Pour chaque niveau, les scores sont standardisés par rapport aux scores des élèves scolarisés en 5^e dans les établissements qui ne mettent en place aucune forme de regroupement au cours de l'année scolaire 2024-2025. Les erreurs-types sont clusterisées au niveau de l'établissement fréquenté en 5^e. Les intervalles de confiance à 95 % sont affichés sur le graphique. La 6^e est le niveau de référence.

Champ : Éléves scolarisés en 5^e au sein des établissements de l'enquête EEC ayant mis en place des groupes constitués selon le profil scolaire des élèves ou des regroupements hétérogènes en 5^e pour la discipline. Cela correspond à 1 146 collèges pour le français et 1 151 collèges pour les mathématiques. Pour chaque niveau et pour chaque domaine, le score par domaine est calculé uniquement pour les élèves ayant un score dans chacune des compétences retenues pour ce domaine.

Sources : DEPP, Évaluations nationales en français et mathématiques de CP (2018), de mi-CP (2019), de CE1 (2019), de 6^e (2023) et de 4^e (2025) ; enquête EEC, questionnaires Chefs d'établissement, printemps 2025 ; Constat de rentrée 2025.

FIGURE 27 • Différence de scores aux évaluations nationales avant et après la mise en place de la réforme (comparaison groupes composés selon le profil scolaire des élèves/groupes hétérogènes)



Lecture : Le coefficient de l'analyse événementielle est égal à -0,01 pour le français en CP, par rapport à la 6^e qui est le niveau de référence. Cela signifie qu'entre le CP et la 6^e, le score moyen dans cette discipline a davantage progressé (ou moins diminué) pour les élèves scolarisés en 5^e au sein d'établissements qui mettent en place des groupes constitués selon le profil scolaire des élèves que pour les élèves des établissements qui instaurent des regroupements hétérogènes. En 4^e, le coefficient associé à cette discipline est égal à -0,01 : cela signifie que le score des élèves scolarisés en 5^e au sein d'établissements qui mettent en place des groupes constitués selon le profil scolaire des élèves a diminué (ou a moins augmenté) par rapport à celui des élèves des établissements avec des regroupements hétérogènes. Aucun de ces deux coefficients n'est cependant significatif au seuil de 5 %.

Note : Pour chaque niveau, les scores sont standardisés par rapport aux scores des élèves scolarisés en 5^e dans les établissements qui ne mettent en place aucune forme de regroupement au cours de l'année scolaire 2024-2025. Les erreurs-types sont clusterisées au niveau de l'établissement fréquenté en 5^e. Les intervalles de confiance à 95 % sont affichés sur le graphique. La 6^e est le niveau de référence.

Champ : Élèves scolarisés en 5^e au sein des établissements de l'enquête EEC ayant mis en place des groupes constitués selon le profil scolaire des élèves ou des regroupements hétérogènes en 5^e pour la discipline. Cela correspond à 1 146 collèges pour le français et 1 151 collèges pour les mathématiques. Pour chaque niveau et pour chaque discipline, le score par discipline est calculé uniquement pour les élèves ayant un score dans chacun des domaines retenus pour cette discipline.

Sources : DEPP, Évaluations nationales en français et mathématiques de CP (2018), de mi-CP (2019), de CE1 (2019), de 6^e (2023) et de 4^e (2025) ; enquête EEC, questionnaires Chefs d'établissement, printemps 2025 ; Constat de rentrée 2025.

L'absence d'effet moyen des nouvelles organisations pédagogiques sur la réussite des élèves confirmée par la seconde approche

Dans l'ensemble, l'exposition à l'une des formes de regroupement d'élèves ne semble pas, en moyenne, avoir d'impact sur leur réussite. Les coefficients associés aux effets des regroupements selon le profil scolaire des élèves sont presque tous négatifs dans les deux modèles de régression (**Tableaux 10 & 11**, et **Annexe 7**), mais ces effets sont d'une faible ampleur (inférieure à 2 % d'un écart-type en valeur absolue), et aucun coefficient n'est significatif au seuil 5 % : en d'autres termes, les données disponibles ne permettent pas de conclure avec un niveau de confiance suffisant à l'existence d'un quelconque effet du type de regroupement. À nouveau, la puissance statistique du dispositif d'évaluation est bonne puisque des effets positifs ou négatifs petits (3 à 4 % d'un écart-type) pourraient être détectés s'ils étaient présents. Ces observations sont valables pour les deux disciplines et pour les domaines qui les composent (**Annexe 8**). Pour les regroupements hétérogènes, les coefficients du modèle à effets fixes sont presque tous égaux à 0, tandis que le modèle de doubles différences sans effet fixe et mobilisant uniquement les données de 6^e et de 4^e produit quelques coefficients négatifs de faible ampleur et non significatifs : les regroupements hétérogènes n'ont donc pas eu d'effet sur la réussite des élèves en comparaison avec l'absence de regroupement. Enfin, la comparaison des élèves exposés aux groupes hétérogènes et de ceux concernés par les regroupements selon leur profil scolaire fait de nouveau apparaître une absence d'effet statistiquement significatif.

Ainsi, et en cohérence avec les analyses menées dans la partie précédente, ces premières estimations ne permettent pas d'affirmer que le regroupement des élèves selon leur profil scolaire en classe de 5^e favorise ou dégrade les apprentissages, ni que la constitution de groupes hétérogènes (qui

s’accompagnait possiblement d’une légère réduction de la tailles des groupes, quoique cette hypothèse reste à confirmer) a une influence sur la réussite scolaire des élèves, mesurée au regard des résultats aux évaluations nationales de français et de mathématiques.

TABLEAU 10 • Coefficients de l’estimation des effets des organisations pédagogiques sur les scores aux évaluations nationales de français de 4^e

Comparaison	Français					
	Hétérogène vs pas de regroupement	Hétérogène vs pas de regroupement	Selon le profil scolaire des élèves vs pas de regroupement	Selon le profil scolaire des élèves vs pas de regroupement	Selon le profil scolaire des élèves vs hétérogène	Selon le profil scolaire des élèves vs hétérogène
Modèle	DiD	TWFE	DiD	TWFE	DiD	TWFE
Coefficient	-0,012 (0,018)	0 (0,02)	-0,012 (0,018)	-0,008 (0,018)	-0,009 (0,012)	-0,008 (0,012)
Crédibilité des tendances parallèles	Oui	Oui	Oui	Oui	Oui	Oui
Effets fixes	Non	Id élève + niveau	Non	Id élève + niveau	Non	Id élève + niveau
Cluster	UAI en 5 ^e	UAI en 5 ^e	UAI en 5 ^e	UAI en 5 ^e	UAI en 5 ^e	UAI en 5 ^e
N obs.	91 296	217 422	186 440	448 801	228 156	548 607
N élèves	53 111	55 836	108 884	114 641	133 047	140 217
N établissements	484	484	948	948	1 146	1 146

Lecture : Pour le français, le modèle de doubles différences sans effet fixe estime un coefficient de -0,012 pour l’effet des regroupements hétérogènes : le fait d’appartenir à la cohorte d’élèves qui entre en 5^e en 2024-2025 et d’avoir été scolarisé dans un établissement qui met en place des groupes hétérogènes plutôt que dans un établissement qui ne met en place aucune forme de regroupement diminue en moyenne le score en français de 0,012, soit 1,2 % d’un écart-type. Ce coefficient n’est cependant pas significatif.

Note : Les modèles DiD correspondent aux modèles de doubles différences classiques (Équation 5) tandis que les modèles TWFE désignent les modèles avec effets fixes (Équation 6). Les erreurs types clusterisées au niveau de l’établissement fréquenté en 5^e sont présentées entre parenthèses. Seuils de significativité : « *** », pour un coefficient significatif à 0,1 %, « ** », pour un coefficient significatif à 1 %, « * », pour un coefficient significatif à 5 % et « . » pour un coefficient significatif à 10 %.

Champ : Élèves scolarisés en 5^e au sein des établissements de l’enquête EEC ayant adopté une forme unique de regroupement ou aucune en 5^e pour le français. Cela correspond à 1 289 collèges. Pour chaque niveau, le score en français est calculé uniquement pour les élèves ayant un score dans chacun des domaines retenus pour cette discipline.

Sources : DEPP, Évaluations nationales en français de CP (2018), de mi-CP (2019), de CE1 (2019), de 6^e (2023) et de 4^e (2025) ; enquête EEC, questionnaires Chefs d’établissement, printemps 2025 ; Constat de rentrée 2025.

TABLEAU 11 • Coefficients de l'estimation des effets des organisations pédagogiques sur les scores aux évaluations nationales en mathématiques de 4^e

Comparaison	Mathématiques					
	Hétérogène vs pas de regroupement	Hétérogène vs pas de regroupement	Selon le profil scolaire des élèves vs pas de regroupement	Selon le profil scolaire des élèves vs pas de regroupement	Selon le profil scolaire des élèves vs hétérogène	Selon le profil scolaire des élèves vs hétérogène
Modèle	DiD	TWFE	DiD	TWFE	DiD	TWFE
Coefficient	-0,017 (0,016)	0,001 (0,018)	-0,017 (0,016)	-0,015 (0,017)	-0,002 (0,009)	-0,016 (0,011)
Crédibilité des tendances parallèles	Non, écart significatif en CP	Non, écart significatif en CP	Oui	Oui	Non, écart significatif en CE1	Non, écart significatif en CE1
Effets fixes	Non	Id élève + niveau	Non	Id élève + niveau	Non	Id élève + niveau
Cluster	UAI en 5 ^e	UAI en 5 ^e	UAI en 5 ^e	UAI en 5 ^e	UAI en 5 ^e	UAI en 5 ^e
N obs.	104 330	229 342	223 753	491 004	270 463	593 978
N élèves	54 534	54 846	117 189	117 830	141 675	142 458
N établissements	473	473	958	958	1 151	1 151

Lecture : Pour les mathématiques, le modèle de doubles différences sans effet fixe estime un coefficient de -0,017 pour l'effet des regroupements hétérogènes : le fait d'appartenir à la cohorte d'élèves qui entre en 5^e en 2024-2025 et d'avoir été scolarisé dans un établissement qui met en place des groupes hétérogènes plutôt que dans un établissement qui ne met en place aucune forme de regroupement diminue en moyenne le score en mathématiques de 0,017, soit 1,7 % d'un écart-type. Ce coefficient n'est cependant pas significatif.

Note : Les modèles DiD correspondent aux modèles de doubles différences classiques (Équation 5) tandis que les modèles TWFE désignent les modèles avec effets fixes (Équation 6). Les erreurs types clusterisées au niveau de l'établissement fréquenté en 5^e sont présentées entre parenthèses. Seuils de significativité : « *** », pour un coefficient significatif à 0,1 %, « ** », pour un coefficient significatif à 1 %, « * », pour un coefficient significatif à 5 % et « . » pour un coefficient significatif à 10 %.

Champ : Élèves scolarisés en 5^e au sein des établissements de l'enquête EEC ayant adopté une forme unique de regroupement ou aucune en 5^e pour les mathématiques. Cela correspond à 1 291 collèges. Pour chaque niveau, le score en mathématiques est calculé uniquement pour les élèves ayant un score dans chacun des domaines retenus pour cette discipline.

Sources : DEPP, Évaluations nationales en mathématiques de CP (2018), de mi-CP (2019), de CE1 (2019), de 6^e (2023) et de 4^e (2025) ; enquête EEC, questionnaires Chefs d'établissement, printemps 2025 ; Constat de rentrée 2025.

Comme pour la première approche de différences-de-différences, reposant sur la comparaison des cohortes d'élèves avant et après la réforme, une analyse d'hétérogénéité a été réalisée à partir de la seconde approche fondée sur le suivi des élèves au sein d'une cohorte, en fonction de leur sexe et de leur niveau de performance en 6^e (mesuré par tercile). Les résultats présentés en annexe ne révèlent aucun effet significatif selon ces deux dimensions (Annexe 7). Cela suggère que la réforme n'a pas eu d'impact différencié selon le profil des élèves : ni les niveaux les plus faibles, ni les plus forts, ni les filles, ni les garçons n'ont été affectés de manière distincte par rapport au reste de la population étudiée. Autrement dit, dans la limite de ce que permettent d'observer les données disponibles, la réforme n'a ni réduit, ni accentué les écarts de performance entre élèves de niveaux plus forts et ceux de niveaux plus faibles, ni entre filles et garçons.

Discussion des limites associées aux résultats

Si ces premiers résultats d'estimation suggèrent, pour les deux approches d'estimation, une absence d'effet des différentes organisations pédagogiques mises en œuvre à la rentrée 2024 sur les performances scolaires en français et en mathématiques, plusieurs éléments invitent à la prudence et appellent des analyses complémentaires. En premier lieu, l'analyse proposée ici ne concerne que la première année de mise en place de la réforme, qui s'est faite dans des conditions de définition et d'appropriation progressive de la mesure par les différents acteurs et au sein d'un contexte initialement marqué par de fortes incertitudes autour de la mise en place et du devenir de la réforme. L'effet est donc mesuré sur des élèves qui n'ont été exposés à la réforme qu'au cours d'une seule année scolaire, en classe de 5^e. En économie de l'éducation, il est souvent observé que l'effet d'une politique éducative est d'autant plus important que la politique est mise en place tôt dans la scolarité ou que les élèves y sont exposés pendant plusieurs années (Heckman, 2006). Ainsi, on ne peut pas extrapoler ces résultats d'évaluation à d'autres situations, par exemple les élèves de 6^e exposés aux

regroupements pendant deux années scolaires, en 2024-2025 en classe de 6^e et en 2025-2026 en classe de 5^e.

D'autre part, la méthode des différences de différences repose généralement sur l'hypothèse d'un « traitement » unique et bien défini. Dans le cadre de la politique étudiée ici, chaque établissement a pu mettre en œuvre des organisations pédagogiques spécifiques et ainsi définir son propre « traitement ». Dans les analyses qui précèdent, ce qui est mesuré sous les appellations de « groupes hétérogènes » ou de « groupes homogènes » peut ainsi recouvrir des réalités différentes d'un établissement à l'autre, susceptibles d'être associées à des effets eux-mêmes différents : l'absence d'impact moyen pourrait ainsi masquer des réalités locales variées. Cela est d'autant plus important que l'enquête relative à l'organisation des enseignements au collège s'appuie sur les réponses des acteurs interrogés, ce qui les rend sujettes à des erreurs de déclaration ou de mesure, par exemple si les chefs d'établissement ou les coordonnateurs ne se retrouvent pas tout à fait dans les modalités de réponse proposées dans le questionnaire. Une même approche méthodologique fondée sur des mesures objectivées de l'hétérogénéité au sein des groupes ou des classes pourrait ainsi aboutir à des conclusions différentes en limitant ces biais d'atténuation.

Par ailleurs, la mise en place de nouvelles organisations pédagogiques est ici définie à l'échelle des établissements, sans prendre en compte les groupes d'appartenance des élèves, comme si tous avaient vécu la même expérience de la réforme. Les effets présentés sont des moyennes qui pourraient masquer des effets différenciés en fonction du groupe d'appartenance des élèves. D'autres travaux pourront se pencher sur ces dimensions afin de documenter plus précisément l'effet des regroupements d'élèves sur leurs performances scolaires, tout en tenant compte d'éventuels effets différenciés des nouvelles organisations pédagogiques selon le niveau initial des élèves, leur genre ou leur milieu social.

En outre, l'hypothèse des tendances parallèles suppose, pour être valide, qu'aucun événement simultané ne produise des effets susceptibles de se confondre avec ceux de la réforme étudiée. Dans le cas de la réforme des organisations pédagogiques, une forme de traitement non observée pourrait exister si certains établissements, ne se reconnaissant pas dans les catégories proposées dans le questionnaire, ont profité de cette période de réforme pour développer d'autres dispositifs, avec ou sans moyens supplémentaires. Il pourrait s'agir, par exemple, de la généralisation de « Devoirs faits » dans l'établissement, du recours à des demi-groupes dans différentes matières, du maintien d'options attractives ou encore du développement de pratiques de co-intervention. De plus, dans les commentaires libres recueillis dans le cadre de l'enquête menée pour documenter la mise en œuvre de la politique, les chefs d'établissement et les coordonnateurs soulignent l'importance des absences non remplacées, qui ont entravé le déploiement du dispositif tout au long de l'année scolaire. Les nouvelles organisations pédagogiques ont fortement contribué à réduire le nombre de professeurs disponibles pour assurer les remplacements, ce qui a pu affecter le nombre d'heures effectivement réalisées en groupe ainsi que le nombre d'heures d'enseignement effectives pour les élèves, par rapport aux années précédentes. Cette dimension n'est toutefois pas mesurée dans cette étude.

Enfin, la première approche de doubles différences qui repose sur la différence entre les cohortes successives d'élèves est ici limitée par la faible profondeur temporelle des données disponibles. En effet, la crédibilité de l'hypothèse de tendances parallèles après la mise en œuvre de la politique repose en grande partie sur l'observation de tendances comparables au cours de la période qui l'a précédée. Or cette période se limite ici à deux années, 2023 et 2024, ce qui est insuffisant pour apprécier pleinement la validité de cette hypothèse. De plus, sur les deux années disponibles, l'hypothèse de tendances parallèles ne semble pas vérifiée en français pour la comparaison entre les regroupements opérés selon le profil scolaire des élèves et l'absence de regroupements. Ce manque de profondeur temporelle tient principalement au fait que les évaluations de 4^e n'ont été généralisées que récemment. D'autres indicateurs du parcours scolaire des élèves pourraient à

l'avenir être mobilisés afin d'apprécier les effets de plus long terme de cette réforme, comme par exemple les choix d'orientation après la 3^e ou encore les résultats obtenus au DNB. D'autres méthodes, appliquées à l'échelle des établissements, des groupes ou des élèves pourraient en outre être mobilisées afin de tenter d'identifier autrement l'effet de cette réforme.

➤ Conclusion

Cette étude constitue la première évaluation quantitative des organisations pédagogiques mises en place dans le cadre de la réforme des « groupes de besoins » en français et en mathématiques dans les classes de 6^e et de 5^e à partir de la rentrée de septembre 2024. **Elle dresse un constat contrasté à l'issue de la première année de mise en œuvre du dispositif : les établissements ont largement mis en place les regroupements d'élèves, mais selon des modalités très diverses.** Les réponses au questionnaire auto-administré collectées lors de la deuxième vague d'enquête au printemps 2026 seront prochainement analysées afin de décrire l'évolution de la mise en œuvre des regroupements ainsi que les perceptions des différents acteurs interrogés. L'étude, qui couvre deux années consécutives correspondant aux deux premières années de déploiement du dispositif, permettra d'apprécier son évolution et d'évaluer ses effets de manière plus globale.

Les premières analyses causales de l'effet de ces organisations pédagogiques sur les apprentissages des élèves concernés par la réforme en 5^e en 2024-2025 ne révèlent aucun effet statistiquement significatif sur les résultats moyens aux évaluations nationales de français et de mathématiques à l'entrée en 4^e, que l'on compare les établissements ayant mis en place des groupes hétérogènes, des groupes constitués selon le profil scolaire des élèves, ou n'ayant mis en place aucune nouvelle organisation. **Ce résultat,** obtenu à partir de deux approches complémentaires en doubles différences, **rejoint les conclusions de la littérature scientifique internationale la plus récente sur le regroupement des élèves, qui ne met pas en évidence d'effet moyen de ce type de politique sur la réussite scolaire** (Terrin et Triventi, 2023). Aucune hétérogénéité significative n'a par ailleurs pu être mise en évidence selon le niveau scolaire initial des élèves en 6^e ou selon leur sexe, ce qui ne permet pas de conclure à des effets différenciés du dispositif.

Plusieurs éléments invitent à considérer ces résultats avec une certaine prudence. D'une part, la cohorte d'élèves étudiée dans les analyses d'impact n'a été concernée par les nouvelles organisations pédagogiques que pendant une seule année scolaire, celle de leur scolarisation en 5^e en 2024-2025. Or les effets des politiques éducatives sont souvent d'autant plus marqués que l'exposition des élèves est précoce et prolongée. D'autre part, l'indicateur mobilisé pour estimer leur progression repose sur les évaluations nationales à l'entrée en 4^e. Celles-ci n'ayant été généralisées que récemment, l'analyse ne dispose que d'une profondeur temporelle limitée, ce qui fragilise la validité d'une des méthodes utilisées.

Ensuite, comme décrit en détail dans la première partie de ce document, l'importante hétérogénéité qui caractérise la mise en œuvre du dispositif a pu peser sur les analyses. Le contexte de mise en œuvre a en effet été marqué par de fortes incertitudes, se traduisant par des pratiques de regroupements très variables d'un établissement à l'autre et par des ajustements organisationnels progressifs au cours de l'année. Enfin, il est important de souligner que les effets estimés se concentrent sur les résultats moyens des élèves aux évaluations nationales et donc sur l'évolution des apprentissages. D'éventuels effets pourraient toutefois exister sur des dimensions socio-émotionnelles, pour lesquelles aucun indicateur disponible ne permet de réaliser des estimations comparables à celles présentées dans ce document.

Le prolongement de ces premières analyses selon plusieurs axes permettra d'affiner et de consolider l'estimation des éventuels effets du dispositif. L'extension de l'analyse aux autres cohortes concernées par la réforme, notamment celle entrée au collège en 2024-2025, concernée par les regroupements pendant deux années consécutives en 6^e et 5^e, et celle entrée en 2025-2026 exposée aux regroupements uniquement en 6^e, permettra d'augmenter la précision des résultats et d'estimer de potentiels effets différenciés selon la durée d'exposition des élèves et le moment auquel cette exposition intervient. L'exploitation des résultats des élèves au DNB apportera quant à elle, outre une profondeur temporelle nettement plus importante, la possibilité de mesurer des effets de plus long terme sur les résultats des élèves, une fois le retour en classe entière effectué en 4^e et 3^e.

Les enquêtes menées en 2024-2025 et 2025-2026 ont également permis la collecte de données administratives sur la répartition des élèves au sein des groupes à partir d'extractions d'informations des logiciels de vie scolaire. L'analyse de ces données, encore en cours, permettra de documenter plus finement la composition des groupes constitués au sein de chaque établissement, et de mesurer non seulement l'effet d'être scolarisé dans un établissement ayant mis en place un certain type de regroupement, mais aussi l'effet spécifique de l'appartenance à un type de regroupement donné, ainsi que les caractéristiques des enseignants et des élèves qui ont y été affectés. Ces données permettront par ailleurs d'explorer des dimensions encore peu documentées dans ce premier document de travail, telles que les effets de la réforme sur la mixité sociale et scolaire au sein des établissements ou sur le comportement des élèves grâce aux données collectées au printemps 2026 via les logiciels de vie scolaire.

↳ Bibliographie

Aisenberg, L., & Lobut, C. (2023). Différenciation des apprentissages : quelles modalités pour quels impacts ? *Note IDEE*, n°1, Novembre 2023.

Andreu, S., Ben-Ali, L., Blouet, L., Bressoux, P., Charpentier, A., Cioldi, I., Lacroix, A., Lima, L., Murat, F., Odin-Steiner, D., Raffaelli, C., Rocher, T., & Vourc'h, R. (2021). Évaluation de l'impact de la réduction de la taille des classes de CP et de CE1 en REP+ sur les résultats des élèves et les pratiques des enseignants. *Document de travail – Série études*, n° 2021-E04, DEPP.

Angrist, J. D., Lavy, V., Leder-Luis, J., & Shany, A. (2019). Maimonides rule redux. *American Economic Review: Insights*, 1(3), 309-324.

Antonovics, K., Black, S. E., Cullen, J. B., & Meiselman, A. Y. (2022). Patterns, determinants, and consequences of ability tracking: Evidence from Texas public schools, *NBER Working Paper Series* n° 30370, National Bureau of Economic Research.

Archer, L., Francis, B., Miller, S., Taylor, B., Tereshchenko, A., Mazenod, A., Pepper, D., & Travers, M. C. (2018). The symbolic violence of setting: A Bourdieusian analysis of mixed methods data on secondary students' views about setting. *British Educational Research Journal*, 44(1), 119-140.

Avery, C., & Bacher-Hicks, A. (2026). Inequality and within-school student-teacher assignment patterns. *Economics of Education Review*, 112, 102792.

Banerjee, A., Banerji, R., Berry, J., Duflo, E., Kannan, H., Mukherji, S., Shotland, M., & Walton, M. (2016). Mainstreaming an effective intervention: Evidence from randomized evaluations of "Teaching at the Right Level" in India. *NBER Working Paper Series*, n° 22746, National Bureau of Economic Research.

Baradji, É. (2025). Combien d'élèves devant un enseignant pendant une heure de cours dans les établissements du second degré à la rentrée 2024 ? *Note d'Information*, n° 25.49, DEPP.

Betts, J. R., & Shkolnik, J. L. (2000a). The effects of ability grouping on student achievement and resource allocation in secondary schools. *Economics of Education Review*, 19(1), 1-15.

Betts, J. R., & Shkolnik, J. L. (2000b). Key difficulties in identifying the effects of ability grouping on student achievement. *Economics of Education Review*, 19(1), 21-26.

Billings, S. B., Chyn, E., & Haggag, K. (2021). The long-run effects of school racial diversity on political identity. *American Economic Review: Insights*, 3(3), 267-284.

Boaler, J. (1997). When even the winners are losers: Evaluating the experiences of top set students. *Journal of curriculum studies*, 29(2), 165-182.

Bonneau, C., & Dousset, L. (2025). Gender Gap in High-Stakes Exams: What Role for Exam Preparation?, *Document de travail*.

Booij, A. S., Leuven, E., & Oosterbeek, H. (2017). Ability Peer Effects in University: Evidence from a Randomized Experiment. *The Review of Economic Studies*, 84(2), 547-578.

Bouguen, A., Grenet, J., & Gurgand, M. (2017). La taille des classes influence-t-elle la réussite scolaire ? *Notes IPP*, (28).

Boutchenik, B., & Maillard, S. (2019). Élèves hétérogènes, pairs hétérogènes. *Éducation & formations*, (100), 53-72.

Brodaty, T. (2010). Les effets de Pairs dans l'Éducation : une Revue de Littérature. *Revue d'économie politique*, 120(5), 739-757.

Burke, M. A., & Sass, T. R. (2013). Classroom peer effects and student achievement. *Journal of labor economics*, 31(1), 51-82.

Card, D., & Giuliano, L. (2016). Can tracking raise the test scores of high-ability minority students?. *American Economic Review*, 106(10), 2783-2816.

- Castle, S., Deniz, C. B., & Tortora, M. (2005). Flexible grouping and student learning in a high-needs school. *Education and urban society*, 37(2), 139-150.
- Ceesay, K., Raffaëlli, C., Rugambage, N., & Voisin, C. (2025). Les pratiques enseignantes à l'école élémentaire et au collège en France : les premiers résultats de l'enquête internationale Talis 2024. *Note d'information*, n° 25.55, DEPP.
- Charousset, P., Monnet, M., & Souidi, Y. (2023). Ségrégation sociale en milieu scolaire : appréhender ses causes et déterminer ses effets. *Note IPP*, n° 97.
- Charpentier, A., Longhi, L., Raffaëlli, C., & Solnon, A. (2020). Le métier d'enseignant. *Éducation & formations*, (101), pp-53.
- Claval, B., & Darnault, C. (coord.) (2025). Mise en place des groupes de besoins en français et mathématiques au collège. *Rapport de l'IGÉSR*, n° 24-25 007.
- Collins, C. A., & Gan, L. (2013). Does sorting students improve scores? An analysis of class composition. *NBER Working Paper Series*, n° 18848, National Bureau of Economic Research.
- Coppé, S., Grugeon-Allys, B., Horoks, J., Pilet, J., Solnon, A., Raffaëlli, C., & Charpentier, A. (2021). Premiers résultats de l'enquête sur les pratiques d'enseignement des mathématiques, Praesco en classe de troisième en 2019. *Note d'information*, n°21.11, DEPP.
- Dauphant, F., Evain, F., Guillermin, M., Simon, C., & Rocher, T. (2023). L'indice de position sociale (IPS) : un outil statistique pour décrire les inégalités sociales entre établissements - Focus sur les collèges. *Note d'information*, n° 23.16, DEPP.
- Delaney, J. M., & Devereux, P. J. (2022). Rank effects in education: What do we know so far?. *Handbook of Labor, Human Resources and Population Economics*, 1-24.
- Denning, J. T., Murphy, R., & Weinhardt, F. (2023). Class rank and long-run outcomes. *Review of Economics and Statistics*, 105(6), 1426-1441.
- Dubet, F., & Duru-Bellat, M. (2004). Qu'est-ce qu'une école juste ? *Revue française de pédagogie*, (146), pp. 105-114.
- Duflo, E., Dupas, P., & Kremer, M. (2011). Peer effects, teacher incentives, and the impact of tracking: Evidence from a randomized evaluation in Kenya. *American economic review*, 101(5), 1739-1774.
- Dupriez, V., & Draelants, H. (2004). Classes homogènes versus classes hétérogènes : les apports de la recherche à l'analyse de la problématique. *Revue française de pédagogie*, 148(1), 145-165.
- Duru-Bellat, M., & Mingat, A. (1997). La constitution de classes de niveau dans les collèges ; les effets pervers d'une pratique à visée égalisatrice. *Revue française de sociologie*, 759-789.
- EEF, Education Endowment Foundation. (2021). Within class attainment grouping. *Teaching and Learning Toolkit*. educationendowmentfoundation.org.uk.
- El Bah, F. (2026). Les effets à long terme de la réduction de la taille de classe en éducation prioritaire sur les performances des élèves : analyse sur la première cohorte d'élèves bénéficiaires. *Document de travail*, n° 2026-E01, DEPP.
- Fajeau, M., Grenet, J., Laveissière, E., & Leonetti, O. (2025). Efficacité économique de la réduction de la taille des classes. *Focus du CAE*, (113).
- Figlio, D., & Özek, U. (2024). The unintended consequences of test-based remediation. *American Economic Journal: Applied Economics*, 16(1), 60-89.
- Figlio, D. N., & Page, M. E. (2002). School choice and the distributional effects of ability tracking: does separation increase inequality?. *Journal of Urban Economics*, 51(3), 497-514.
- Fleischmann, M., Huebner, N., Nagengast, B., & Trautwein, U. (2023). The dark side of detracking: Mixed-ability classrooms negatively affect the academic self-concept of students with low academic achievement. *Learning and Instruction*, 86, 101753.

- Fredriksson, P., Öckert, B., & Oosterbeek, H. (2013). Long-term effects of class size. *The Quarterly journal of economics*, 128(1), 249-285.
- Goube, L., Lugand, A., Potrel, D., Raffaëlli, C., Bianco, M., Bressoux, P., Lima, L., & Vibert, A. (2024). Premiers résultats de l'enquête sur les pratiques d'enseignement du français, Praesco, en classe de CM2 en 2021. *Note d'information*, n°24.11, DEPP.
- Grenet, J., & Landais, C. (2025). Éducation : comment mieux orienter la dépense publique. *Notes du conseil d'analyse économique*, 84(1), 1-12.
- Hattie, J. (2023). *Visible Learning: The Sequel. A Synthesis of Over 2,100 Meta-Analyses Relating to Achievement*. Routledge, 512p.
- Hermann, Z., Horváth, H., & Kisfalusi, D. (2024). Are separate classrooms inherently unequal? The effect of within-school sorting on the socioeconomic test score gap in Hungary. *Economics of Education Review*, 103, 102582.
- Hoffer, T. B. (1992). Middle school ability grouping and student achievement in science and mathematics. *Educational evaluation and policy analysis*, 14(3), 205-227.
- Hoxby, C. M., & Weingarth, G. (2005). Taking race out of the equation: School reassignment and the structure of peer effects (No. 7867). *Working paper, Harvard university*.
- Ireson, J., Clark, H., & Hallam, S. (2002). Constructing ability groups in the secondary school: Issues in practice. *School Leadership & Management*, 22(2), 163-176.
- Jepsen, C., & Rivkin, S. (2009). Class size reduction and student achievement: The potential tradeoff between teacher quality and class size. *Journal of human resources*, 44(1), 223-250.
- Jerrim, J. (2021). The association between within-class grouping and children's achievement in mathematics during Year 2, Year 5 and Year 9. *School choices report*, Education Endowment Foundation, March 2021.
- Jerrim, J., Alejandro Lopez-Agudo, L., & David Marcenaro-Gutierrez, O. (2025). Does within-school between-class ability grouping harm the educational outcomes of socio-economically disadvantaged children? International evidence. *European Sociological Review*, 41(5), 722-736.
- Kaplan, E., Spenkuch, J. L., & Tuttle, C. (2025). A Different World: Enduring Effects of School Desegregation on Ideology and Attitudes, *NBER Working Paper Series*, n° 33365, National Bureau of Economic Research.
- Kerckhoff, A. C. (1986). Effects of ability grouping in British secondary schools. *American sociological review*, 842-858.
- Kulik, C. L. C., & Kulik, J. A. (1982). *Effects of ability grouping on secondary school students: A meta-analysis of evaluation findings*. *American educational research journal*, 19(3), 415-428.
- Landaud, F., & Maurin, É. (2026). Tracking when ranking matters. *Economics of Education Review*, 112, 102785.
- Lavy, V., Paserman, M. D., & Schlosser, A. (2012a). Inside the black box of ability peer effects: Evidence from variation in the proportion of low achievers in the classroom. *The Economic Journal*, 122(559), 208-237.
- Lavy, V., Silva, O., & Weinhardt, F. (2012b). The good, the bad, and the average: Evidence on ability peer effects in schools. *Journal of Labor Economics*, 30(2), 367-414.
- Leuven, E., Oosterbeek, H., & Rønning, M. (2008). Quasi-experimental estimates of the effect of class size on achievement in Norway. *The Scandinavian Journal of Economics*, 110(4), 663-693.
- Lou, Y., Abrami, P. C., & Spence, J. C. (2000). Effects of within-class grouping on student achievement: An exploratory model. *The Journal of Educational Research*, 94(2), 101-112.

- Ly, S. T., Maurin, É., & Riegert, A. (2014). La mixité sociale et scolaire en Île-de-France : le rôle des établissements. *Rapport IPP*, n° 4.
- Ly, S. T., & Riegert, A. (2016). « Mixité sociale et scolaire et ségrégation inter—et intra-établissement dans les collèges et lycées français ». Rapport du Conseil national d'évaluation du système scolaire.
- Marsh, H. W., & Parker, J. W. (1984). Determinants of student self-concept: Is it better to be a relatively large fish in a small pond even if you don't learn to swim as well?. *Journal of personality and social psychology*, 47(1), 213.
- de Monestrol, H., Paillet, V., Valdenaire, B., Andreu, S., Bret, A., Cannes, F.-X., Garnero, M., Gill-Sotty, C., Hick, M., Korchia, L., Lacroix, A., Laskowski, C., Léger, A., M'Bafoumou, A., Mas, S., Rocher, T., Rue, G. & Vezon Persem, É. (2026). Évaluations de début de cinquième 2025. Résultats. *Document de travail – Série études*, n° 2025-E13, DEPP.
- Mons, N., Chesné, J.-F., Piquet, N., & Coudroy, T. (2017). Différenciation pédagogique : comment adapter l'enseignement à la réussite de tous les élèves ? *Dossier de synthèse*. Cnesco.
- Monso, O., Fougère, D., Givord, P., & Pirus, C. (2019). Les camarades influencent-ils la réussite et le parcours des élèves ? *Éducation & formations*, (100), 23-52.
- Murat, F., Rocher, T. (2026). Comment progressent les élèves de l'éducation prioritaire au cours de l'école élémentaire ? Écarts de performances observés entre les élèves de REP+ et de REP et les élèves du secteur public hors éducation prioritaire. *Document de travail – Série Études*, n° 2026-E02, DEPP.
- Murphy, R., & Weinhardt, F. (2020). Top of the class: The importance of ordinal rank. *The Review of Economic Studies*, 87(6), 2777-2826.
- OCDE. (2023a). Selecting and Grouping Students. In *PISA 2022 Results (Volume II): Learning During – and From – Disruption*, PISA, OECD Publishing, Paris.
- OCDE (2023b), *PISA 2022 Results (Volume I and II) – Country Notes: France*, OECD Publishing, Paris.
- Pagani, L., Comi, S., & Origo, F. (2021). The effect of school rank on personality traits. *Journal of Human Resources*, 56(4), 1187-1225.
- Piketty, T., & Valdenaire, M. (2006). L'impact de la taille des classes sur la réussite scolaire dans les écoles, collèges et lycées français : estimations à partir du panel primaire 1997 et du panel secondaire 1995. *Les Dossiers de l'Éducation Nationale*, n° 173, Direction de l'évaluation et de la prospective.
- Potrel, D., Raffaëlli, C., & Lima, L. (2024). Premiers résultats de l'enquête sur les pratiques d'enseignement du français, Praesco en classe de troisième en 2021. *Note d'information*, n° 24-36, DEPP.
- Rees, D. I., Brewer, D. J., & Argys, L. M. (2000). How should we measure the effect of ability grouping on student performance?. *Economics of Education Review*, 19(1), 17-20.
- Rosenbaum, J. E. (1980). Track misperceptions and frustrated college plans: An analysis of the effects of tracks and track perceptions in the National Longitudinal Survey. *Sociology of education*, 74-88.
- Sims, D. P. (2009). Crowding Peter to educate Paul: Lessons from a class size reduction externality. *Economics of Education Review*, 28(4), 465-473.
- Slavin, R. E. (1990). Achievement effects of ability grouping in secondary schools: A best-evidence synthesis. *Review of educational research*, 60(3), 471-499.
- Souidi, Y. (2023). « Ségrégation entre classes d'un même établissement : quels effets de la suppression des sections bilingues et européennes ? », Note IPP n°94.
- Steenbergen-Hu, S., Makel, M. C., & Olszewski-Kubilius, P. (2016). What one hundred years of research says about the effects of ability grouping and acceleration on K–12 students' academic achievement: Findings of two second-order meta-analyses. *Review of educational research*, 86(4), 849-899.
- Sukhnandan, L., & Lee, B. (1998). Streaming, setting and grouping by ability. *Slough: NFER*.

Terrin, É., & Triventi, M. (2023). The effect of school tracking on student achievement and inequality: A meta-analysis. *Review of Educational Research*, 93(2), 236-274.

Wu, J., Wei, X., Zhang, H., & Zhou, X. (2019). Elite schools, magnet classes, and academic performances: Regression-discontinuity evidence from China. *China Economic Review*, 55, 143-167.

➤ Annexes

Annexe 1 : Protocole d'enquête

Cadre général

L'étude relative à l'organisation des enseignements au collège (EEC) de la DEPP s'appuie sur un questionnaire auto-administré envoyé à 1 525 établissements publics de France hors Mayotte.

L'objectif de ce questionnaire est d'apprécier l'hétérogénéité des contextes scolaires concernés par la mesure des « groupes de besoins », de rendre compte des organisations pédagogiques retenues pour l'année scolaire 2024-2025 en français et en mathématiques en classes de 6^e et de 5^e et de documenter la réception et l'appropriation de la mesure s'agissant de sa mise en œuvre et de ses potentiels premiers effets perçus. Deux questionnaires distincts ont été envoyés : un premier à destination des chefs d'établissement, et un second adressé aux enseignants coordonnateurs de français et de mathématiques intervenant pour les organisations pédagogiques mises en place en 6^e et en 5^e pour l'année scolaire 2024-2025. Les deux questionnaires partagent un ensemble de questions communes concernant les modalités pratiques de constitution des groupes et le regard porté par les acteurs sur la mesure et ses effets perçus, qui sont complétées par quelques questions spécifiques liées à la fonction de la personne répondante au sein de l'établissement.

Les coordonnateurs de français et de mathématiques ont été renseignés par les chefs d'établissement lors d'une campagne de contact en février-mars 2025, en amont de la passation des questionnaires. Les questionnaires principaux ont été envoyés à partir de fin avril 2025 selon un calendrier échelonné pour tenir compte des vacances de printemps, et les réponses ont été collectées jusqu'à la fin du mois de juin.

Les questionnaires ont été réalisés sur le logiciel Sphinx IQ et diffusés à partir de la plateforme Sphinx Online. Ils ont été élaborés par la DEPP en collaboration étroite avec l'IGÉSR et l'équipe de recherche associée à l'étude EEC (École d'Économie de Paris, Fondation Nationale de Sciences Politiques et Université Paris Dauphine-PSL).

Définition de la base de sondage et de l'échantillon principal

La base de sondage a été constituée à partir des données extraites d'Archipel pour l'année 2022-2023, des informations des fichiers Scolarité pour l'année 2023-2024 et des résultats aux évaluations nationales de 6^e pour la rentrée 2023. Elle correspond aux 5 250 collèges publics de France hors Mayotte ouverts à la rentrée 2024 et comptant au moins 20 élèves en 6^e et 20 élèves en 5^e.

L'échantillon de l'étude EEC est composé de 1 676 établissements représentatifs de cette base de sondage. Les 434 collèges publics de l'échantillon de l'expérimentation cinquième 2024 pour le renouvellement des exercices des évaluations nationales ont tous été sélectionnés pour faire partie de l'étude EEC, de manière à maximiser le nombre d'établissements concernés par les évaluations nationales de début de cinquième en 2024, avant leur généralisation à l'ensemble des collèges à partir de la rentrée scolaire de septembre 2025. Les 1 242 établissements supplémentaires ont été sélectionnés en deux temps à l'aide de procédures de tirage aléatoire simple stratifié afin de garantir la représentativité de l'échantillon vis-à-vis de la base de sondage.

La base de sondage a ainsi été divisée en 81 strates définies à partir de l'appartenance au réseau de l'éducation prioritaire (Hors REP, REP et REP+), de la taille de l'établissement pendant l'année 2022-2023 (en terciles), de la densité communale en 2024 (en trois positions) et de l'hétérogénéité des notes à l'entrée en 6^e en septembre 2023 (en terciles des écarts-types des notes à l'échelle de l'établissement). Afin de garantir une taille minimale de 13 établissements par strate, certaines strates

avec différentes tailles d'établissement ou densités communales ont été fusionnées, pour aboutir à un découpage en 56 strates.

Dans un premier temps, un échantillon principal de 1 066 établissements a été tiré à partir de ces 56 strates, proportionnellement au nombre d'établissements dans chacune des strates. Ces établissements, ainsi que les 434 établissements de l'échantillon d'expérimentation des évaluations nationales de 5^e, ont été contactés afin d'obtenir les coordonnées des coordonnateurs de français et de mathématiques pour l'envoi des questionnaires de l'enquête.

Un échantillon de remplacement de 1 500 établissements a ensuite été constitué à partir des établissements restants dans la base de sondage. L'objectif était de pouvoir sélectionner des établissements de substitution en cas de non-réponse des établissements principaux lors de la première phase de contact. Afin de garantir un minimum de 13 établissements par strate, les 56 strates de tirage ont été regroupées en 40 strates avant le tirage aléatoire stratifié.

Parmi les 1 066 établissements de l'échantillon principal, 176 chefs d'établissements n'ont pas répondu au formulaire lors de la campagne visant à collecter les contacts des coordonnateurs. Un nombre équivalent de collègues a donc été tiré aléatoirement dans l'échantillon de remplacement, au sein des strates d'origine des établissements remplacés, afin de les contacter pour obtenir les informations pour joindre les coordonnateurs. Parmi eux, 151 ont répondu à ce formulaire de contact. Finalement, 25 collègues initialement remplacés ont finalement répondu au formulaire de contact hors de la période de passation initiale et ont été intégrés à l'échantillon final. L'échantillon final auquel a été adressé les questionnaires de l'enquête est ainsi composé des 1 349 établissements de l'échantillon principal qui ont renvoyé le formulaire et des 176 établissements de remplacement, qu'ils aient renvoyé ou non le formulaire.

En plus des 1 525 questionnaires adressés aux chefs d'établissement, 2 774 questionnaires ont été adressés aux coordonnateurs de français et de mathématiques intervenant pour les organisations pédagogiques mises en place en 6^e et en 5^e et désignés par les chefs d'établissement pour l'année scolaire 2024-2025. Les formulaires de contact ont permis d'identifier ces coordonnateurs au sein de 1 180 établissements. Plus précisément, 21 % des établissements n'ont renseigné aucun coordonnateur dans les formulaires retournés, 2 % n'en ont identifié qu'un seul, 58 % en ont déclaré deux et 19 % plus de deux. Par ailleurs, parmi les établissements qui déclarent au moins un coordonnateur, 77 % renseignent au moins un coordonnateur de français, 78 % identifient au moins un coordonnateur de mathématiques et 97 % indiquent qu'il y a au moins un coordonnateur pour chacune des disciplines.

Enfin, un échantillon pilote de 100 établissements a été construit parmi les 2 250 établissements restants dans la base de sondage, à partir d'un tirage aléatoire sur 26 strates. Cet échantillon pilote a permis de tester le matériel d'enquête et ne fait pas partie de l'échantillon final dont les résultats sont présentés au sein de ce document de travail.

En plus des questionnaires, l'étude EEC mobilise des données extraites des logiciels de vie scolaire afin de pouvoir identifier les groupes d'appartenance des élèves au cours de l'année. Les données sont extraites par les équipes de direction des établissements avant d'être transmises à la DEPP via un formulaire dédié.

Annexe 2 : Pondération initiale, non-réponse et redressement

Pondération initiale des établissements

Afin d'assurer la représentation des 5 250 établissements de la base de sondage, des poids de sondage initiaux ont été définis en prenant l'inverse de la probabilité pour chaque établissement d'appartenir à l'échantillon de travail. Les 434 établissements de l'échantillon d'expérimentation cinquième 2024 se sont ainsi vus attribués un poids égal à un. Pour les 1 242 établissements restants, le choix a été fait de faire comme s'ils avaient été tirés en même temps au sein des 56 strates initiales. Cela permet d'éviter de définir un schéma de pondération complexe avec des poids très importants pour tenir compte de la probabilité d'appartenance à l'échantillon des établissements de remplacement. Ce choix a également été motivé par les similarités entre les établissements de l'échantillon principal et ceux de l'échantillon de remplacement, et par l'absence de déformation de l'échantillon principal après ajout des établissements de l'échantillon de remplacement.

La pondération initiale a ensuite été adaptée pour tenir compte de la non-réponse à l'enquête et afin de garantir la représentativité de l'échantillon final vis-à-vis de la base de sondage. Cette repondération a été réalisée en différentes étapes pour être en cohérence avec le protocole d'enquête et pour permettre aux facteurs influençant la non-réponse de varier entre les différentes phases de l'enquête.

Repondération liée au non-renvoi des formulaires de contact

Une première opération de repondération a été réalisée pour tenir compte du non-renvoi des formulaires permettant d'obtenir les informations de contact des coordonnateurs de français et de mathématiques intervenant pour les organisations pédagogiques mises en place en 6^e et en 5^e pour l'année scolaire 2024-2025. Parmi les 1 500 établissements de l'échantillon principal, 151 établissements n'ont jamais renvoyé les formulaires complétés. Au sein de l'échantillon de remplacement, qui vise à compenser cette non-réponse initiale, 25 établissements n'ont pas communiqué les informations de contact demandées.

Ces 176 établissements qui n'ont pas répondu lors de cette première phase de l'enquête représentent 11 % de l'échantillon initial. En moyenne, ces collèges sont des établissements de grande taille, situés dans de grands centres urbains de zone C (région parisienne, Montpellier et Toulouse), et ils se distinguent par une forte hétérogénéité des notes à l'entrée en 6^e.

Deux jeux de poids ont ici été définis pour corriger cette non-réponse initiale. D'une part, le poids a été révisé pour définir une nouvelle pondération des établissements répondants à appliquer aux réponses des coordonnateurs. En effet, les coordonnateurs ne sont connus que pour les établissements ayant renseigné leurs coordonnées, il est donc nécessaire de tenir compte de cette première phase de l'enquête dans la définition de leur poids individuel. Pour cela, un modèle logistique a été mobilisé afin de mesurer la probabilité de renvoi du formulaire de contact à partir des caractéristiques de l'établissement. Le poids initial a été multiplié par l'inverse de la probabilité de réponse estimée pour définir ce poids corrigé des établissements. Ce poids a été attribué à chaque coordonnateur renseigné par l'établissement pour définir leur pondération initiale, sur le modèle conventionnel utilisé par l'Insee pour calculer le poids des individus à partir de leur logement dans le recensement²³.

D'autre part, le poids initial a été révisé pour tenir compte de la probabilité que l'établissement soit présent dans la suite de l'enquête et donc dans l'échantillon final de travail, afin de tenir compte de cette première phase de l'enquête dans la construction des poids finaux de l'établissement à appliquer aux réponses des chefs d'établissement. Pour cela, les poids utilisés pour les établissements

²³ [Recensement de la population – Les pondérations, Juin 2024.](#)

de l'échantillon d'expérimentation cinquième 2024 et pour les établissements de l'échantillon principal sont les poids corrigés par le modèle de régression logistique comme précédemment, mais les poids initiaux des établissements de l'échantillon de remplacement ne sont pas corrigés car l'ensemble de ces établissements ont été destinataires des questionnaires adressés aux chefs d'établissement, même en cas de non-réponse aux formulaires de contact. Les poids de ces établissements doivent donc être modifiés pour le volet « coordonnateurs » de l'enquête DEPP, mais pas pour le volet « chefs d'établissement ».

Repondération liée à la non-réponse aux questionnaires adressés aux chefs d'établissement

La seconde opération de repondération consiste à corriger les poids des établissements pour tenir compte de la non-réponse des chefs d'établissement au questionnaire principal de l'enquête. Celle-ci recouvre à la fois la non-réponse totale (questionnaire non reçu) et la non-réponse partielle trop importante pour que le questionnaire puisse être exploité à des fins statistiques.

À l'issue de la période de collecte fin juin 2025, 1 315 questionnaires ont été reçus et retenus pour l'analyse, ce qui représente 86 % des établissements interrogés et 78 % des établissements de l'échantillon initial. Les établissements qui n'ont pas répondu appartiennent davantage au réseau de l'éducation prioritaire, ils sont surreprésentés dans les grands centres urbains et sous-représentés dans les communes rurales et se situent davantage dans les DROM.

À nouveau, une régression logistique a été mobilisée pour estimer la probabilité d'avoir répondu de façon suffisamment complète au questionnaire à partir des caractéristiques de l'établissement, des informations issues des formulaires de contact et de variables construites lors du processus de collecte des données. Les poids finaux des établissements utilisés pour l'exploitation des réponses aux questionnaires des chefs d'établissement sont donc égaux aux poids initiaux des établissements corrigés pour tenir compte du non-renvoi des formulaires de contact (sauf pour les établissements de remplacement) multipliés par l'inverse de la probabilité de répondre au questionnaire final, estimée par la régression logistique.

Calage sur marges

L'échantillon final compte 1 315 établissements et sa représentativité vis-à-vis de la base de sondage a été conservée grâce au travail de repondération. Cependant, la distribution finale des établissements entre les différentes tranches de densité communale suggère un léger déséquilibre par rapport à la population de référence, présent dès le tirage initial des échantillons. Pour corriger ces écarts, un calage sur marges a été réalisé pour conclure la phase de redressement des données d'enquête. L'échantillon final a été calé sur les données de stratification (appartenance à l'éducation prioritaire, taille de l'établissement, densité communale, hétérogénéité des scores à l'entrée en 6^e) avec la méthode logit à l'aide du package Icarus sur R. La représentativité finale de l'échantillon des chefs d'établissement s'en trouve ainsi renforcée (**Tableau 12 ; Figure 28**).

TABLEAU 12 • Représentativité de l'échantillon final – Questionnaires Chefs d'établissement

Variable	Échantillon initial pondéré	Échantillon final repondéré	P-value
Effectif pondéré	5250,0	5250,0	
Appartenance à un réseau d'éducation prioritaire :			
Hors EP	4187,0 (79,8 %)	4187,0 (79,8 %)	1,000
REP	713,0 (13,6 %)	713,0 (13,6 %)	
REP+	350,0 (6,7 %)	350,0 (6,7 %)	
Densité communale :			
Commune rurale	1647,7 (31,4 %)	1637,0 (31,2 %)	0,962
Communes de densité intermédiaire	1788,9 (34,1 %)	1817,0 (34,6 %)	
Grands centres urbains	1813,4 (34,5 %)	1796,0 (34,2 %)	
Appartenance à l'échantillon d'expérimentation des évaluations nationales de 5 ^e	434,0 (8,3 %)	428,1 (8,2 %)	0,863
Taille de l'établissement : Élèves	499,48 (183,17)	498,68 (182,67)	0,917
Taille de l'établissement : Enseignants	38,27 (11,71)	38,26 (11,74)	0,987
Effectifs d'élèves de niveau collège : Total niveau collège	485,06 (177,29)	484,25 (176,95)	0,913
Effectifs détaillés de l'enseignement adapté : Total enseignement adapté	14,42 (26,09)	14,43 (25,57)	0,991
Effectifs d'élèves de l'enseignement adapté : 6 ^e SEGPA	11,52 (5,09)	11,32 (4,55)	0,634
Effectifs d'élèves de l'enseignement adapté : 5 ^e SEGPA	14,36 (5,07)	14,21 (4,47)	0,694
Pourcentage de filles dans l'établissement	49,09 (2,75)	49,08 (2,79)	0,942
IPS de l'établissement : Ensemble des élèves	100,49 (14,53)	100,51 (14,42)	0,972
Pourcentage de boursiers par échelon (collège) : Ensemble	0,27 (0,15)	0,27 (0,15)	0,974
IPS par niveau de formation : Entrants en 6 ^e	100,80 (14,67)	100,80 (14,61)	1,000
Hétérogénéité sociale des élèves : Ensemble des élèves	31,15 (4,47)	31,17 (4,48)	0,902
IPS par niveau de formation : Niveau collège hors SEGPA	100,98 (14,46)	101,00 (14,34)	0,977
Score moyen des élèves de 6 ^e : Score en français	251,98 (16,03)	252,01 (15,87)	0,958
Score moyen des élèves de 6 ^e : Score en mathématiques	249,53 (18,23)	249,62 (18,09)	0,912
Hétérogénéité scolaire (écart-type de la moyenne des évaluations de 6 ^e 2023)	42,75 (5,06)	42,75 (4,98)	0,982
Note à l'écrit au DNB : Note moyenne à l'écrit - Série Générale	10,95 (1,37)	10,95 (1,35)	0,990
Taux d'accès 6 ^e -3 ^e	0,92 (0,04)	0,92 (0,04)	0,949

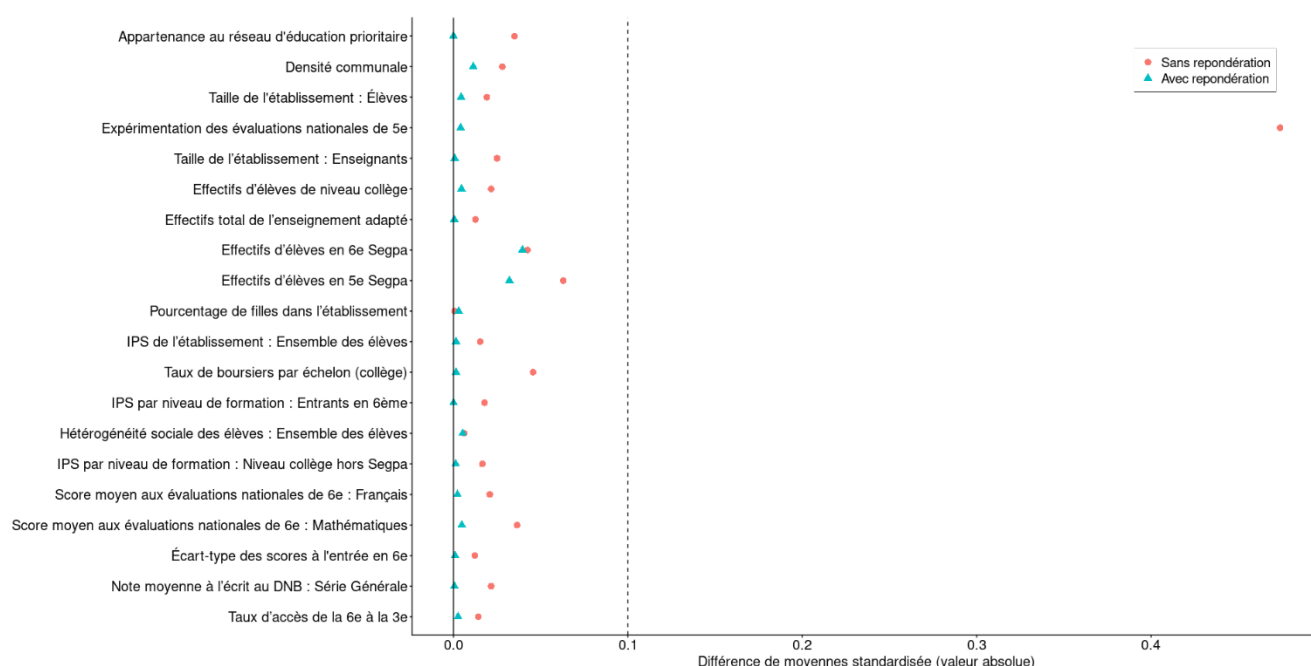
Lecture : Dans l'échantillon initial avec poids d'échantillonnage, 4 187 collèges sur les 5250 n'appartiennent pas au réseau de l'éducation prioritaire. Cela représente 79,8 % de l'échantillon. Ce même effectif et cette même proportion se retrouvent dans l'échantillon final en utilisant les poids corrigés pour tenir compte de la non-réponse.

Note : Pour les variables catégorielles, les proportions sont présentées entre parenthèses et la p-value est celle du test du χ^2 . Pour les variables numériques, l'écart-type est présenté entre parenthèses et la p-value est celle d'un test de Student.

Champ : Collèges publics de France hors Mayotte, ouverts à la rentrée 2024 et accueillant au moins 20 élèves de niveau 6^e et 20 élèves de niveau 5^e.

Sources : DEPP, enquête EEC, questionnaires Chefs d'établissement, printemps 2025, données finales repondérées ; Évaluations nationales à l'entrée en 6^e 2023, CYCLADES-DNB 2023, Constat de rentrée 2023.

FIGURE 28 • Différences entre l'échantillon initial pondéré, l'échantillon final non-pondéré et l'échantillon final repondéré – Chefs d'établissement



Lecture : La différence de moyennes standardisées pour les distributions des trois modalités de la variable d'appartenance au réseau de l'éducation prioritaire (Hors REP, REP et REP+) entre l'échantillon initial pondéré et l'échantillon final non-pondéré est égale à 0,04, ce qui correspond à une différence très faible : la non-réponse n'a pas véritablement modifié la représentation de ces trois modalités dans l'échantillon final. Une fois que la pondération de l'échantillon final est prise en compte, la différence standardisée entre l'échantillon initial et l'échantillon final tombe à 0, ce qui signifie que les très faibles déséquilibres liés à la non-réponse sont corrigés par la procédure de repondération, au moins du point de vue des moyennes.

Note : Les différences de moyennes standardisées ont été calculées à partir du package *tableone* sur R. L'écart-type pris en compte pour la standardisation est donc celui des observations empilées (échantillon initial pondéré et échantillon final non pondéré, échantillon initial pondéré et échantillon final repondéré). Pour les variables catégorielles, une seule différence de moyennes standardisée est calculée pour la variable, plutôt qu'une différence par modalité.

Champ : Collèges publics de France hors Mayotte, ouverts à la rentrée 2024 et accueillant au moins 20 élèves de niveau 6^e et 20 élèves de niveau 5^e.

Sources : DEPP, enquête EEC, questionnaires Chefs d'établissement, printemps 2025, données finales repondérées ; Évaluations nationales à l'entrée en 6^e 2023, CYCLADES-DNB 2023, Constat de rentrée 2023.

Pondération initiale des coordonnateurs

L'utilisation des questionnaires adressés aux coordonnateurs de français et de mathématiques intervenant pour les organisations pédagogiques mises en place en 6^e et en 5^e pour l'année scolaire 2024-2025 nécessite de leur attribuer une pondération spécifique. Le poids initial de chaque coordonnateur est égal au poids de son établissement après correction du non-renvoi des formulaires de contact (cf. *supra*), ce qui permet de réaliser des statistiques à leur échelle plutôt qu'à celle de leur établissement.

Repondération des questionnaires adressés aux coordonnateurs

Un modèle logistique de repondération a été appliqué pour tenir compte de la non-réponse des coordonnateurs. Comme pour les chefs d'établissement, la non-réponse recouvre ici la non-réponse totale (1 187 coordonnateurs sur les 2 774 contactés n'ont pas répondu au questionnaire) et les questionnaires trop incomplets pour être pris en compte dans l'analyse (166 questionnaires ont ainsi été exclus). Ainsi, le taux de réponse au questionnaire est de 51 %. Il est plus important pour les coordonnateurs de mathématiques (58 %) que pour les coordonnateurs de français (45 %).

Le modèle de repondération s'appuie sur les caractéristiques des établissements et sur les données obtenues dans les formulaires de contact (discipline et présence éventuelle d'autres coordonnateurs). À l'issue de cette procédure de repondération, la distribution des coordonnateurs ayant répondu au questionnaire est similaire à celle des coordonnateurs auxquels le questionnaire avait été adressé (**Tableau 13** ; **Figure 29**).

TABLEAU 13 • Représentativité de l'échantillon final – Questionnaires Coordonnateurs

Variable	Échantillon initial pondéré	Échantillon final repondéré	P-value
Effectif pondéré	9783,5	9783,1	
Discipline des coordonnateurs : Nombre de coordonnateurs de français Nombre de coordonnateurs de mathématiques	4841,2 (49,5 %) 4942,3 (50,5 %)	4834,4 (49,4 %) 4948,7 (50,6 %)	0,944
Appartenance à un réseau d'éducation prioritaire : Hors EP REP REP+	7690,6 (78,6 %) 1345,2 (13,7 %) 747,6 (7,6 %)	7671,4 (78,4 %) 1357,9 (13,9 %) 753,8 (7,7 %)	0,981
Densité communale : Commune rurale Communes de densité intermédiaire Grands centres urbains	2901,8 (29,7 %) 3354,4 (34,3 %) 3527,2 (36,1 %)	2923,9 (29,9 %) 3348,4 (34,2 %) 3510,8 (35,9 %)	0,974
Appartenance à l'échantillon d'expérimentation des évaluations nationales de 5e	798,2 (8,2 %)	778,2 (8,0 %)	0,579
Taille de l'établissement : Élèves	507,69 (185,14)	507,06 (182,82)	0,883
Taille de l'établissement : Enseignants	38,94 (12,04)	38,76 (11,72)	0,519
Effectifs d'élèves de niveau collège : Total niveau collège	492,45 (179,19)	491,92 (177,92)	0,899
Effectifs détaillés de l'enseignement adapté : Total enseignement adapté	15,24 (27,73)	15,14 (27,84)	0,867
Effectifs d'élèves de l'enseignement adapté : 6e SEGPA	11,83 (5,97)	11,73 (6,16)	0,731
Effectifs d'élèves de l'enseignement adapté : 5e SEGPA	14,82 (5,98)	14,97 (5,77)	0,587
Pourcentage de filles dans l'établissement	49,09 (2,75)	49,10 (2,75)	0,849
IPS de l'établissement : Ensemble des élèves	100,48 (15,02)	100,47 (15,05)	0,971
Pourcentage de boursiers par échelon (collège) : Ensemble	0,28 (0,15)	0,28 (0,16)	0,995
IPS par niveau de formation : Entrants en 6e	100,90 (15,05)	100,91 (15,04)	0,990
Hétérogénéité sociale des élèves : Ensemble des élèves	31,14 (4,57)	31,14 (4,64)	0,940
IPS par niveau de formation : Niveau collège hors SEGPA	101,00 (14,94)	100,99 (14,98)	0,975
Score moyen des élèves de 6e : Score en français	251,73 (16,44)	251,42 (16,54)	0,534
Score moyen des élèves de 6e : Score en mathématiques	249,22 (18,79)	249,02 (18,91)	0,717
Hétérogénéité scolaire (écart-type de la moyenne des évaluations de 6e 2023)	42,86 (5,15)	42,84 (5,21)	0,824
Note à l'écrit au DNB : Note moyenne à l'écrit - Série Générale	10,94 (1,40)	10,95 (1,40)	0,854
Taux d'accès 6e-3e	0,92 (0,04)	0,92 (0,04)	0,431

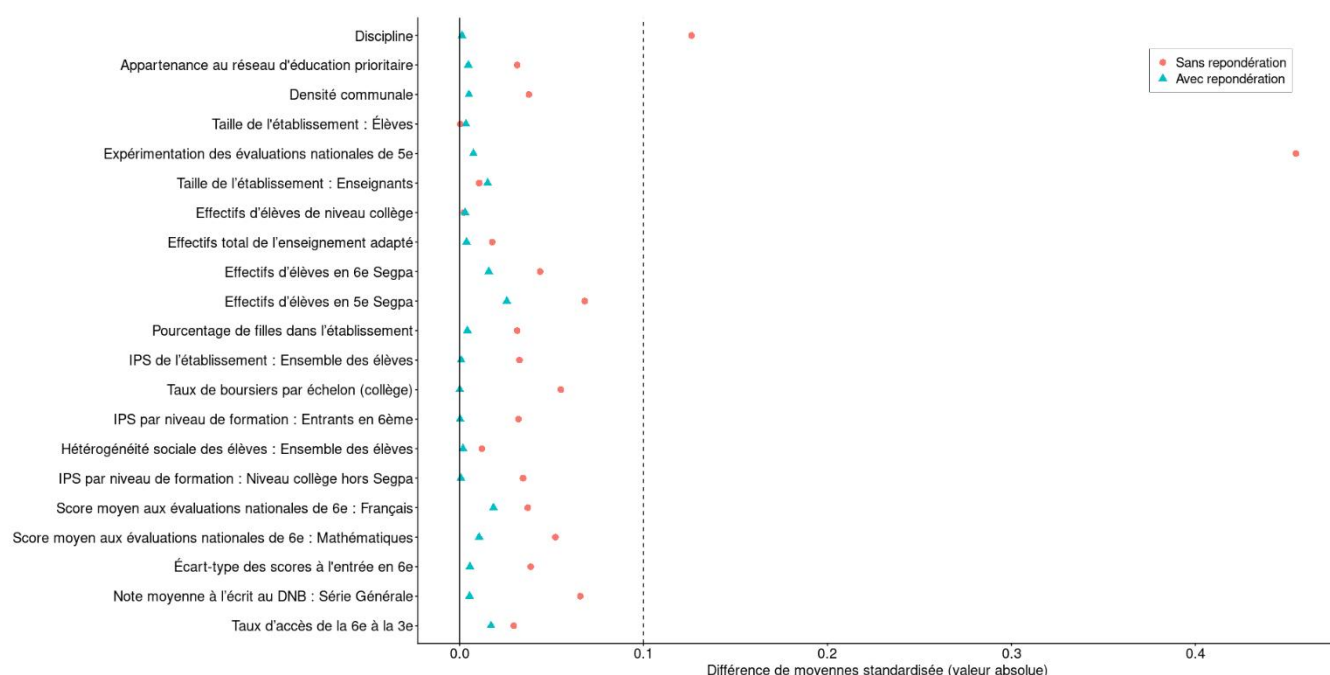
Lecture : D'après les poids initiaux des coordonnateurs, correspondant aux poids de l'établissement corrigés du non-renvoi initial des formulaires de contact, 7 690,6 coordonnateurs enseignent hors du réseau de l'éducation prioritaire, ce qui représente 78,6 % de l'échantillon. Dans l'échantillon final, en utilisant les poids corrigés pour tenir compte de la non-réponse, le nombre de coordonnateurs qui enseignent hors de l'enseignement prioritaire est égal à 7 671,4, ce qui représente 78,4 % de l'échantillon. Il semble donc y avoir relativement plus d'enseignants dans le réseau d'éducation prioritaire avec les poids finaux, mais cette différence n'est pas significative.

Note : Pour les variables catégorielles, les proportions sont présentées entre parenthèses et la p-value est celle du test du χ^2 . Pour les variables numériques, l'écart-type est présenté entre parenthèses et la p-value est celle d'un test de Student.

Champ : Coordonnateurs intervenants pour les enseignements de français et de mathématiques en classes de 6^e et de 5^e au sein de collèges de France hors Mayotte, ouverts à la rentrée 2024 et accueillant au moins 20 élèves de niveau 6^e et 20 élèves de niveau 5^e.

Sources : DEPP, enquête EEC, questionnaires Chefs d'établissement, printemps 2025, données finales repondérées ; Évaluations nationales à l'entrée en 6^e 2023, CYCLADES-DNB 2023, Constat de rentrée 2023.

FIGURE 29 • Différences entre l'échantillon initial pondéré, l'échantillon final non pondéré et l'échantillon final repondéré – Coordonnateurs



Lecture : La différence de moyennes standardisées pour les distributions des deux disciplines (français ou mathématiques) entre l'échantillon initial pondéré et l'échantillon final non-pondéré est égale à 0,13, ce qui correspond à un déséquilibre léger en raison de taux de réponses légèrement plus importants en mathématiques : la non-réponse a ainsi légèrement modifié la représentation de de l'échantillon final. Une fois que la pondération de l'échantillon final est prise en compte, la différence standardisée entre l'échantillon initial et l'échantillon final tombe à 0, ce qui signifie que les très faibles déséquilibres liés à la non-réponse sont corrigés par la procédure de repondération, au moins du point de vue des moyennes.

Note : Les différences de moyennes standardisées ont été calculées à partir du package *tableone* sur R. L'écart-type pris en compte pour la standardisation est donc celui des observations empilées (échantillon initial pondéré et échantillon final non pondéré, échantillon initial pondéré et échantillon final repondéré). Pour les variables catégorielles, une seule différence de moyennes standardisée est calculée pour la variable, plutôt qu'une différence par modalité.

Champ : Coordonnateurs intervenants pour les enseignements de français et de mathématiques en classes de 6^e et de 5^e au sein de collèges de France hors Mayotte, ouverts à la rentrée 2024 et accueillant au moins 20 élèves de niveau 6^e et 20 élèves de niveau 5^e

Sources : DEPP, enquête EEC, questionnaires Coordonnateurs, printemps 2025, données finales repondérées ; Évaluations nationales à l'entrée en 6^e 2023, CYCLADES-DNB 2023, Constat de rentrée 2023.

Annexe 3 : Construction de la typologie des établissements

Le choix par certains établissements de mettre en œuvre plusieurs modalités de regroupement ne permet pas d'associer immédiatement chaque collège à une seule catégorie. Pour résoudre ces situations, une typologie spécifique a été définie, permettant de définir une modalité de regroupement unique par établissement et à l'échelle du niveau et de la discipline, pour le français en 6^e, le français en 5^e, les mathématiques en 6^e et les mathématiques en 5^e. Cette typologie s'appuie exclusivement sur les réponses des chefs d'établissement.

Après avoir indiqué qu'au moins une partie des classes de 6^e ou de 5^e étaient concernées par des regroupements d'élèves en français ou en mathématiques, les chefs d'établissement ont renseigné les modalités selon lesquelles ces regroupements d'élèves ont été constitués à partir d'une liste de six modalités, avec la possibilité de renseigner plusieurs modalités pour un même niveau et une même matière.

Les six modalités, telles que présentées aux chefs d'établissement, sont les suivantes :

- « Groupes constitués de manière à ce que chaque groupe accueille des élèves de niveau similaire dans la discipline (groupes homogènes) » ;
- « Groupes constitués de manière à ce que chaque groupe accueille des élèves de niveau similaire en français et en mathématiques » ;
- « Groupes constitués de manière à ce que chaque groupe accueille des élèves de niveau similaire pour une compétence ou un domaine spécifique des programmes » ;
- « Groupes constitués de manière à ce que chaque groupe accueille des élèves de niveaux variés (groupes hétérogènes) » ;
- « Un ou plusieurs groupes constitués des élèves les plus performants dans la matière, et les autres groupes constitués indépendamment du niveau des élèves » ;
- « Un ou plusieurs groupes constitués des élèves identifiés comme rencontrant le plus de difficultés dans la matière, et les autres groupes constitués indépendamment du niveau des élèves ».

Dans les tableaux et les graphiques, ces modalités sont respectivement qualifiées de « groupes homogènes », « groupes selon le niveau scolaire dans les deux matières », « groupes par compétence », « groupes hétérogènes », « groupes constitués des élèves les plus performants » et « groupes constitués des élèves en difficulté ».

La typologie mise en œuvre dans ce rapport repose sur plusieurs règles de gestion, définies dans l'ordre suivant :

1. En cas de regroupement unique pour un niveau et une discipline, par exemple en 6^e pour les enseignements de français, c'est cette modalité de regroupement qui est utilisée.
2. En cas de regroupements multiples faisant intervenir les groupes à destination des élèves les plus performants ou les groupes à destination des élèves rencontrant le plus de difficultés d'une part, et des groupes hétérogènes d'autre part, c'est la modalité de regroupement pour les élèves les plus performants (ou, le cas échéant, pour les élèves rencontrant le plus de difficultés) qui est choisie pour caractériser la modalité de regroupement à l'intersection du niveau et de la discipline. Cela découle de la définition de ces deux modalités de regroupement, qui précisent que les autres groupes sont constitués indépendamment du niveau des élèves.
3. En revanche, en présence à la fois de groupes à destination des élèves les plus performants et de groupes à destination des élèves rencontrant le plus de difficultés et de groupes hétérogènes, alors la modalité retenue est celle des groupes homogènes qui correspond davantage au découpage de la classe de référence en différents groupes à partir du niveau des élèves dans la discipline.

4. En cas de regroupements multiples faisant intervenir des groupes à destination des élèves les plus performants et/ou des groupes à destination des élèves rencontrant le plus de difficultés d'une part, et des groupes homogènes d'autre part, c'est cette dernière modalité qui est retenue. Cela se justifie par le fait qu'il s'agit finalement, à l'échelle du niveau et de la discipline, d'un regroupement des élèves selon leur niveau dans la discipline.
5. Pour les groupes constitués selon le niveau dans les deux disciplines et les groupes constitués selon le niveau pour une compétence ou domaine spécifique : en présence regroupements complémentaires à destination des élèves les plus en difficulté ou des élèves les plus performants, c'est la modalité de constitution plus générale qui est retenue (niveau dans les deux disciplines ou compétence spécifique).
6. En cas de regroupements multiples faisant intervenir des groupes constitués selon le niveau dans les deux disciplines ou des groupes constitués selon le niveau pour une compétence ou un domaine spécifique d'une part et des groupes homogènes (et éventuellement des groupes à destination des élèves les plus en difficulté ou les plus performants, d'après les règles de gestion précédentes) d'autre part, c'est la première modalité qui est retenue (groupes constitués selon le niveau dans les deux disciplines ou pour une compétence spécifique). L'idée derrière ce regroupement est de conserver le plus haut niveau de détail dans la constitution de groupes selon le niveau des élèves.
7. Pour les autres cas de regroupements multiples faisant intervenir plusieurs modalités qui concernent le niveau des élèves sans aucune mention de groupes hétérogènes, c'est la modalité homogène qui est retenue pour son caractère plus général (par exemple en présence d'une déclaration de groupes à destination des élèves les plus en difficulté et de groupes selon le niveau dans les deux disciplines et de groupes constitués par rapport à une compétence ou domaine spécifique).
8. Enfin, pour les chefs d'établissements qui déclarent à la fois des modalités de regroupement homogènes et une modalité hétérogène, la modalité retenue est définie comme « modalités multiples » pour l'intersection du niveau et de la discipline. C'est par exemple le cas pour un établissement qui déclarerait avoir mis en place en 5^e à la fois des groupes hétérogènes et des groupes constitués selon le niveau dans les deux matières pour les enseignements de français.

À partir des modalités uniques ainsi définies par niveau et discipline, il est possible de remonter à l'échelle de l'établissement afin d'associer chaque établissement à un mode de regroupement. Une première approche consiste à associer de manière unique un établissement à une catégorie, tout en restant au plus près des déclarations des chefs d'établissement. Pour cela, si une modalité unique a été choisie pour toutes les classes de 6^e et de 5^e en français et en mathématiques, alors c'est cette modalité qui est retenue. C'est par exemple le cas des établissements ayant défini des groupes hétérogènes pour tous ces élèves. Pour tous les autres cas qui font intervenir au moins deux modalités de regroupement, l'établissement n'est pas associé à une modalité de regroupement en particulier et se retrouve donc dans la catégorie « modalités multiples ». Ici, la modalité « Pas de groupe » est considérée comme une modalité à part entière à l'échelle du niveau et de la discipline. Ainsi, un établissement qui met en place des groupes homogènes en mathématiques et qui ne met pas en place de groupes en français se retrouve dans la catégorie « modalités multiples ».

Une seconde approche consiste à identifier les établissements au sein desquels les différentes modalités ont été mises en place pour au moins une partie des élèves. Pour cela, le point de départ n'est plus la modalité de regroupement unique à l'intersection d'un niveau et d'une discipline, comme définie précédemment, mais la déclaration initiale du chef d'établissement dans le questionnaire. Un établissement peut ainsi se retrouver à la fois parmi les établissements qui ont défini des groupes homogènes et parmi les établissements qui ont regroupé les élèves de manière hétérogène dès lors que ces différentes modalités ont été choisies, quels que soient le niveau et la discipline. Les règles de gestion définies ci-dessus ne s'appliquent pas pour cette seconde approche :

un établissement qui indique regrouper les élèves rencontrant les plus de difficulté et faire des groupes hétérogènes pour les classes de français en 6^e est compté comme expérimentant à la fois les groupes hétérogènes et les groupes à destination des élèves les plus en difficulté. Avec cette seconde approche, les établissements ne sont donc pas associés à une modalité unique de regroupement, la somme des établissements associés à chacune des modalités dépasse l'effectif total des établissements, puisque certains sont comptés dans différentes catégories. Ce découpage permet néanmoins de comparer les caractéristiques des établissements qui ont expérimenté une certaine modalité de regroupement pour au moins une partie des élèves par rapport à ceux qui ne l'ont pas fait.

Annexe 4 : Évolution de la répartition des élèves dans les groupes de performance aux évaluations nationales entre 2024 et 2025

ENCADRE 5 Scores par discipline et groupes de performance

Les résultats des élèves aux évaluations nationales de début d'année sont présentés par domaine ou par discipline (français et mathématiques). Par discipline, les scores sont calculés à partir de modèles psychométriques et les résultats sont ensuite standardisés afin que les évaluations nationales pour un même niveau puissent être comparées à travers le temps. Dans le second degré, les scores sont construits de manière à avoir une moyenne de 250 et un écart-type de 50 lors de la première année de passation. Il s'agit d'une échelle sans valeur normative, et le score de 250 ne correspond pas à un seuil de compétence minimal à atteindre (Ballereau et al., 2025).

Afin de faciliter l'interprétation de ces scores, mais également de se rapprocher des standards internationaux, les élèves sont répartis en six groupes de performance pour chacune des disciplines, selon le score obtenu aux évaluations nationales. Les résultats peuvent ainsi être présentés selon ces groupes ou selon une répartition en trois catégories, comme cela est également fait dans ce document de travail. Pour chaque discipline, le groupe des élèves les moins performants est ainsi composé chaque année des élèves ayant un score inférieur à 225, le groupe de performance moyenne regroupe les élèves avec des scores entre 225 et 275 et le groupe de performance élevée désigne les élèves qui ont obtenu un score supérieur à 275.

Au lieu de présenter les scores moyens, on peut répartir les élèves dans les différents groupes de performance²⁴. Cette perspective reste cohérente avec l'évolution observée des scores moyens (**Figure 30**). En effet, la faible diminution des scores de français de 4^e entre septembre 2024 et 2025 se traduit par une légère hausse de la proportion d'élèves situés dans le groupe de performance faible en 2025, et par une très faible baisse des proportions d'élèves situés dans les groupes de performance élevée ou moyenne. À l'inverse, en mathématiques, la légère augmentation du score entre septembre 2024 et 2025 s'accompagne d'une faible augmentation de la proportion d'élèves dans le groupe de performance moyenne ou élevée entre ces deux années.

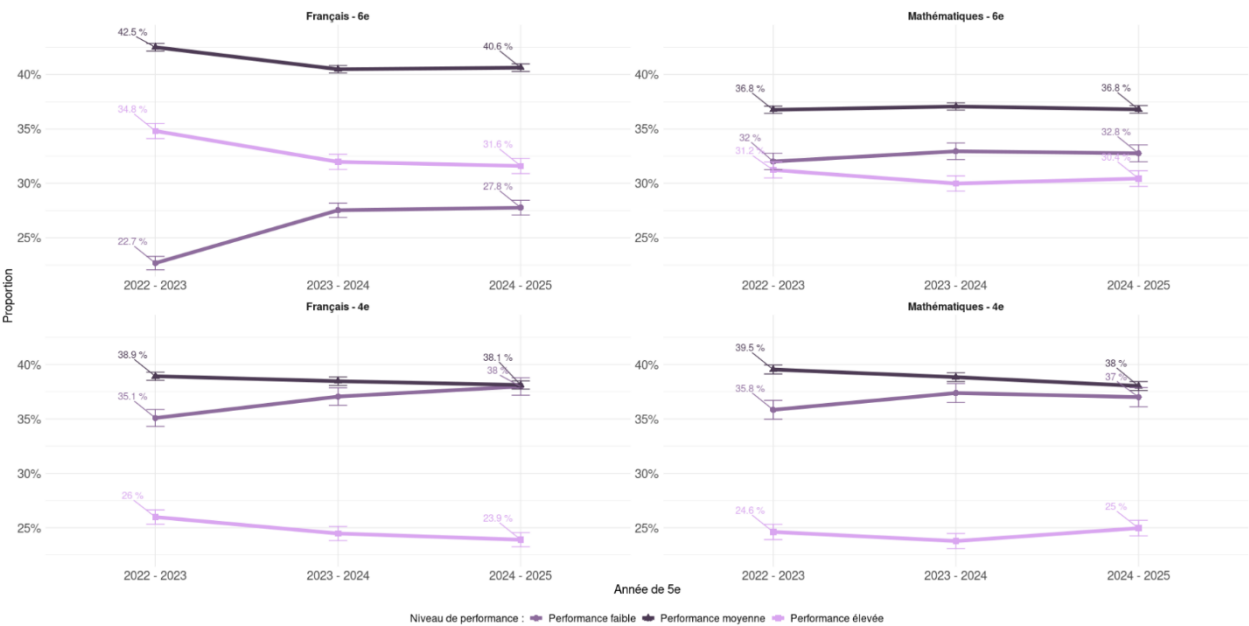
Comme pour les résultats moyens, il est intéressant de décliner ces évolutions selon le type de regroupement mis en place, afin d'identifier des variations qui pourraient éventuellement suggérer un effet de la réforme (**Figures 31 & 32**).

Comme pour les scores moyens, on observe peu de différences majeures entre les types d'établissement dans l'évolution, entre septembre 2024 et septembre 2025, de la répartition des élèves en groupes de performance définis à partir de leurs scores aux évaluations nationales. En français, une légère progression de la part des élèves les plus faibles entre septembre 2024 et 2025 semble être compensée par une diminution de la part des élèves de performance élevée. À nouveau, le profil le plus atypique est celui des établissements qui n'ont pas mis en place de regroupements d'élèves : la progression de la part des élèves dans le groupe de performance faible y a été plus marquée entre septembre 2024 et 2025 qu'entre septembre 2023 et 2024.

En mathématiques, aucune évolution significative n'est observée pour les établissements qui ont mis en place des regroupements hétérogènes lors de l'année scolaire 2024-2025. Au contraire, pour les établissements qui ont construit des groupes selon le profil scolaire des élèves, une diminution de la part d'élèves dans le groupe de performance moyenne peut être observée, ainsi qu'une progression de la part du groupe de performance élevée entre septembre 2024 et 2025. Enfin, pour les établissements qui n'ont pas mis en place de regroupement d'élèves, seule la progression de la part d'élèves dans le groupe de performance élevée est significative entre ces deux années.

²⁴ Afin d'apporter une information plus détaillée sur le niveau des élèves, les résultats aux évaluations nationales sont également proposés par groupe de performance. La distribution des élèves est présentée dans six groupes selon les compétences maîtrisées, regroupant les moins performants (groupe 1 - score inférieur à 200 et groupe 2 - score compris entre 200 et 225), ceux de performance moyenne (groupe 3 - score compris entre 225 et 250 et groupe 4 - score compris entre 250 et 275) et les plus performants (groupe 5 - score compris entre 275 et 300 et groupe 6 - score supérieur ou égal à 300). Pour plus d'informations voir Euzéby et al. (2025).

FIGURE 30 • Évolution de la répartition des élèves de 5^e des établissements de l'échantillon dans les groupes de performance aux évaluations nationales de 6^e et de 4^e selon leur année de scolarisation en 5^e



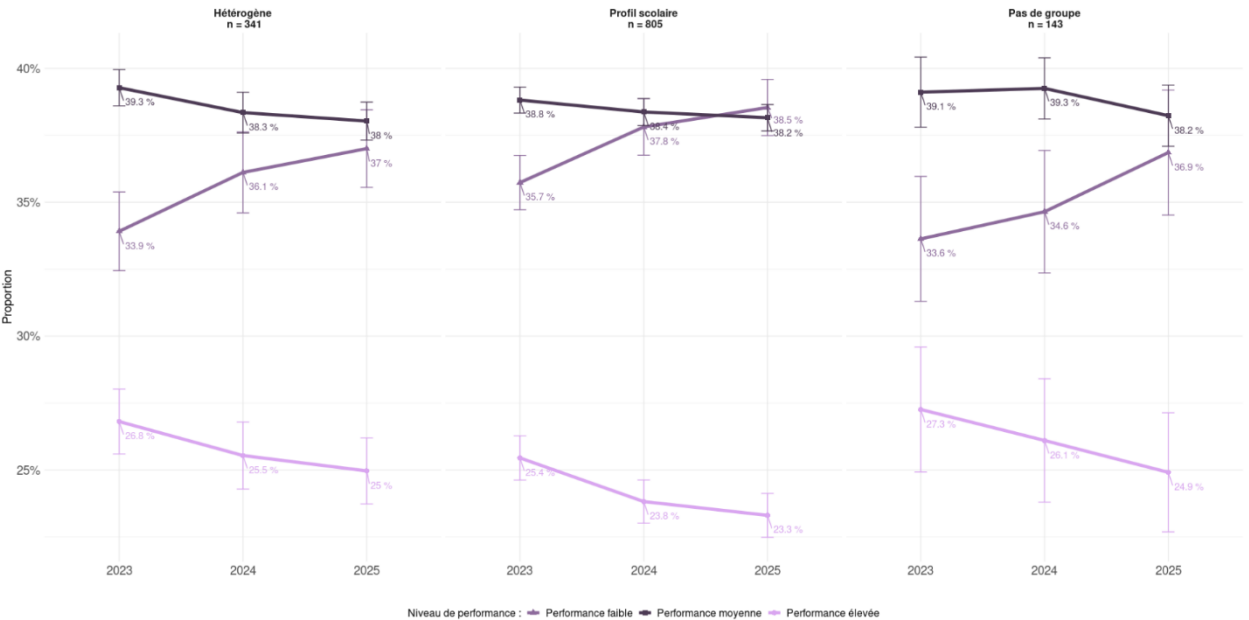
Lecture : Parmi les élèves scolarisés en 5e au cours de l'année 2022-2023 au sein d'un établissement qui met en place une seule ou aucune forme de regroupement au cours de l'année 2024-2025, 35,1 % appartiennent au groupe de performance faible à leur entrée en 4e. Cette proportion est de 22,7 % pour la 6e.

Note : Les intervalles de confiance à 95 % sont affichés sur le graphique.

Champ : Élèves scolarisés en 5^e au sein des établissements de l'enquête EEC ayant adopté une forme unique de regroupement ou aucune en 5^e pour la discipline. Cela correspond à 1 289 collèges pour le français et à 1 291 collèges pour les mathématiques.

Sources : DEPP, Évaluations nationales en français et mathématiques de 6^e (2021, 2022, 2023) et de 4^e (2023, 2024, 2025) ; enquête EEC, questionnaires Chefs d'établissement, printemps 2025 ; Constat de rentrée 2025.

FIGURE 31 • Répartition des élèves dans les groupes de performance aux évaluations de français de 4^e selon le type de regroupement mis en œuvre



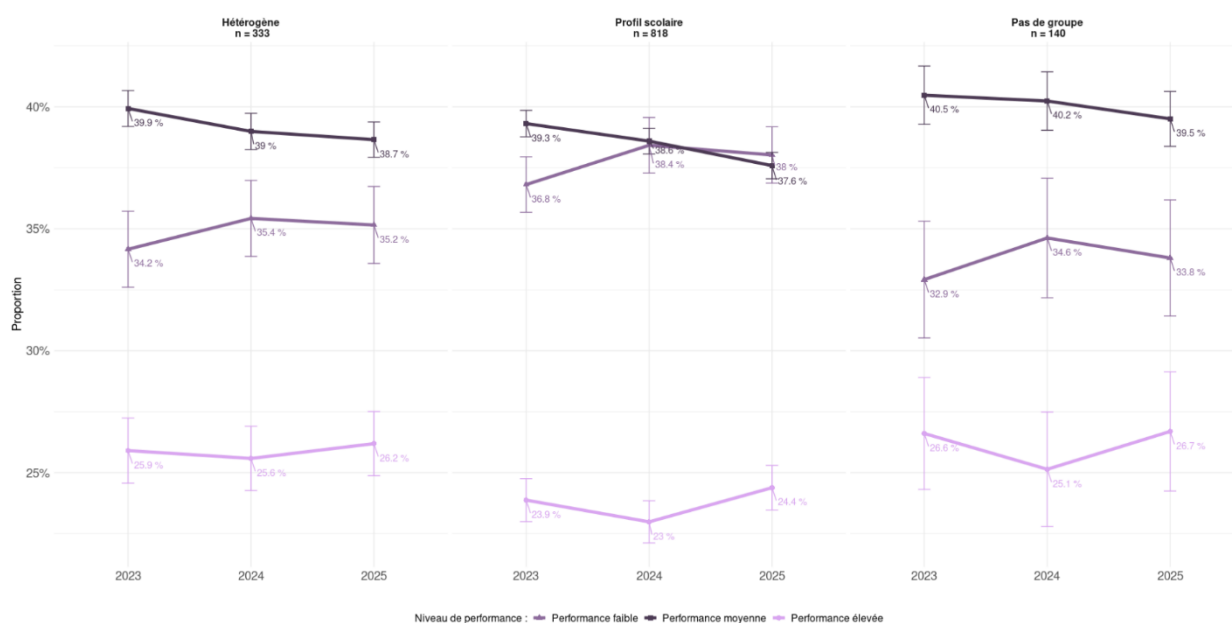
Lecture : Parmi les élèves scolarisés en 5e au cours de l'année 2022-2023 au sein d'établissements qui mettent en place des regroupements hétérogènes au cours de l'année scolaire 2024-2025 et qui passent leurs évaluations de début de 4^e en 2023, 26,8 % appartiennent au groupe de performance élevée en français. Cette proportion est de 27,3 % dans les collèges qui ne mettent en place aucune forme de regroupement au cours de l'année scolaire 2024-2025.

Note : Les intervalles de confiance à 95 % sont affichés sur le graphique.

Champ : Élèves scolarisés en 5^e au sein des établissements de l'enquête EEC ayant adopté une forme unique de regroupement ou aucune en 5^e pour le français. Cela correspond à 1 289 collèges.

Sources : DEPP, Évaluations nationales en français de 4^e (2023, 2024, 2025) ; enquête EEC, questionnaires Chefs d'établissement, printemps 2025 ; Constat de rentrée 2025.

FIGURE 32 • Répartition des élèves dans les groupes de performance aux évaluations de mathématiques de 4^e selon le type de regroupement mis en œuvre



Lecture : Parmi les élèves scolarisés en 5e au cours de l'année 2022-2023 au sein d'établissements qui mettent en place des regroupements hétérogènes au cours de l'année scolaire 2024-2025 et qui passent leurs évaluations de début de 4e en 2023, 25,9 % appartiennent à la catégorie de performance élevée en mathématiques. Cette proportion est de 26,6 % dans les collèges qui ne mettent en place aucune forme de regroupement au cours de l'année scolaire 2024-2025.

Note : Les intervalles de confiance à 95 % sont affichés sur le graphique.

Champ : Éléves scolarisés en 5^e au sein des établissements de l'enquête EEC ayant adopté une forme unique de regroupement ou aucune en 5^e pour les mathématiques. Cela correspond à 1 291 collèges.

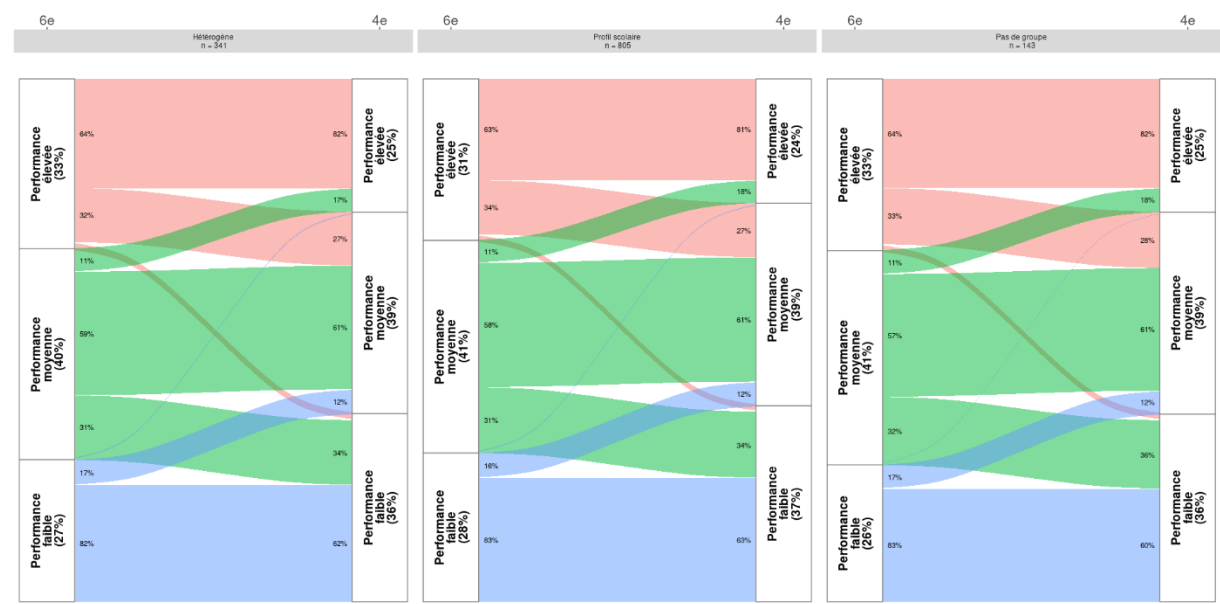
Sources : DEPP, Évaluations nationales en mathématiques de 4^e (2023, 2024, 2025) ; enquête EEC, questionnaires Chefs d'établissement, printemps 2025 ; Constat de rentrée 2025.

Annexe 5 : Évolution de la répartition des élèves dans les groupes de performance aux évaluations nationales entre la 6^e et la 4^e, selon le type de regroupement mis en œuvre en 2024-2025

Au-delà de la comparaison des scores moyens obtenus aux évaluations nationales, il est possible de mesurer la progression des élèves en observant leurs éventuels changements de groupes entre les évaluations de 6^e et celles de 4^e, selon les organisations pédagogiques auxquelles ils ont été exposés en 5^e.

En français et en mathématiques, les élèves connaissent des dynamiques de progression similaires quel que soit le type de regroupement auquel ils sont exposés en 5^e. Ainsi, en français, pour tous les types de regroupement, environ 80 % des élèves de niveau initial faible en 6^e sont toujours dans le groupe de performance faible aux évaluations nationales de 4^e (**Figure 33**). Ce taux de persistance est légèrement inférieur à 60 % pour les élèves de niveau initial moyen et légèrement supérieur à 60 % pour les élèves de niveau initial élevé. Dans ces deux derniers cas, environ 30 % des élèves sont en 4^e dans un groupe de performance inférieure à celui auquel ils appartenaient en 6^e (performance faible pour la performance initiale moyenne, performance moyenne pour la performance initiale élevée).

FIGURE 33 • Mobilité entre groupes de performance aux évaluations nationales de français entre la 6^e et la 4^e, selon le type de regroupement mis en œuvre en 5^e en 2024-2025



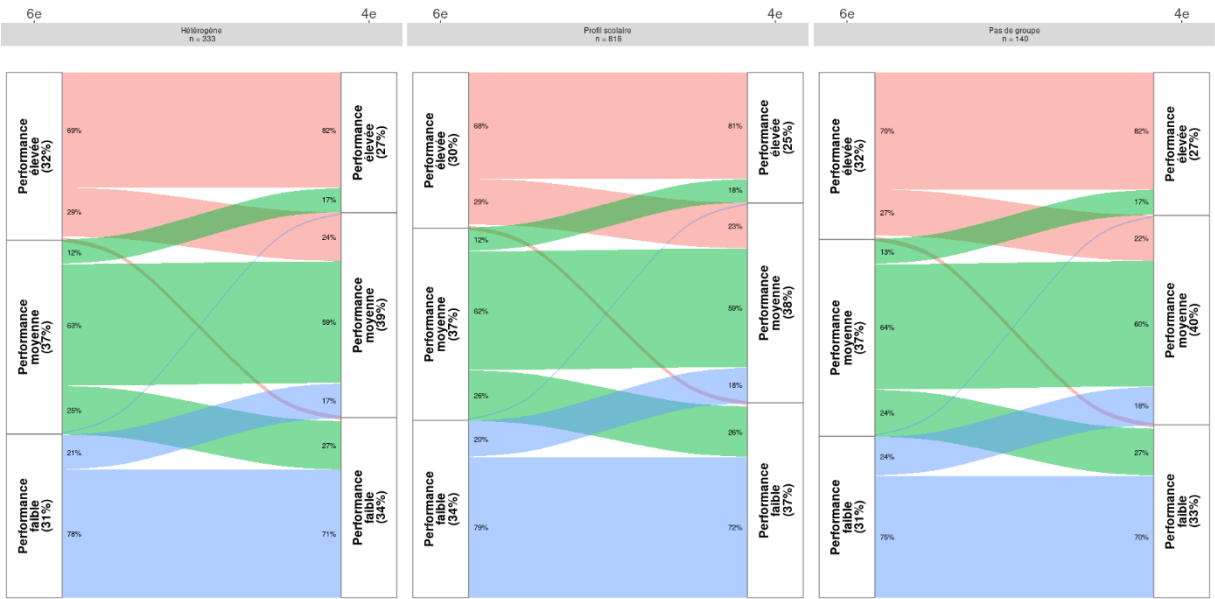
Lecture : Parmi les élèves qui appartiennent de niveau initial élevé en 6^e pour le français, qui effectuent leur 5^e en 2024-2025 au sein des établissements ayant mis en place des regroupements hétérogènes en 5^e en français et qui ont passé les évaluations nationales de début de 4^e en 2025, 64 % ont toujours un niveau de performance élevé en début de 4^e pour le français. 82 % des élèves de niveau élevé en français en 4^e avaient déjà des performances élevées à l'entrée en 6^e. Ces proportions sont identiques pour les élèves des établissements qui n'ont mis en place aucune forme de regroupement, et inférieures d'un point de pourcentage pour les élèves des établissements qui ont constitué des groupes selon le profil scolaire des élèves (respectivement 63 % et 81 %).

Champ : Élèves scolarisés en 5^e au cours de l'année scolaire 2024-2025 au sein des établissements de l'enquête EEC ayant adopté une forme unique de regroupement ou aucune en 5^e pour le français. Cela correspond à 1 289 collèves. Seuls les élèves retrouvés à la fois dans les évaluations nationales de français de début de 6^e et dans celles de début de 4^e sont pris en compte dans l'analyse.

Sources : DEPP, Évaluations nationales en français de 6^e de 2023 et de 4^e de 2025 ; enquête EEC, questionnaires Chefs d'établissement, printemps 2025 ; Constat de rentrée 2025.

En mathématiques, la proportion d'élèves qui conservent des performances faibles entre la 6^e et la 4^e est légèrement inférieure à celle observée en français, et varie autour de 75 % selon les types de regroupement (**Figure 34**). Ce taux de persistance est juste en-dessous de 65 % pour les élèves de niveau initial moyen et légèrement supérieur à 65 % pour ceux de niveau initial élevé. Ces taux varient peu selon les différents types de regroupement mis en œuvre.

FIGURE 34 • Mobilité entre groupes de performance aux évaluations nationales de mathématiques entre la 6^e et la 4^e, selon le type de regroupement mis en œuvre en 5^e en 2024-2025



Lecture : Parmi les élèves de niveau initial élevé en 6^e pour les mathématiques, qui effectuent leur 5^e en 2024-2025 au sein des établissements qui mettent en place des regroupements hétérogènes en 5^e en mathématiques et qui passent leurs évaluations nationales de début de 4^e en 2025, 69 % appartiennent toujours à la catégorie de performance élevée en début de 4^e. Les élèves qui étaient déjà dans la catégorie de performance élevée en 6^e représentent 82 % des élèves de ce même groupe à l’entrée en 4^e. Ces proportions sont respectivement de 68 % et 81 % pour les élèves des établissements qui constituent des groupes selon le profil scolaire des élèves et de 70 % et 82 % pour les élèves des établissements qui ne mettent en place aucune forme de regroupement.

Champ : Élèves scolarisés en 5^e au cours de l’année scolaire 2024-2025 au sein des établissements de l’enquête EEC ayant adopté une forme unique de regroupement ou aucune en 5^e pour les mathématiques. Cela correspond à 1 291 collèges. Seuls les élèves retrouvés à la fois dans les évaluations nationales de mathématiques de début de 6^e et dans celles de début de 4^e ont été pris en compte dans l’analyse.

Sources : DEPP, Évaluations nationales en mathématiques de 6^e (2021, 2022, 2023) et de 4^e (2023, 2024, 2025) ; enquête EEC, questionnaires Chefs d’établissement, printemps 2025 ; Constat de rentrée 2025.

Ainsi, l’analyse ne permet pas de faire ressortir d’évolution majeure dans les mobilités entre groupes de performance aux évaluations nationales entre la 6^e et la 4^e selon la nature des regroupements mis en place dans le cadre du dispositif. Les performances des élèves apparaissent au contraire relativement stables dans le temps.

Annexe 6 : Données des évaluations nationales et scores standardisés

Afin de mesurer l'évolution des performances scolaires des élèves après la mise en place de la réforme, les données des évaluations nationales ont été mobilisées²⁵. Pour les analyses corrélationnelles et la première approche par différence de différences, l'ensemble des évaluations de début de 4^e disponibles ont été mobilisées : il s'agit des années 2023, 2024 et 2025. Pour bien identifier les éventuels effets associés aux différentes formes de regroupement, le traitement est défini par rapport à l'année de 5^e : l'échantillon d'analyse correspond ainsi aux élèves scolarisés en 5^e au sein des établissements de l'enquête. Par cohérence, la même approche a été adoptée pour les années précédentes : les élèves qui ont effectué leur année de 5^e en 2024-2025 et qui ont passé leurs évaluations nationales de début de 4^e en 2025 sont comparés aux élèves ayant effectué leur 5^e en 2023-2024 et qui ont passé leurs évaluations nationales de début de 4^e en septembre 2024. Cette étude ne prend donc pas en compte les élèves qui redoublent leur 4^e et passent leurs évaluations nationales de début de 4^e en septembre 2023, 2024 ou 2025.

La seconde approche par doubles différences consiste à suivre la cohorte des élèves qui sont exposés à la réforme en 5^e au cours de l'année scolaire 2024-2025 et qui passent les évaluations nationales de 4^e à la rentrée 2025. Pour cela, toutes les évaluations nationales disponibles pour cette cohorte, à l'exception des évaluations de début de 5^e (**Encadré 6**), sont exploitées : les évaluations de CP de la rentrée 2018, le point d'étape de CP en janvier 2019, les évaluations de début de CE1 en 2019, les évaluations de début de 6^e en 2023 et les évaluations de début de 4^e en 2025. Les élèves qui redoublent une fois entre le CP et la 4^e sortent ainsi du panel.

ENCADRE 6 Les évaluations nationales de début de 5^e

Les évaluations nationales de 5^e ont été généralisées à partir de la rentrée 2025. Ces données sont donc disponibles pour l'ensemble des élèves ayant effectué leur 6^e en 2024-2025, et, à ce titre, potentiellement concernées par la mise en œuvre des nouvelles organisations pédagogiques. Il n'y a pas assez de profondeur temporelle pour mettre en œuvre la première approche par doubles différences qui repose sur la comparaison des résultats d'une cohorte par rapport aux résultats des cohortes précédentes.

En revanche, il est théoriquement possible de mobiliser la seconde approche par doubles différences, qui repose sur une mesure de la progression des élèves entre la 6^e à la 5^e. Cependant, les évaluations de 5^e n'évaluent pas les mêmes domaines que les évaluations passées lors des années précédentes. En effet, le domaine « compétences langagières et linguistiques » en français ne se retrouve que pour la 5^e, et le domaine « automatismes » n'existe que pour le second degré. Seul le domaine « fluence » se retrouve depuis le point d'étape à mi-CP. Il serait ainsi possible d'analyser l'effet des regroupements d'élèves sur la fluence, mais pas sur les autres domaines, et pas au niveau de la discipline. L'analyse de la fluence en 5^e n'est pas développée dans ce document de travail mais pourra faire l'objet de travaux futurs.

Pour plus d'informations sur les évaluations de 5^e, voir de Monestrol et al., 2026.

Quelques traitements préliminaires ont été réalisés afin de garantir la qualité du travail produit à partir de ces sources de données. Dans un premier temps, les élèves qui ne pouvaient pas être uniquement associés à un seul score ont été retirés de l'analyse : il s'agit en premier lieu des lignes sans identifiant d'élève et des doublons associés à un même identifiant. Dans ce dernier cas, toutes les lignes doublonnées sont retirées. Cela représente entre 0,1 % et 0,4 % des lignes des fichiers des évaluations nationales chaque année. La proportion de lignes sans identifiant unique est particulièrement importante pour les évaluations nationales de 2024 et atteint 1,02 % de l'ensemble des lignes de la base de données. Les identifiants individuels mobilisés dans cette étude sont des clés uniques créées à partir d'un cryptage des identifiants nationaux (INE) afin de garantir l'anonymat des élèves.

²⁵ Pour plus d'informations sur ces données (description, dictionnaires), voir le [catalogue des données de la DEPP](https://catalogue.depp.education.fr) (<https://catalogue.depp.education.fr>).

Pour les analyses causales, l'établissement de traitement est le collège fréquenté au cours de la 5^e, pour les élèves évalués en début de 4^e en 2025. Pour chaque élève, le collège fréquenté en 5^e a été identifié grâce aux données du constat de rentrée de l'année de 4^e, à partir de la variable indiquant l'établissement fréquenté l'année précédente. Certains élèves n'ont pas été retrouvés de cette manière : par exemple, parmi les élèves qui passent les évaluations de 4^e à la rentrée 2025, 3,3 % n'ont pas été retrouvés dans le constat et n'ont donc pas d'établissement de 5^e. Ce taux est cohérent avec les taux de redoublement nationaux qui sont égaux à 2,6 % pour le passage entre la 5^e et la 4^e en 2025 (sachant que les taux de redoublement sont plus élevés dans le secteur public, duquel provient l'intégralité des collèges de l'étude). Inversement, moins de 1 % des élèves qui ont un établissement de 5^e et qui sont retrouvés en 4^e ne sont pas présents dans les bases des évaluations nationales. Cela peut être lié à une absence lors des évaluations nationales ou à la présence d'identifiants ou de scores doublonnés.

Les scores aux évaluations nationales par discipline dans le second degré sont standardisés de manière à ce que le score moyen la première année soit égal à 250, avec un écart-type de 50. Pour la première approche par doubles différences, qui compare les scores aux évaluations nationales de début de 4^e, les scores de chaque année ont été standardisés de manière à ce que la moyenne des établissements qui n'ont pas mis en place de groupes soit égale à 0 chaque année et ait un écart-type égal à 1. Cette échelle, qui permet d'apprécier les éventuels effets de la politique en pourcentage d'écart-type, est couramment mobilisée en économie de l'éducation. Elle offre des estimations dont l'ampleur est facilement interprétable, et permet de comparer la magnitude des effets à d'autres politiques publiques alternatives (Kraft, 2020).

Pour chaque discipline et pour chaque année, le score standardisé correspond au score de l'élève diminué du score moyen des élèves du groupe témoin (qui correspond aux établissements sans regroupement d'élèves) puis divisé par l'écart-type des scores de ce même groupe :

$$Score_std_{i,d,n} = \frac{Score_{i,d,n} - \overline{Score}_{contrôle,d,n}}{\sqrt{Var(Score_{contrôle,d,n})}} \quad (8)$$

Où :

- $Score_{i,d,n}$ est le score obtenu par l'élève i pour la discipline d et l'année n .
- $\overline{Score}_{contrôle,d,n}$ est le score moyen obtenu pour la discipline d durant l'année n par les élèves qui fréquentent en 5^e un établissement qui ne met pas en place de regroupements d'élèves.
- $\sqrt{Var(Score_{contrôle,d,n})}$ est l'écart-type des scores obtenus pour la discipline d durant l'année n par les élèves qui fréquentent en 5^e un établissement qui ne met pas en place de regroupements d'élèves.

La seconde approche par doubles-différences implique le suivi des mêmes élèves dans le temps et la comparaison de leurs scores d'un niveau à l'autre, de l'entrée en CP aux évaluations nationales de début de 4^e. Les scores aux évaluations nationales sont construits pour être comparables d'une année à l'autre pour un niveau donné, mais pas pour être comparés d'un niveau à l'autre. Ainsi, les scores aux évaluations de 6^e en 2024 sont comparables aux scores de début de 6^e en 2025, ce qui permet de mesurer les différences de performances entre cohortes. En revanche, les scores aux évaluations nationales de 6^e ne sont pas comparables aux évaluations nationales de 5^e ou de 4^e : il n'est ainsi pas directement possible d'obtenir une mesure de la progression d'un même élève d'un niveau à l'autre.

Par ailleurs, contrairement au second degré, les évaluations du second degré ne donnent pas lieu au calcul d'un score synthétique par discipline : les données disponibles correspondent aux scores obtenus pour chacune des compétences évaluées, qui sont ensuite associées à un domaine, lui-

même rattaché à une discipline. Suivre les élèves dans le temps implique donc de construire des scores standardisés par domaine et par discipline, qui soient cohérents du CP à la 4^e.

Afin de garantir cette cohérence, seuls les domaines présents dans les évaluations de CP, de mi-CP, de CE1, de 6^e et de 4^e sont retenus. Il s'agit de « compréhension de l'oral », « compréhension de l'écrit » et « fluence » pour le français, et de « nombres et calculs » et « résolution des problèmes » pour les mathématiques. Les compétences évaluées au sein de chacun des domaines peuvent être retrouvées dans les documents de travail de la DEPP, en particulier dans ceux sur les points d'étape en CP. La fluence n'est évaluée qu'à partir de mi-CP, mais elle a été conservée en raison de son caractère particulièrement représentatif du niveau des élèves en français. Par ailleurs, les valeurs extrêmes en fluence ont été censurées au 995^e millime, c'est-à-dire que toutes les valeurs dépassant ce seuil ont été définies comme égales à ce seuil.

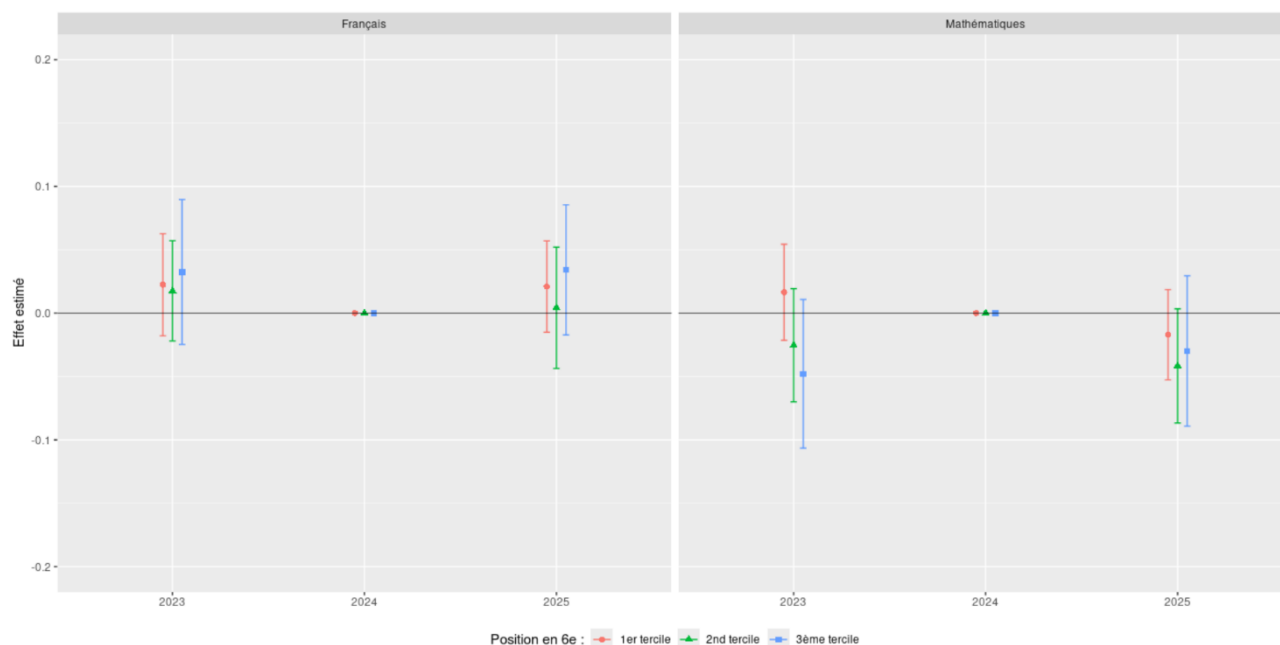
Des scores de discipline sont ensuite construits pour chaque année à partir des scores dans les différents domaines appartenant à la discipline. Comme pour la première approche de différences-de-différences, les scores relatifs aux compétences de chaque domaine sont d'abord standardisés en retranchant la moyenne des scores des élèves scolarisés en 5^e au sein des établissements qui n'ont mis en place aucune forme de regroupement au cours de l'année 2024-2025 et en divisant la grandeur ainsi obtenue par l'écart-type des scores obtenus par ces mêmes élèves (**Équation 8**, avec d la compétence associée à un domaine). Cette procédure permet de mettre toutes les évaluations de compétences sur la même échelle, le barème et les notes maximales variant d'une compétence à l'autre. Ainsi, lors de l'agrégation des scores de chacune des compétences par domaine ou par discipline, la standardisation permet d'éviter de donner plus d'importance à l'une des compétences ou des domaines. Pour chacun des domaines, la moyenne des scores standardisés pour les différentes compétences est calculée afin de définir un score standardisé par domaine. Afin d'assurer la comparabilité des scores entre les élèves, le score par domaine n'est calculé que pour ceux qui ont un score dans chacune des compétences du domaine. Ces scores sont standardisés à nouveau selon les scores standardisés par domaine des élèves du groupe témoin (**Équation 8**, avec d le domaine). Enfin, la procédure est répétée pour définir des scores au niveau de la discipline, eux-mêmes standardisés par rapport aux résultats du groupe témoin (**Équation 8**, avec d la discipline). À nouveau, le score par discipline n'est calculé que pour les élèves qui ont un score dans chacun des domaines associés à la discipline. Le nombre d'élèves considérés dans les estimations peut varier d'un niveau à l'autre, car l'estimation nécessite la présence de l'ensemble des scores dans les compétences et domaines considérés dans l'analyse. Pour étudier la sensibilité des estimations à l'échantillon retenu, une variante d'analyse consiste à définir un échantillon parfaitement stable - dit « complet » - constitué de l'ensemble des élèves dont les résultats à l'intégralité des scores sont présents pour chaque niveau. Ces analyses, qui aboutissent à un taux d'attrition particulièrement élevé (seuls 42 % des élèves des collèges de l'échantillon vérifient ces critères pour le français, et 57 % pour les mathématiques), donnent des résultats d'estimation qualitativement similaires à ceux obtenus par l'approche plus flexible considérée dans ce rapport.

Par construction, pour chaque niveau, la moyenne des élèves du groupe témoin est égale à 0. Ainsi, pour un même niveau, le score standardisé d'un élève indique à quel point cet élève s'éloigne de la moyenne du groupe témoin : si le score est positif, l'élève fait mieux que la moyenne du groupe témoin, sinon il fait moins bien. Pour un même élève au cours du temps, la comparaison de ces scores standardisés permet de mesurer de combien un élève s'est approché ou s'est éloigné de la moyenne du groupe témoin entre deux années. Elle ne permet cependant pas de dire si c'est la performance individuelle de l'élève qui a changé ou si c'est la performance moyenne du groupe témoin qui a évolué. Cette procédure de standardisation permet ainsi de comparer les différentes évaluations à la fois pour un même niveau entre les élèves et entre les niveaux pour un même élève.

Annexe 7 : Hétérogénéité des résultats selon le niveau des élèves à l'entrée en 6^e et leur sexe

Première approche

FIGURE 35 • Différence de scores aux évaluations nationales en français et mathématiques à l'entrée en 4^e avant et après la mise en place de la réforme (comparaison : groupes hétérogènes/pas de regroupement), selon le niveau initial à l'entrée en 6^e



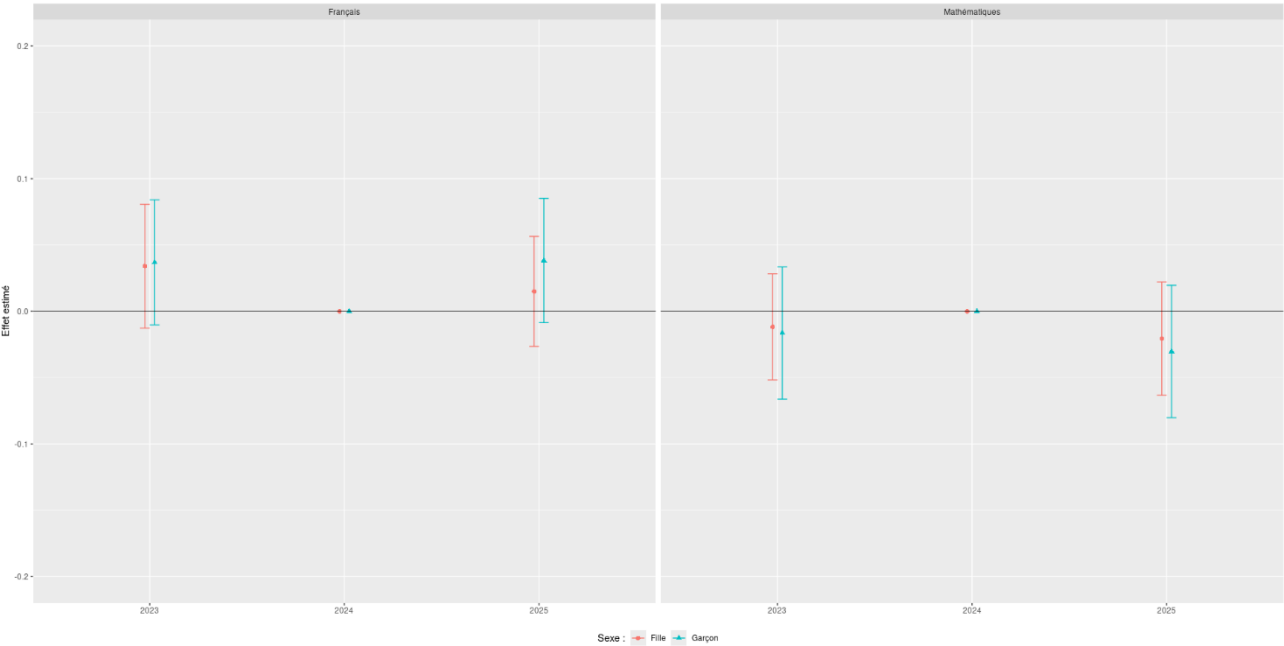
Lecture : En français, les coefficients de l'analyse événementielle pour les évaluations de septembre 2023 sont égaux à 0,02 pour les élèves qui appartenaient au premier tertile de leur établissement pour les scores de français aux évaluations de 6^e, à 0,02 pour les élèves du second tertile et à 0,03 pour les élèves du troisième tertile, c'est-à-dire ceux qui avaient les meilleurs résultats au sein de l'établissement où ils ont passé leur évaluation de début de 6^e. Aucun de ces coefficients n'est significatif au seuil de 5 % et le coefficient de chacun des tertiles appartient aux intervalles de confiance des deux autres. L'évolution de l'écart sur les scores de français entre les élèves qui effectuent leur 5^e dans un établissement qui met en place des groupes hétérogènes et ceux des établissements sans aucune forme de regroupement, entre 2023 et 2024, ne varie donc pas selon le niveau de l'élève en français en 6^e, regroupé en tertiles.

Note : Pour chaque année, les scores sont standardisés par rapport aux scores des élèves scolarisés en 5^e dans les établissements qui ne mettent en place aucune forme de regroupement au cours de l'année scolaire 2024-2025. Les erreurs-types sont clusterisées au niveau de l'établissement fréquenté en 5^e. Les intervalles de confiance à 95 % sont affichés sur le graphique. L'année 2024 est l'année de référence.

Champ : Élèves scolarisés en 5^e au sein des établissements de l'enquête EEC ayant mis en place des regroupements hétérogènes ou aucune forme de regroupement en 5^e pour la discipline. Cela correspond à 484 collèges pour le français et 473 collèges pour les mathématiques.

Sources : DEPP, Évaluations nationales en français et mathématiques de 4^e (2023, 2024, 2025) ; Évaluations nationales en français et mathématiques de 6^e (2021, 2022, 2023) ; enquête EEC, questionnaires Chefs d'établissement, printemps 2025 ; Constat de rentrée 2025.

FIGURE 36 • Différence de scores aux évaluations nationales en français et mathématiques à l'entrée en 4^e avant et après la mise en place de la réforme (comparaison : groupes hétérogènes/pas de regroupement), selon le sexe



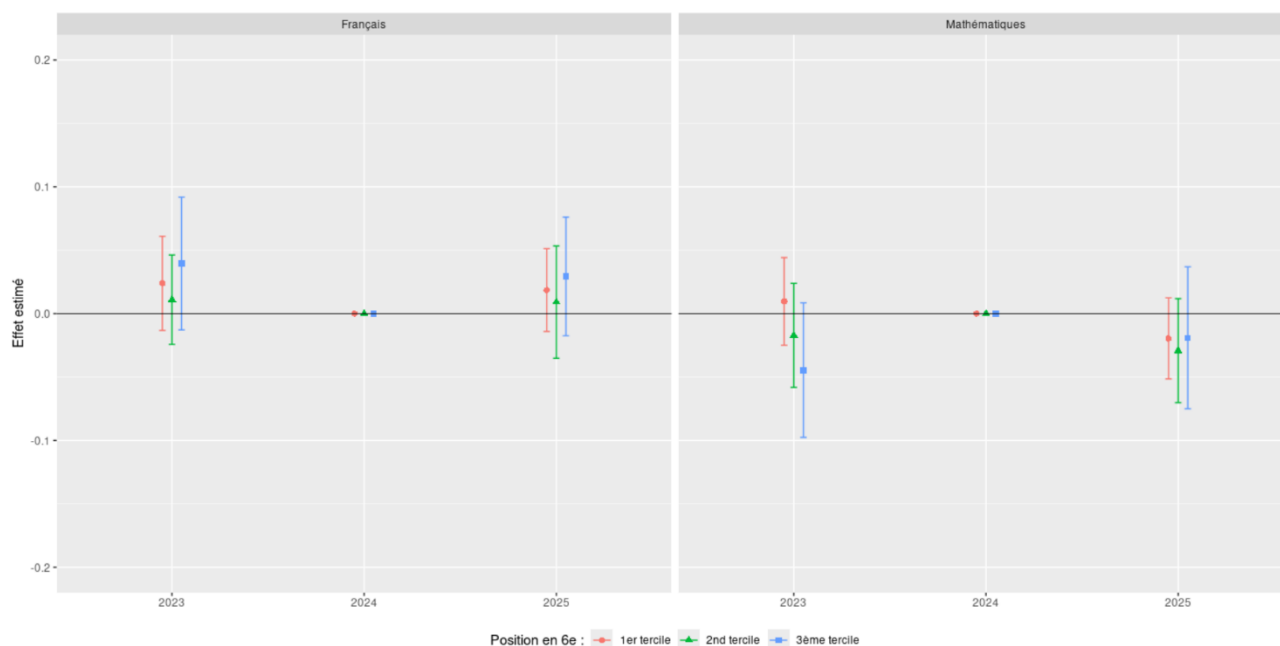
Lecture : En français, les coefficients de l’analyse événementielle pour les évaluations de septembre 2023 sont égaux à 0,03 pour les filles et à 0,04 pour les garçons. Aucun de ces coefficients n’est significatif au seuil de 5 % et le coefficient associé à chacun des sexes appartient à l’intervalle de confiance de l’autre. L’évolution de l’écart sur les scores de français entre les élèves qui effectuent leur 5^e dans un établissement qui met en place des groupes hétérogènes et ceux des établissements sans aucune forme de regroupement, entre 2023 et 2024, ne varie donc pas selon le sexe de l’élève.

Note : Pour chaque année, les scores sont standardisés par rapport aux scores des élèves scolarisés en 5^e dans les établissements qui ne mettent en place aucune forme de regroupement au cours de l’année scolaire 2024-2025. Les erreurs-types sont clusterisées au niveau de l’établissement fréquenté en 5^e. Les intervalles de confiance à 95 % sont affichés sur le graphique. L’année 2024 est l’année de référence.

Champ : Élèves scolarisés en 5^e au sein des établissements de l’enquête EEC ayant mis en place des regroupements hétérogènes ou aucune forme de regroupement en 5^e pour la discipline. Cela correspond à 484 collèges pour le français et 473 collèges pour les mathématiques.

Sources : DEPP, Évaluations nationales en français et mathématiques de 4^e (2023, 2024, 2025) ; enquête EEC, questionnaires Chefs d’établissement, printemps 2025 ; Constat de rentrée 2025.

FIGURE 37 • Différence de scores aux évaluations nationales en français et mathématiques à l'entrée en 4^e avant et après la mise en place de la réforme (comparaison : groupes composés selon le profil scolaire des élèves/pas de regroupement), selon le niveau initial à l'entrée en 6^e



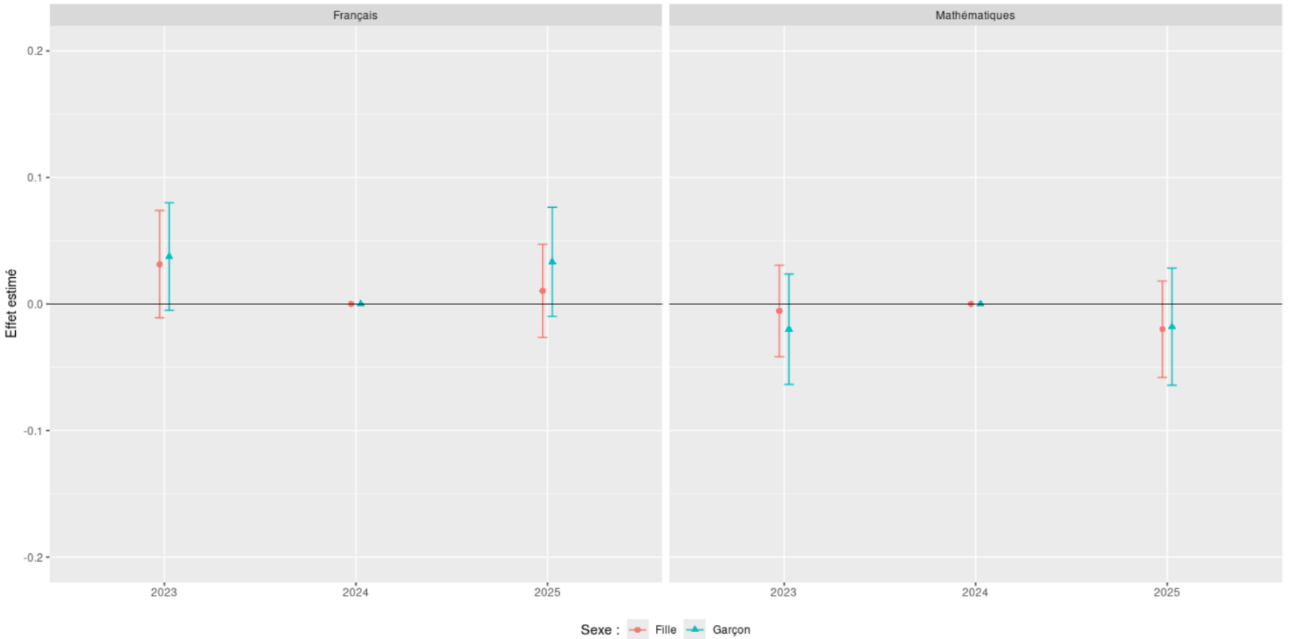
Lecture : En français, les coefficients de l'analyse événementielle pour les évaluations de septembre 2023 sont égaux à 0,02 pour les élèves qui appartenaient au premier tercile de leur établissement pour les scores de français aux évaluations de 6^e, à 0,01 pour les élèves du second tercile et à 0,04 pour les élèves du troisième tercile, c'est-à-dire ceux qui avaient les meilleurs résultats au sein de l'établissement où ils ont passé leur évaluation de début de 6^e. Aucun de ces coefficients n'est significatif au seuil de 5 % et le coefficient de chacun des terciles appartient aux intervalles de confiance des deux autres. L'évolution de l'écart sur les scores de français entre les élèves qui effectuent leur 5^e dans un établissement qui met en place des groupes constitués selon le profil scolaire des élèves et ceux des établissements sans aucune forme de regroupement, entre 2023 et 2024, ne varie donc pas selon le niveau de l'élève en français en 6^e, regroupé en terciles.

Note : Pour chaque année, les scores sont standardisés par rapport aux scores des élèves scolarisés en 5^e dans les établissements qui ne mettent en place aucune forme de regroupement au cours de l'année scolaire 2024-2025. Les erreurs-types sont clusterisées au niveau de l'établissement fréquenté en 5^e. Les intervalles de confiance à 95 % sont affichés sur le graphique. L'année 2024 est l'année de référence.

Champ : Élèves scolarisés en 5^e au sein des établissements de l'enquête EEC ayant mis en place des regroupements selon le profil scolaire des élèves ou aucune forme de regroupement en 5^e pour la discipline. Cela correspond à 948 collèges pour le français et 958 collèges pour les mathématiques.

Sources : DEPP, Évaluations nationales en français et mathématiques de 4^e (2023, 2024, 2025) ; Évaluations nationales en français et mathématiques de 6^e (2021, 2022, 2023) ; enquête EEC, questionnaires Chefs d'établissement, printemps 2025 ; Constat de rentrée 2025.

FIGURE 38 • Différence de scores aux évaluations nationales en français et mathématiques à l'entrée en 4^e avant et après la mise en place de la réforme (comparaison : groupes composés selon le profil scolaire des élèves/pas de regroupement), selon le sexe



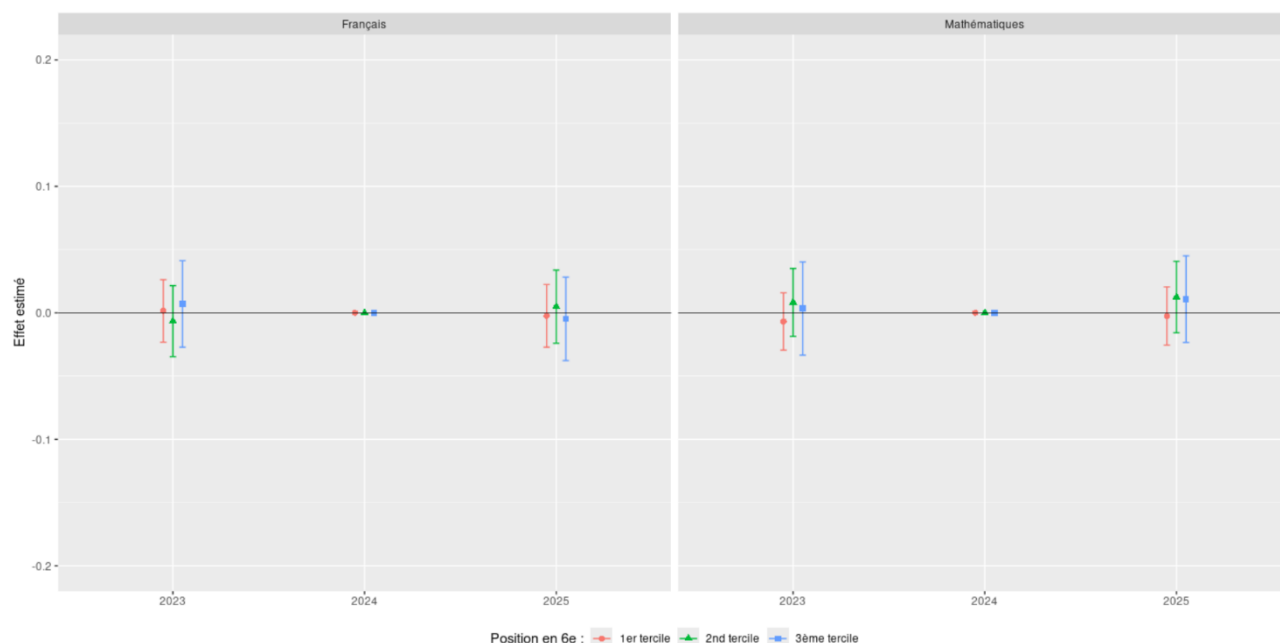
Lecture : En français, les coefficients de l'analyse événementielle pour les évaluations de septembre 2023 sont égaux à 0,03 pour les filles et à 0,04 pour les garçons. Aucun de ces coefficients n'est significatif au seuil de 5 % et le coefficient associé à chacun des sexes appartient à l'intervalle de confiance de l'autre. L'évolution de l'écart sur les scores de français entre les élèves qui effectuent leur 5^e dans un établissement qui met en place des groupes constitués selon le profil scolaire des élèves et ceux des établissements sans aucune forme de regroupement, entre 2023 et 2024, ne varie donc pas selon le sexe de l'élève.

Note : Pour chaque année, les scores sont standardisés par rapport aux scores des élèves scolarisés en 5^e dans les établissements qui ne mettent en place aucune forme de regroupement au cours de l'année scolaire 2024-2025. Les erreurs-types sont clusterisées au niveau de l'établissement fréquenté en 5^e. Les intervalles de confiance à 95 % sont affichés sur le graphique. L'année 2024 est l'année de référence.

Champ : Élèves scolarisés en 5^e au sein des établissements de l'enquête EEC ayant mis en place des regroupements selon le profil scolaire des élèves ou aucune forme de regroupement en 5^e pour la discipline. Cela correspond à 948 collèges pour le français et 958 collèges pour les mathématiques.

Sources : DEPP, Évaluations nationales en français et mathématiques de 4^e (2023, 2024, 2025) ; enquête EEC, questionnaires Chefs d'établissement, printemps 2025 ; Constat de rentrée 2025.

FIGURE 39 • Différence de scores aux évaluations nationales en français et mathématiques à l'entrée en 4^e avant et après la mise en place de la réforme (comparaison : groupes composés selon le profil scolaire des élèves/groupes hétérogènes), selon le niveau initial à l'entrée en 6^e



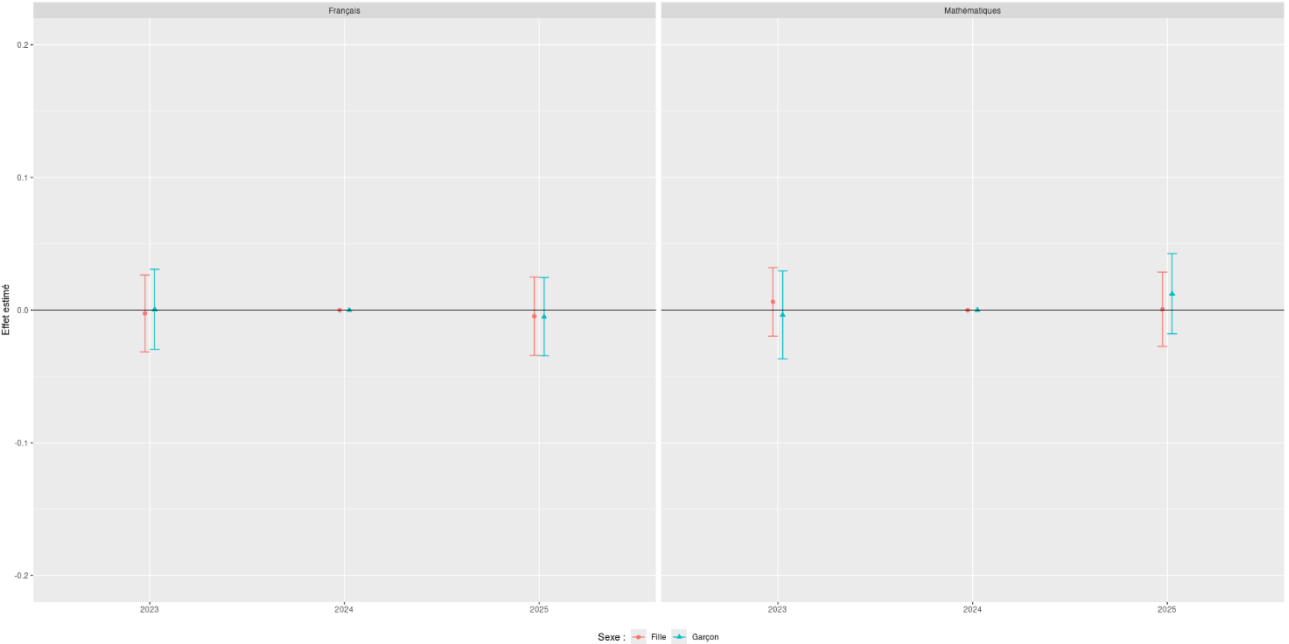
Lecture : En français, les coefficients de l'analyse événementielle pour les évaluations de septembre 2023 sont égaux à 0,002 pour les élèves qui appartenaient au premier tercile de leur établissement pour les scores de français aux évaluations de 6^e, à -0,007 pour les élèves du second tercile et à 0,007 pour les élèves du troisième tercile, c'est-à-dire ceux qui avaient les meilleurs résultats au sein de l'établissement où ils ont passé leur évaluation de début de 6^e. Aucun de ces coefficients n'est significatif au seuil de 5 % et le coefficient de chacun des terciles appartient aux intervalles de confiance des deux autres. L'évolution de l'écart sur les scores de français entre les élèves qui effectuent leur 5^e dans un établissement qui met en place des groupes constitués selon le profil scolaire des élèves et ceux des établissements avec des regroupements hétérogènes, entre 2023 et 2024, ne varie donc pas selon le niveau de l'élève en français en 6^e, regroupé en terciles.

Note : Pour chaque année, les scores sont standardisés par rapport aux scores des élèves scolarisés en 5^e dans les établissements qui ne mettent en place aucune forme de regroupement au cours de l'année scolaire 2024-2025. Les erreurs-types sont clusterisées au niveau de l'établissement fréquenté en 5^e. Les intervalles de confiance à 95 % sont affichés sur le graphique. L'année 2024 est l'année de référence.

Champ : Élèves scolarisés en 5^e au sein des établissements de l'enquête EEC ayant mis en place des regroupements selon le profil scolaire des élèves ou des groupes hétérogènes en 5^e pour la discipline. Cela correspond à 1 146 collèges pour le français et 1 151 collèges pour les mathématiques.

Sources : DEPP, Évaluations nationales en français et mathématiques de 4^e (2023, 2024, 2025) ; Évaluations nationales en français et mathématiques de 6^e (2021, 2022, 2023) ; enquête EEC, questionnaires Chefs d'établissement, printemps 2025 ; Constat de rentrée 2025.

FIGURE 40 • Différence de scores aux évaluations nationales en français et mathématiques à l'entrée en 4^e avant et après la mise en place de la réforme (comparaison : groupes composés selon le profil scolaire des élèves/groupes hétérogènes), selon le sexe



Lecture : En français, les coefficients de l'analyse événementielle pour les évaluations de septembre 2023 sont égaux à -0,003 pour les filles et à 0,001 pour les garçons. Aucun de ces coefficients n'est significatif au seuil de 5 % et le coefficient associé à chacun des sexes appartient à l'intervalle de confiance de l'autre. L'évolution de l'écart sur les scores de français entre les élèves qui effectuent leur 5^e dans un établissement qui met en place des groupes hétérogènes et ceux des établissements sans aucune forme de regroupement, entre 2023 et 2024, ne varie donc pas selon le sexe de l'élève.

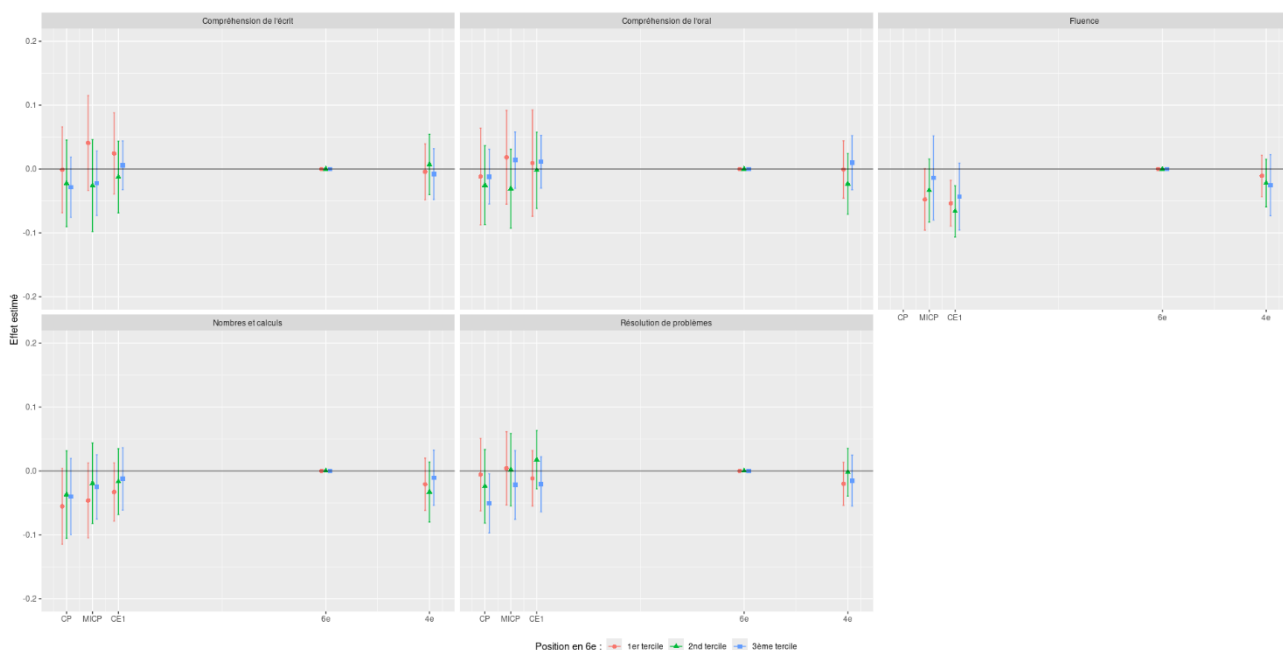
Note : Pour chaque année, les scores sont standardisés par rapport aux scores des élèves scolarisés en 5^e dans les établissements qui ne mettent en place aucune forme de regroupement au cours de l'année scolaire 2024-2025. Les erreurs-types sont clusterisées au niveau de l'établissement fréquenté en 5^e. Les intervalles de confiance à 95 % sont affichés sur le graphique. L'année 2024 est l'année de référence.

Champ : Élèves scolarisés en 5^e au sein des établissements de l'enquête EEC ayant mis en place des regroupements selon le profil scolaire des élèves ou des groupes hétérogènes en 5^e pour la discipline. Cela correspond à 1 146 collèges pour le français et 1 151 collèges pour les mathématiques.

Sources : DEPP, Évaluations nationales en français et mathématiques de 4^e (2023, 2024, 2025) ; enquête EEC, questionnaires Chefs d'établissement, printemps 2025 ; Constat de rentrée 2025.

Seconde approche

FIGURE 41 • Différence de scores par domaine aux évaluations nationales en français et mathématiques avant et après la mise en place de la réforme (comparaison : groupes hétérogènes/pas de regroupement), selon le niveau initial à l'entrée en 6^e



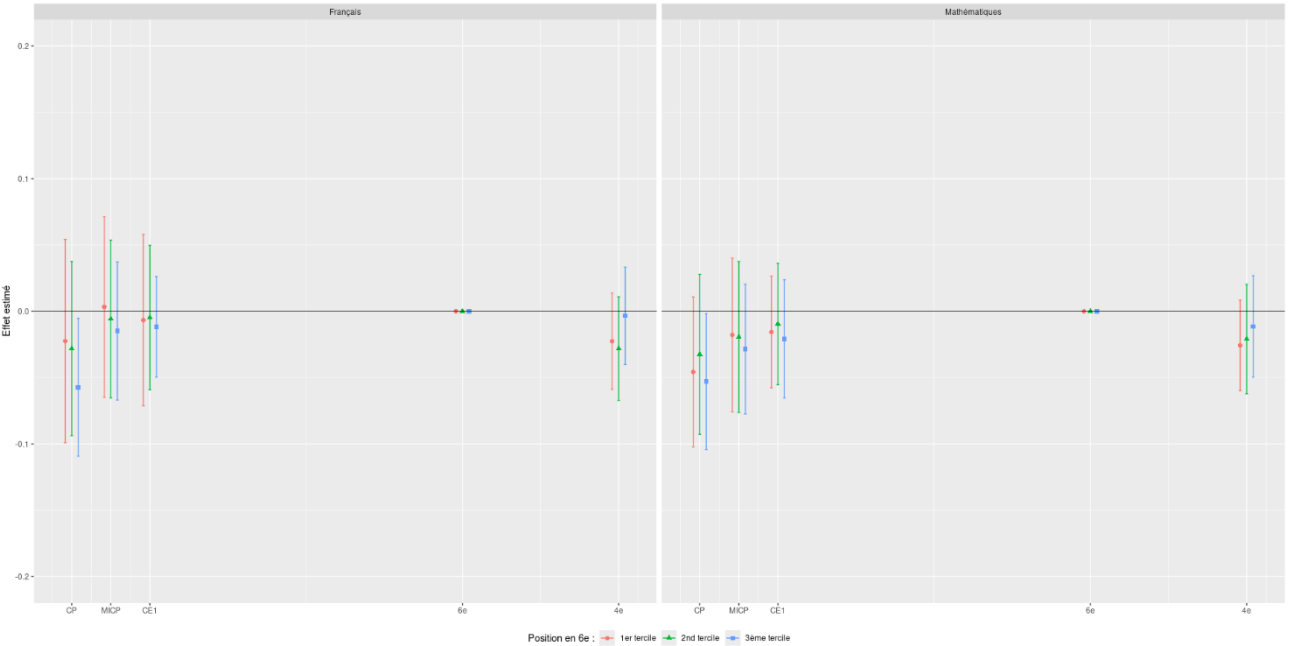
Lecture : Pour la compréhension de l'écrit, les coefficients de l'analyse événementielle pour les évaluations de début de CP sont égaux à -0,001 pour les élèves qui appartenaient au premier tertile de leur établissement pour les scores de compréhension de l'écrit aux évaluations de début de 6^e, à -0,02 pour les élèves du second tertile et à -0,03 pour les élèves du troisième tertile, c'est-à-dire ceux qui avaient les meilleurs résultats au sein de l'établissement où ils ont passé leur évaluation de début de 6^e. Aucun de ces coefficients n'est significatif au seuil de 5 % et le coefficient de chacun des tertiles appartient aux intervalles de confiance des deux autres. L'évolution de l'écart sur les scores de compréhension de l'écrit entre les élèves qui effectuent leur 5^e dans un établissement qui met en place des groupes hétérogènes et ceux des établissements sans aucune forme de regroupement, entre le CP et la 6^e, ne varie donc pas selon le niveau de l'élève en compréhension de l'écrit en 6^e, regroupé en tertiles.

Note : Pour chaque niveau, les scores sont standardisés par rapport aux scores des élèves scolarisés en 5^e dans les établissements qui ne mettent en place aucune forme de regroupement au cours de l'année scolaire 2024-2025. Les erreurs-types sont clusterisées au niveau de l'établissement fréquenté en 5^e. Les intervalles de confiance à 95 % sont affichés sur le graphique. La 6^e est le niveau de référence.

Champ : Élèves scolarisés en 5^e au sein des établissements de l'enquête EEC ayant mis en place des regroupements hétérogènes ou aucune forme de regroupement en 5^e pour la discipline. Cela correspond à 484 collèges pour le français et 473 collèges pour les mathématiques. Pour chaque niveau et pour chaque domaine, le score par domaine est calculé uniquement pour les élèves ayant un score dans chacune des compétences retenues pour ce domaine.

Sources : DEPP, Évaluations nationales en français et mathématiques de CP (2018), de mi-CP (2019), de CE1 (2019), de 6^e (2023) et de 4^e (2025) ; enquête EEC, questionnaires Chefs d'établissement, printemps 2025 ; Constat de rentrée 2025.

FIGURE 42 • Différence de scores aux évaluations nationales en français et mathématiques avant et après la mise en place de la réforme (comparaison : groupes hétérogènes/pas de regroupement), selon le niveau initial à l'entrée en 6^e



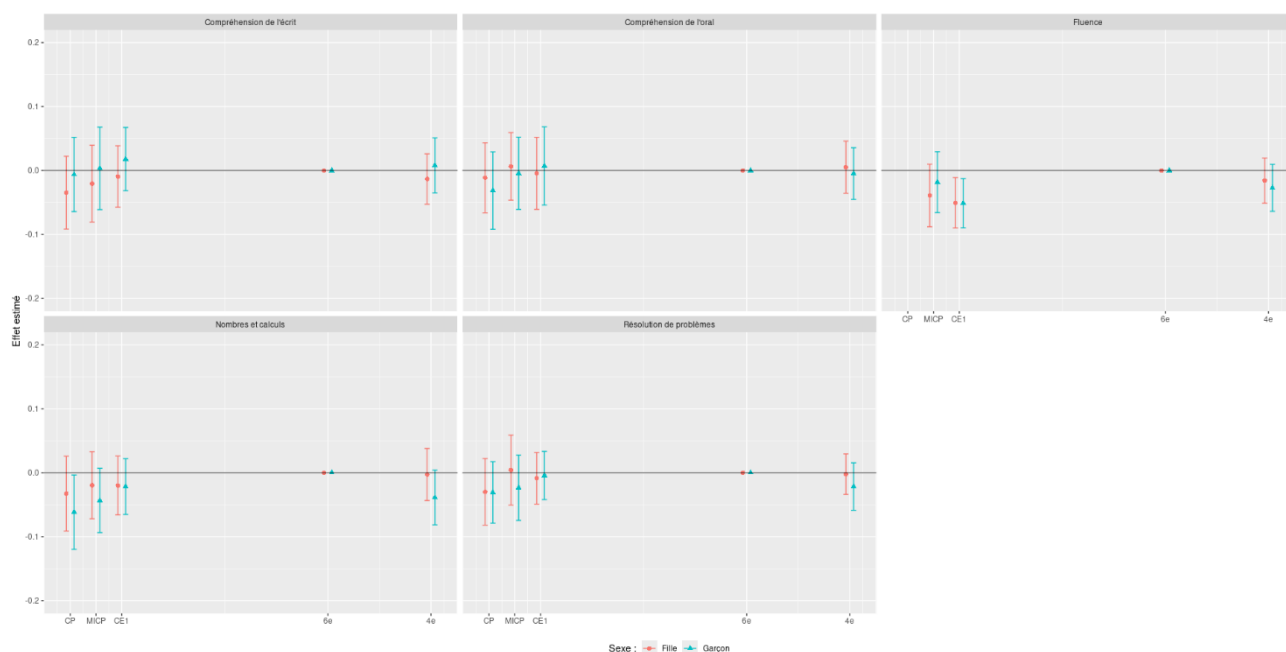
Lecture : Pour le français, les coefficients de l'analyse événementielle pour les évaluations de début de CP sont égaux à -0,02 pour les élèves qui appartenaient au premier tertile de leur établissement pour les scores de français aux évaluations de début de 6^e, à -0,03 pour les élèves du second tertile et à -0,06 pour les élèves du troisième tertile, c'est-à-dire ceux qui avaient les meilleurs résultats au sein de l'établissement où ils ont passé leur évaluation de début de 6^e. Seul ce dernier coefficient est significatif au seuil de 5 %, et il n'est pas significativement différent des coefficients estimés pour les deux autres tertiles.

Note : Pour chaque niveau, les scores sont standardisés par rapport aux scores des élèves scolarisés en 5^e dans les établissements qui ne mettent en place aucune forme de regroupement au cours de l'année scolaire 2024-2025. Les erreurs-types sont clusterisées au niveau de l'établissement fréquenté en 5^e. Les intervalles de confiance à 95 % sont affichés sur le graphique. La 6^e est le niveau de référence.

Champ : Élèves scolarisés en 5^e au sein des établissements de l'enquête EEC ayant mis en place des regroupements hétérogènes ou aucune forme de regroupement en 5^e pour la discipline. Cela correspond à 484 collèges pour le français et 473 collèges pour les mathématiques. Pour chaque niveau et pour chaque discipline, le score par discipline est calculé uniquement pour les élèves ayant un score dans chacun des domaines retenus pour cette discipline.

Sources : DEPP, Évaluations nationales en français et mathématiques de CP (2018), de mi-CP (2019), de CE1 (2019), de 6^e (2023) et de 4^e (2025) ; enquête EEC, questionnaires Chefs d'établissement, printemps 2025 ; Constat de rentrée 2025.

FIGURE 43 • Différence de scores par domaine aux évaluations nationales en français et mathématiques avant et après la mise en place de la réforme (comparaison : groupes hétérogènes/pas de regroupement), selon le sexe



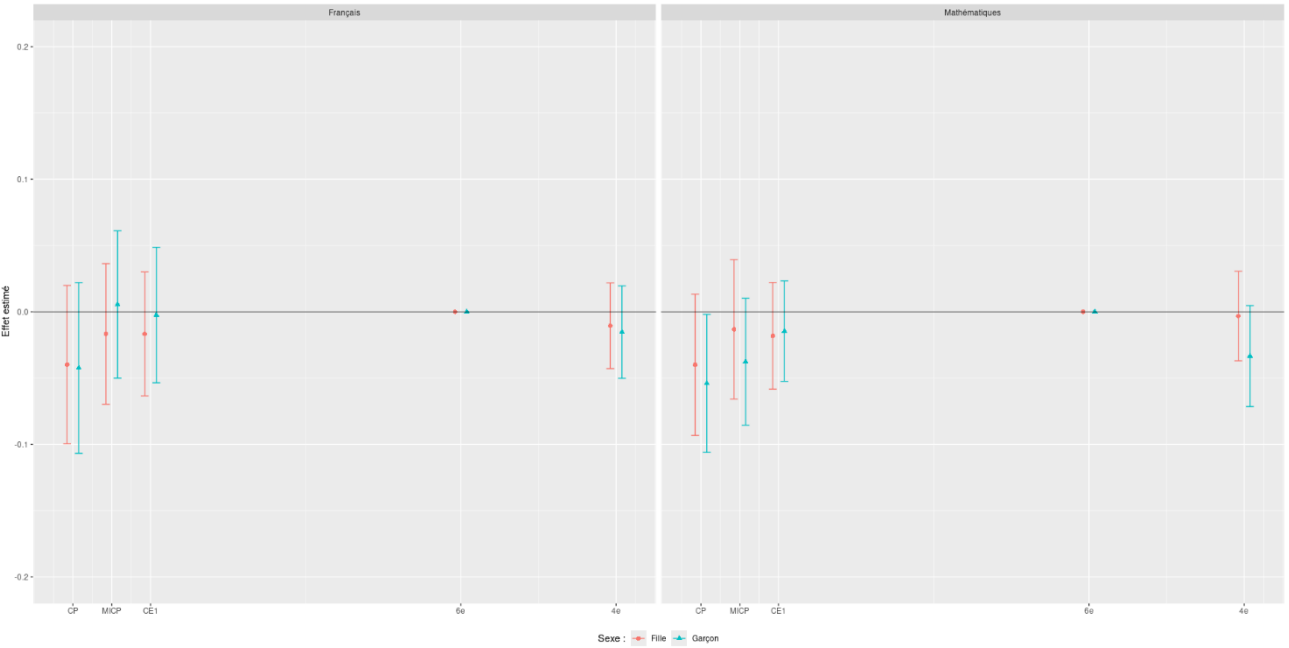
Lecture : Pour la compréhension de l'écrit, les coefficients de l'analyse événementielle pour les évaluations de début de CP sont égaux à -0,03 pour filles et à -0,006 pour les garçons. Aucun de ces deux coefficients n'est significatif au seuil de 5 % et le coefficient estimé pour chaque sexe appartient à l'intervalle de confiance de l'autre. L'évolution de l'écart sur les scores de compréhension de l'écrit entre les élèves qui effectuent leur 5^e dans un établissement qui met en place des groupes hétérogènes et ceux des établissements sans aucune forme de regroupement, entre le CP et la 6^e, ne varie donc pas selon le sexe de l'élève.

Note : Pour chaque niveau, les scores sont standardisés par rapport aux scores des élèves scolarisés en 5^e dans les établissements qui ne mettent en place aucune forme de regroupement au cours de l'année scolaire 2024-2025. Les erreurs-types sont clusterisées au niveau de l'établissement fréquenté en 5^e. Les intervalles de confiance à 95 % sont affichés sur le graphique. La 6^e est le niveau de référence.

Champ : Élèves scolarisés en 5^e au sein des établissements de l'enquête EEC ayant mis en place des regroupements hétérogènes ou aucune forme de regroupement en 5^e pour la discipline. Cela correspond à 484 collèges pour le français et 473 collèges pour les mathématiques. Pour chaque niveau et pour chaque domaine, le score par domaine est calculé uniquement pour les élèves ayant un score dans chacune des compétences retenues pour ce domaine.

Sources : DEPP, Évaluations nationales en français et mathématiques de CP (2018), de mi-CP (2019), de CE1 (2019), de 6e (2023) et de 4e (2025) ; enquête EEC, questionnaires Chefs d'établissement, printemps 2025 ; Constat de rentrée 2025.

FIGURE 44 • Différence de scores aux évaluations nationales en français et mathématiques avant et après la mise en place de la réforme (comparaison : groupes hétérogènes/pas de regroupement), selon le sexe



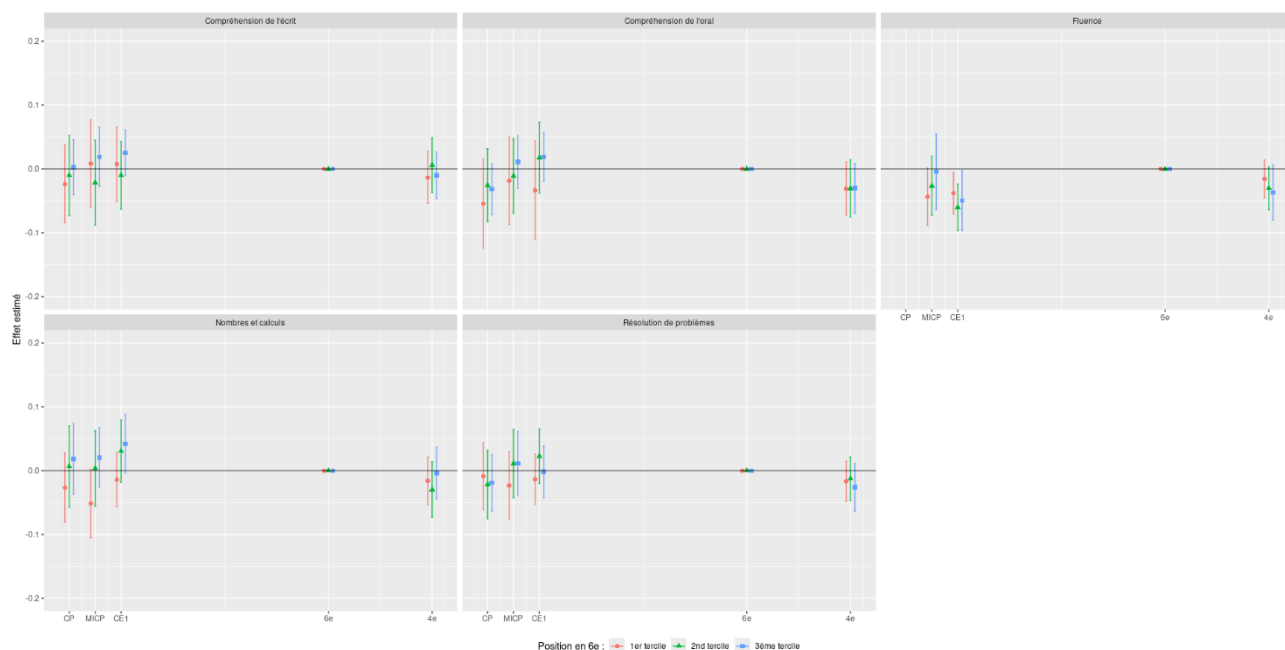
Lecture : Pour le français, les coefficients de l'analyse événementielle pour les évaluations de début de CP sont égaux à -0,04 pour filles et à -0,04 pour les garçons. Aucun de ces deux coefficients n'est significatif au seuil de 5 % et le coefficient estimé pour chaque sexe appartient à l'intervalle de confiance de l'autre. L'évolution de l'écart sur les scores de français entre les élèves qui effectuent leur 5^e dans un établissement qui met en place des groupes hétérogènes et ceux des établissements sans aucune forme de regroupement, entre le CP et la 6^e, ne varie donc pas selon le sexe de l'élève.

Note : Pour chaque niveau, les scores sont standardisés par rapport aux scores des élèves scolarisés en 5^e dans les établissements qui ne mettent en place aucune forme de regroupement au cours de l'année scolaire 2024-2025. Les erreurs-types sont clusterisées au niveau de l'établissement fréquenté en 5^e. Les intervalles de confiance au seuil de 5 % sont affichés sur le graphique. La 6^e est le niveau de référence.

Champ : Élèves scolarisés en 5^e au sein des établissements de l'enquête EEC ayant mis en place des regroupements hétérogènes ou aucune forme de regroupement en 5^e pour la discipline. Cela correspond à 484 collèges pour le français et 473 collèges pour les mathématiques. Pour chaque niveau et pour chaque discipline, le score par discipline est calculé uniquement pour les élèves ayant un score dans chacun des domaines retenus pour cette discipline.

Sources : DEPP, Évaluations nationales en français et mathématiques de CP (2018), de mi-CP (2019), de CE1 (2019), de 6^e (2023) et de 4^e (2025) ; enquête EEC, questionnaires Chefs d'établissement, printemps 2025 ; Constat de rentrée 2025.

FIGURE 45 • Différence de scores par domaine aux évaluations nationales en français et mathématiques avant et après la mise en place de la réforme (comparaison : groupes composés selon le profil scolaire des élèves/pas de regroupement), selon le niveau initial à l'entrée en 6^e



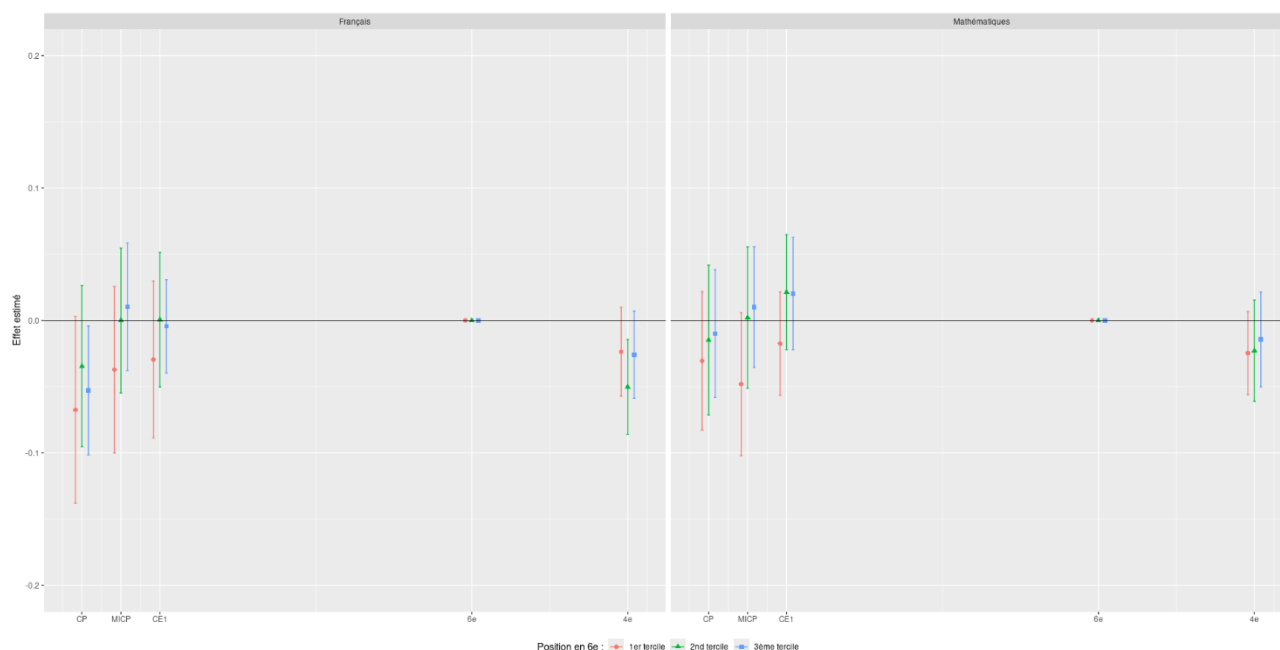
Lecture : Pour la compréhension de l'écrit, les coefficients de l'analyse événementielle pour les évaluations de début de CP sont égaux à -0,02 pour les élèves qui appartenaient au premier tercile de leur établissement pour les scores de compréhension de l'écrit aux évaluations de début de 6^e, à -0,01 pour les élèves du second tercile et à 0,003 pour les élèves du troisième tercile, c'est-à-dire ceux qui avaient les meilleurs résultats au sein de l'établissement où ils ont passé leur évaluation de début de 6^e. Aucun de ces coefficients n'est significatif au seuil de 5 % et le coefficient de chacun des terciles appartient aux intervalles de confiance des deux autres. L'évolution de l'écart sur les scores de compréhension de l'écrit entre les élèves qui effectuent leur 5^e dans un établissement qui met en place des groupes constitués selon le profil scolaire des élèves et ceux des établissements sans aucune forme de regroupement, entre le CP et la 6^e, ne varie donc pas selon le niveau de l'élève en compréhension de l'écrit en 6^e, regroupé en terciles.

Note : Pour chaque niveau, les scores sont standardisés par rapport aux scores des élèves scolarisés en 5^e dans les établissements qui ne mettent en place aucune forme de regroupement au cours de l'année scolaire 2024-2025. Les erreurs-types sont clusterisées au niveau de l'établissement fréquenté en 5^e. Les intervalles de confiance à 95 % sont affichés sur le graphique. La 6^e est le niveau de référence.

Champ : Élèves scolarisés en 5^e au sein des établissements de l'enquête EEC ayant mis en place des regroupements hétérogènes ou aucune forme de regroupement en 5^e pour la discipline. Cela correspond à 484 collèges pour le français et 473 collèges pour les mathématiques. Pour chaque niveau et pour chaque domaine, le score par domaine est calculé uniquement pour les élèves ayant un score dans chacune des compétences retenues pour ce domaine.

Sources : DEPP, Évaluations nationales en français et mathématiques de CP (2018), de mi-CP (2019), de CE1 (2019), de 6^e (2023) et de 4^e (2025) ; enquête EEC, questionnaires Chefs d'établissement, printemps 2025 ; Constat de rentrée 2025.

FIGURE 46 • Différence de scores aux évaluations nationales en français et mathématiques avant et après la mise en place de la réforme (comparaison : groupes composés selon le profil scolaire des élèves/pas de regroupement), selon le niveau initial à l'entrée en 6^e



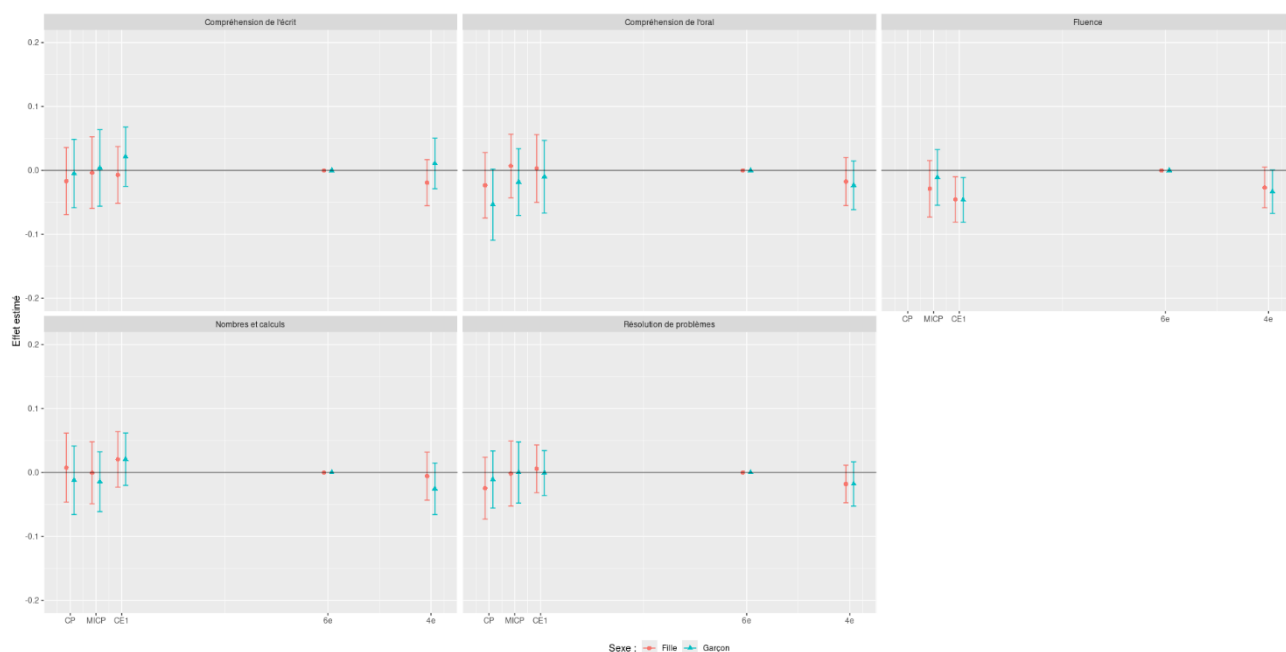
Lecture : Pour le français, les coefficients de l'analyse événementielle pour les évaluations de début de CP sont égaux à -0,02 pour les élèves qui appartenaient au premier tercile de leur établissement pour les scores de français aux évaluations de début de 6^e, à -0,03 pour les élèves du second tercile et à -0,06 pour les élèves du troisième tercile, c'est-à-dire ceux qui avaient les meilleurs résultats au sein de l'établissement où ils ont passé leur évaluation de début de 6^e. Seul ce dernier coefficient est significatif au seuil de 5 %, et il n'est pas significativement différent des coefficients estimés pour les deux autres terciles.

Note : Pour chaque niveau, les scores sont standardisés par rapport aux scores des élèves scolarisés en 5^e dans les établissements qui ne mettent en place aucune forme de regroupement au cours de l'année scolaire 2024-2025. Les erreurs-types sont clusterisées au niveau de l'établissement fréquenté en 5^e. Les intervalles de confiance à 95 % sont affichés sur le graphique. La 6^e est le niveau de référence.

Champ : Élèves scolarisés en 5^e au sein des établissements de l'enquête EEC ayant mis en place des regroupements hétérogènes ou aucune forme de regroupement en 5^e pour la discipline. Cela correspond à 484 collèges pour le français et 473 collèges pour les mathématiques. Pour chaque niveau et pour chaque discipline, le score par discipline est calculé uniquement pour les élèves ayant un score dans chacun des domaines retenus pour cette discipline.

Sources : DEPP, Évaluations nationales en français et mathématiques de CP (2018), de mi-CP (2019), de CE1 (2019), de 6^e (2023) et de 4^e (2025) ; enquête EEC, questionnaires Chefs d'établissement, printemps 2025 ; Constat de rentrée 2025.

FIGURE 47 • Différence de scores par domaine aux évaluations nationales en français et mathématiques avant et après la mise en place de la réforme (comparaison : groupes composés selon le profil scolaire des élèves/pas de regroupement), selon le sexe



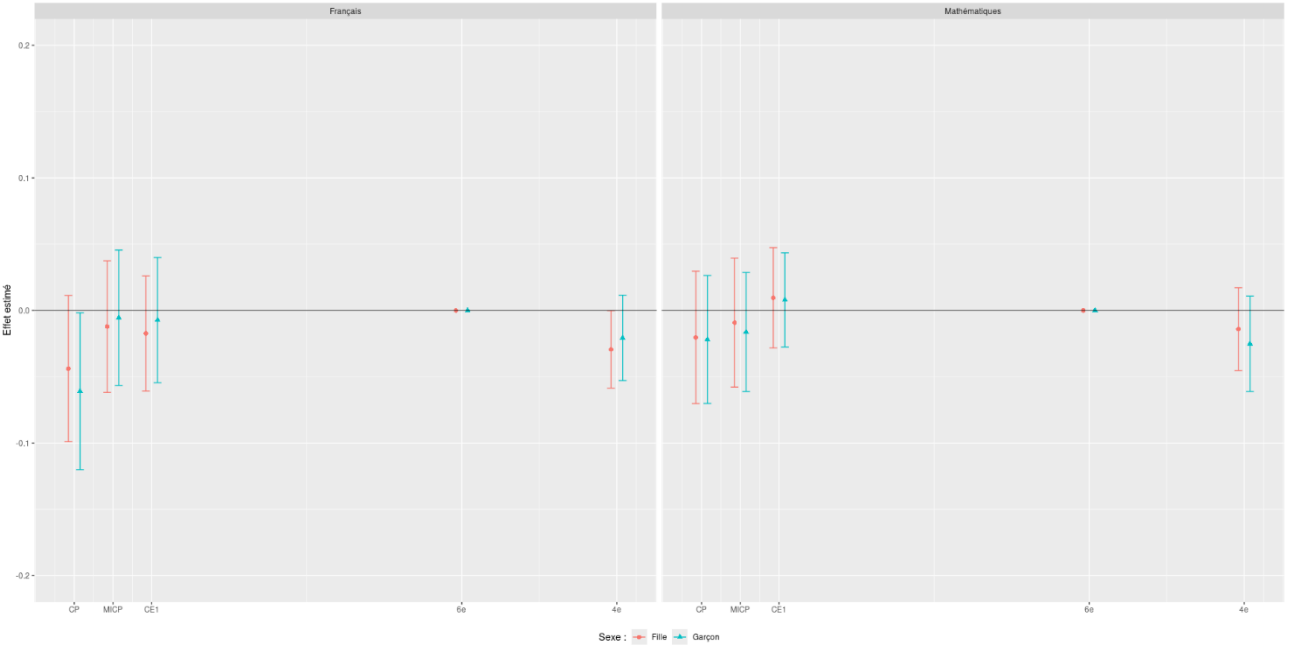
Lecture : Pour la compréhension de l'écrit, les coefficients de l'analyse événementielle pour les évaluations de début de CP sont égaux à -0,02 pour filles et à -0,005 pour les garçons. Aucun de ces deux coefficients n'est significatif au seuil de 5 % et le coefficient estimé pour chaque sexe appartient à l'intervalle de confiance de l'autre. L'évolution de l'écart sur les scores de compréhension de l'écrit entre les élèves qui effectuent leur 5^e dans un établissement qui met en place des groupes constitués selon le profil scolaire des élèves et ceux des établissements sans aucune forme de regroupement, entre le CP et la 6^e, ne varie donc pas selon le sexe de l'élève.

Note : Pour chaque niveau, les scores sont standardisés par rapport aux scores des élèves scolarisés en 5^e dans les établissements qui ne mettent en place aucune forme de regroupement au cours de l'année scolaire 2024-2025. Les erreurs-types sont clusterisées au niveau de l'établissement fréquenté en 5^e. Les intervalles de confiance à 95 % sont affichés sur le graphique. La 6^e est le niveau de référence.

Champ : Élèves scolarisés en 5^e au sein des établissements de l'enquête EEC ayant mis en place des regroupements hétérogènes ou aucune forme de regroupement en 5^e pour la discipline. Cela correspond à 484 collèges pour le français et 473 collèges pour les mathématiques. Pour chaque niveau et pour chaque domaine, le score par domaine est calculé uniquement pour les élèves ayant un score dans chacune des compétences retenues pour ce domaine.

Sources : DEPP, Évaluations nationales en français et mathématiques de CP (2018), de mi-CP (2019), de CE1 (2019), de 6^e (2023) et de 4^e (2025) ; enquête EEC, questionnaires Chefs d'établissement, printemps 2025 ; Constat de rentrée 2025.

FIGURE 48 • Différence de scores aux évaluations nationales en français et mathématiques avant et après la mise en place de la réforme (comparaison : groupes composés selon le profil scolaire des élèves/pas de regroupement), selon le sexe



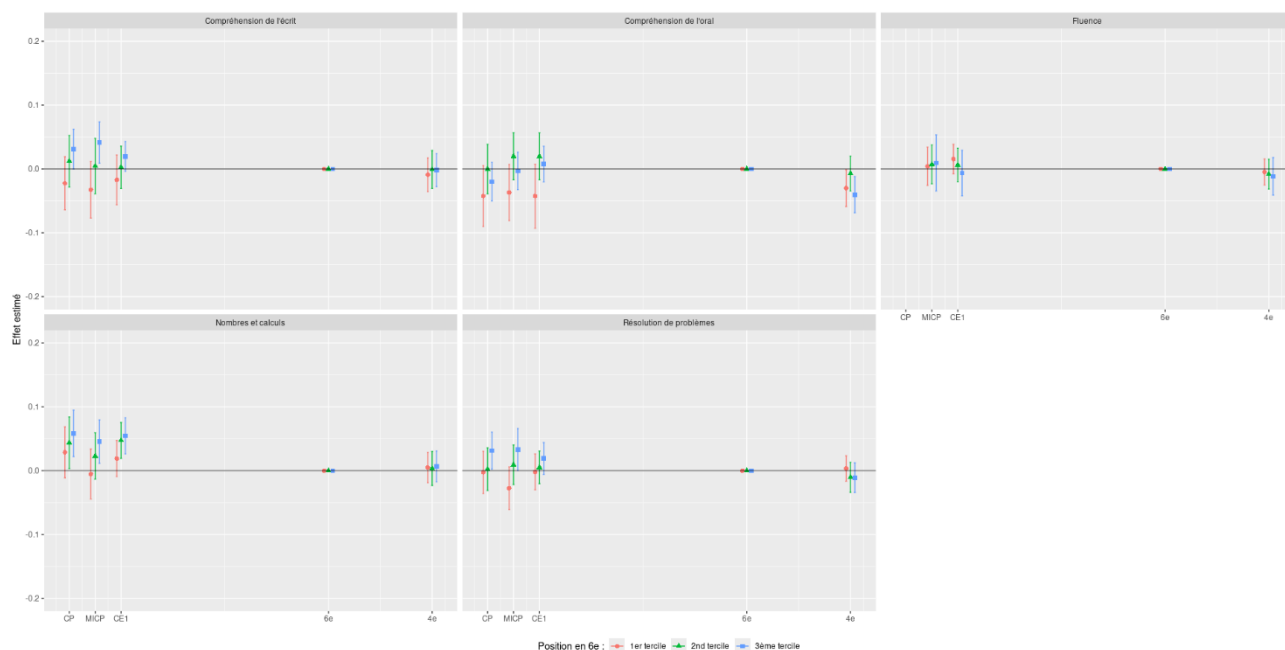
Lecture : Pour le français, les coefficients de l’analyse événementielle pour les évaluations de début de CP sont égaux à -0,04 pour filles et à -0,06 pour les garçons. Seul ce dernier coefficient est significatif au seuil de 5 %, et il n’est pas significativement différent du coefficient estimé pour les filles.

Note : Pour chaque niveau, les scores sont standardisés par rapport aux scores des élèves scolarisés en 5^e dans les établissements qui ne mettent en place aucune forme de regroupement au cours de l’année scolaire 2024-2025. Les erreurs-types sont clusterisées au niveau de l’établissement fréquenté en 5^e. Les intervalles de confiance à 95 % sont affichés sur le graphique. La 6^e est le niveau de référence.

Champ : Élèves scolarisés en 5^e au sein des établissements de l’enquête EEC ayant mis en place des regroupements hétérogènes ou aucune forme de regroupement en 5^e pour la discipline. Cela correspond à 484 collèges pour le français et 473 collèges pour les mathématiques. Pour chaque niveau et pour chaque discipline, le score par discipline est calculé uniquement pour les élèves ayant un score dans chacun des domaines retenus pour cette discipline.

Sources : DEPP, Évaluations nationales en français et mathématiques de CP (2018), de mi-CP (2019), de CE1 (2019), de 6^e (2023) et de 4^e (2025) ; enquête EEC, questionnaires Chefs d’établissement, printemps 2025 ; Constat de rentrée 2025.

FIGURE 49 • Différence de scores par domaine aux évaluations nationales en français et mathématiques avant et après la mise en place de la réforme (comparaison : groupes composés selon le profil scolaire des élèves/groupes hétérogènes), selon le niveau initial à l'entrée en 6^e



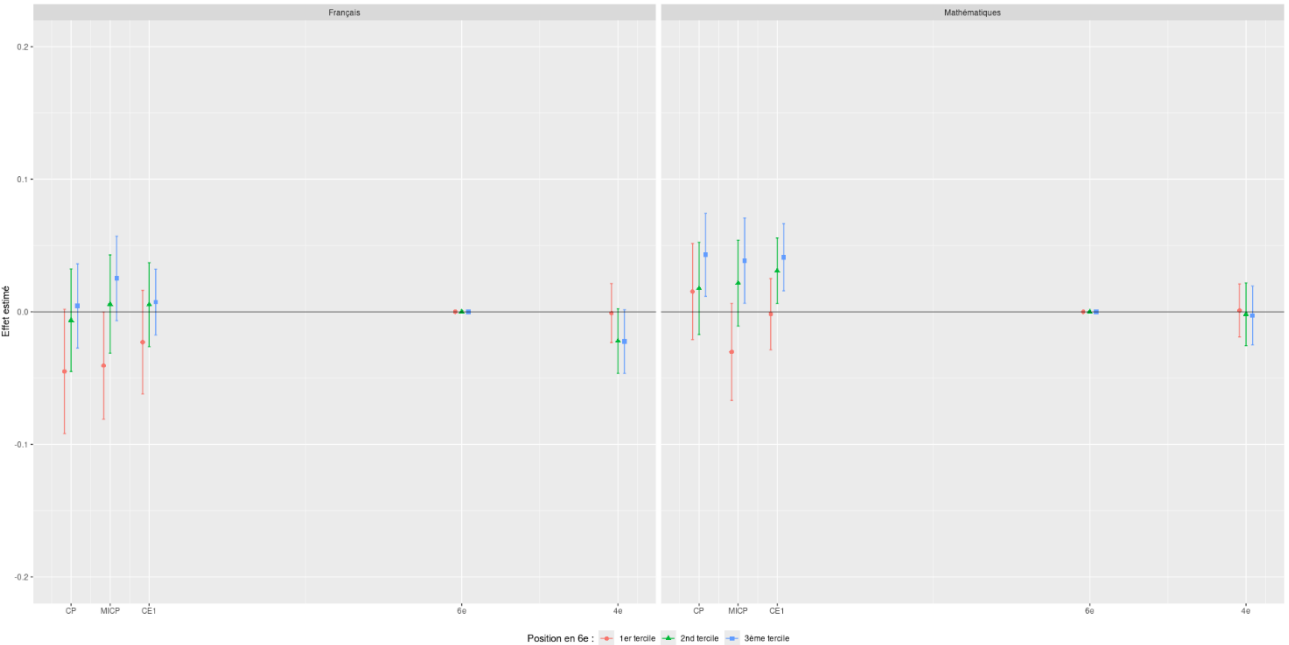
Lecture : Pour la compréhension de l'écrit, les coefficients de l'analyse événementielle pour les évaluations de début de CP sont égaux à -0,02 pour les élèves qui appartenaient au premier tercile de leur établissement pour les scores de compréhension de l'écrit aux évaluations de début de 6^e, à 0,01 pour les élèves du second tercile et à 0,03 pour les élèves du troisième tercile, c'est-à-dire ceux qui avaient les meilleurs résultats au sein de l'établissement où ils ont passé leur évaluation de début de 6^e. Ce dernier coefficient est significatif au seuil de 5 % et est significativement différent de celui associé au premier tercile. L'évolution de l'écart sur les scores de compréhension de l'écrit entre les élèves qui effectuent leur 5^e dans un établissement qui met en place des groupes constitués selon le profil scolaire des élèves et ceux des établissements avec des regroupements hétérogènes, entre le CP et la 6^e, diffère donc selon le niveau de l'élève en compréhension de l'écrit en 6^e, regroupé en terciles.

Note : Pour chaque niveau, les scores sont standardisés par rapport aux scores des élèves scolarisés en 5^e dans les établissements qui ne mettent en place aucune forme de regroupement au cours de l'année scolaire 2024-2025. Les erreurs-types sont clusterisées au niveau de l'établissement fréquenté en 5^e. Les intervalles de confiance à 95 % sont affichés sur le graphique. La 6^e est le niveau de référence.

Champ : Élèves scolarisés en 5^e au sein des établissements de l'enquête EEC ayant mis en place des regroupements hétérogènes ou aucune forme de regroupement en 5^e pour la discipline. Cela correspond à 484 collèges pour le français et 473 collèges pour les mathématiques. Pour chaque niveau et pour chaque domaine, le score par domaine est calculé uniquement pour les élèves ayant un score dans chacune des compétences retenues pour ce domaine.

Sources : DEPP, Évaluations nationales en français et mathématiques de CP (2018), de mi-CP (2019), de CE1 (2019), de 6^e (2023) et de 4^e (2025) ; enquête EEC, questionnaires Chefs d'établissement, printemps 2025 ; Constat de rentrée 2025.

FIGURE 50 • Différence de scores aux évaluations nationales en français et mathématiques avant et après la mise en place de la réforme (comparaison : groupes composés selon le profil scolaire des élèves/groupes hétérogènes), selon le niveau initial à l'entrée en 6^e



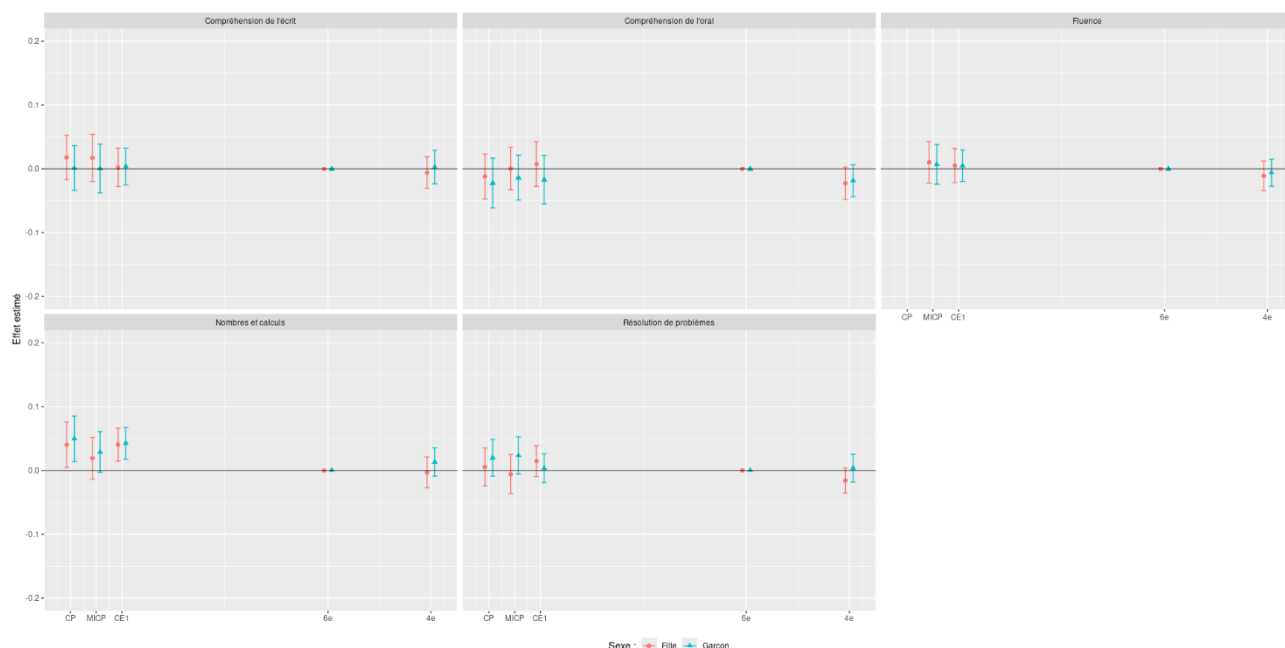
Lecture : Pour le français, les coefficients de l'analyse événementielle pour les évaluations de début de CP sont égaux à -0,04 pour les élèves qui appartenaient au premier tertile de leur établissement pour les scores de français aux évaluations de début de 6^e, à -0,006 pour les élèves du second tertile et à 0,004 pour les élèves du troisième tertile, c'est-à-dire ceux qui avaient les meilleurs résultats au sein de l'établissement où ils ont passé leur évaluation de début de 6^e. Aucun de ces coefficients n'est significatif au seuil de 5 % et le coefficient de chacun des tertiles appartient aux intervalles de confiance des deux autres. L'évolution de l'écart sur les scores de français entre les élèves qui effectuent leur 5^e dans un établissement qui met en place des groupes constitués selon le profil scolaire des élèves et ceux des établissements avec des regroupements hétérogènes, entre le CP et la 6^e, ne varie donc pas selon le niveau de l'élève en français en 6^e, regroupé en tertiles.

Note : Pour chaque niveau, les scores sont standardisés par rapport aux scores des élèves scolarisés en 5^e dans les établissements qui ne mettent en place aucune forme de regroupement au cours de l'année scolaire 2024-2025. Les erreurs-types sont clusterisées au niveau de l'établissement fréquenté en 5^e. Les intervalles de confiance à 95 % sont affichés sur le graphique. La 6^e est le niveau de référence.

Champ : Élèves scolarisés en 5^e au sein des établissements de l'enquête EEC ayant mis en place des regroupements hétérogènes ou aucune forme de regroupement en 5^e pour la discipline. Cela correspond à 484 collèges pour le français et 473 collèges pour les mathématiques. Pour chaque niveau et pour chaque discipline, le score par discipline est calculé uniquement pour les élèves ayant un score dans chacun des domaines retenus pour cette discipline.

Sources : DEPP, Évaluations nationales en français et mathématiques de CP (2018), de mi-CP (2019), de CE1 (2019), de 6^e (2023) et de 4^e (2025) ; enquête EEC, questionnaires Chefs d'établissement, printemps 2025 ; Constat de rentrée 2025.

FIGURE 51 • Différence de scores par domaine aux évaluations nationales en français et mathématiques avant et après la mise en place de la réforme (comparaison : groupes composés selon le profil scolaire des élèves/groupes hétérogènes), selon le sexe



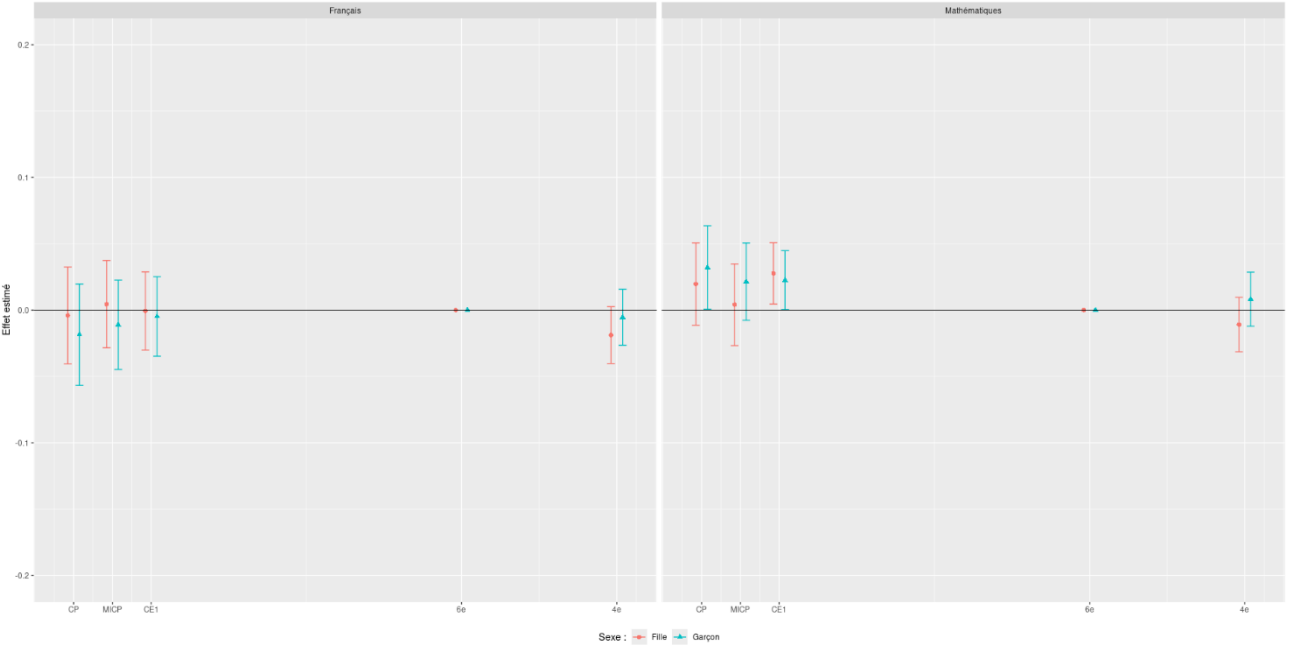
Lecture : Pour la compréhension de l'écrit, les coefficients de l'analyse événementielle pour les évaluations de début de CP sont égaux à 0,02 pour filles et à 0,001 pour les garçons. Aucun de ces deux coefficients n'est significatif au seuil de 5 % et le coefficient estimé pour chaque sexe appartient à l'intervalle de confiance de l'autre. L'évolution de l'écart sur les scores de compréhension de l'écrit entre les élèves qui effectuent leur 5^e dans un établissement qui met en place des groupes constitués selon le profil scolaire des élèves et ceux des établissements avec des regroupements hétérogènes, entre le CP et la 6^e, ne varie donc pas selon le sexe de l'élève.

Note : Pour chaque niveau, les scores sont standardisés par rapport aux scores des élèves scolarisés en 5^e dans les établissements qui ne mettent en place aucune forme de regroupement au cours de l'année scolaire 2024-2025. Les erreurs-types sont clusterisées au niveau de l'établissement fréquenté en 5^e. Les intervalles de confiance à 95 % sont affichés sur le graphique. La 6^e est le niveau de référence.

Champ : Élèves scolarisés en 5^e au sein des établissements de l'enquête EEC ayant mis en place des regroupements hétérogènes ou aucune forme de regroupement en 5^e pour la discipline. Cela correspond à 484 collèges pour le français et 473 collèges pour les mathématiques. Pour chaque niveau et pour chaque domaine, le score par domaine est calculé uniquement pour les élèves ayant un score dans chacune des compétences retenues pour ce domaine.

Sources : DEPP, Évaluations nationales en français et mathématiques de CP (2018), de mi-CP (2019), de CE1 (2019), de 6^e (2023) et de 4^e (2025) ; enquête EEC, questionnaires Chefs d'établissement, printemps 2025 ; Constat de rentrée 2025.

FIGURE 52 • Différence de scores aux évaluations nationales en français et mathématiques avant et après la mise en place de la réforme (comparaison : groupes composés selon le profil scolaire des élèves/groupes hétérogènes), selon le sexe



Lecture : Pour le français, les coefficients de l’analyse événementielle pour les évaluations de début de CP sont égaux à -0,004 pour filles et à -0,02 pour les garçons. Aucun de ces deux coefficients n’est significatif au seuil de 5 % et le coefficient estimé pour chaque sexe appartient à l’intervalle de confiance de l’autre. L’évolution de l’écart sur les scores de français entre les élèves qui effectuent leur 5^e dans un établissement qui met en place des groupes constitués selon le profil scolaire des élèves et ceux des établissements avec des regroupements hétérogènes, entre le CP et la 6^e, ne varie donc pas selon le sexe de l’élève.

Note : Pour chaque niveau, les scores sont standardisés par rapport aux scores des élèves scolarisés en 5^e dans les établissements qui ne mettent en place aucune forme de regroupement au cours de l’année scolaire 2024-2025. Les erreurs-types sont clusterisées au niveau de l’établissement fréquenté en 5^e. Les intervalles de confiance à 95 % sont affichés sur le graphique. La 6^e est le niveau de référence.

Champ : Élèves scolarisés en 5^e au sein des établissements de l’enquête EEC ayant mis en place des regroupements hétérogènes ou aucune forme de regroupement en 5^e pour la discipline. Cela correspond à 484 collèges pour le français et 473 collèges pour les mathématiques. Pour chaque niveau et pour chaque discipline, le score par discipline est calculé uniquement pour les élèves ayant un score dans chacun des domaines retenus pour cette discipline.

Source : DEPP, Évaluations nationales en français et mathématiques de CP (2018), de mi-CP (2019), de CE1 (2019), de 6^e (2023) et de 4^e (2025) ; enquête EEC, questionnaires Chefs d’établissement, printemps 2025 ; Constat de rentrée 2025.

Annexe 8 : Résultats par domaine évalué en français et mathématiques

TABLEAU 14 • Coefficients de l’estimation des effets des organisations pédagogiques sur les scores par domaine aux évaluations nationales en français de 4^e

Domaine	Compréhension de l’écrit					
Comparaison	Hétérogène vs pas de regroupement	Hétérogène vs pas de regroupement	Selon le profil scolaire des élèves vs pas de regroupement	Selon le profil scolaire des élèves vs pas de regroupement	Selon le profil scolaire des élèves vs hétérogène	Selon le profil scolaire des élèves vs hétérogène
Modèle	DiD	TWFE	DiD	TWFE	DiD	TWFE
Coefficient	0,002 (0,02)	0,002 (0,021)	0,002 (0,02)	-0,004 (0,019)	-0,005 (0,012)	-0,006 (0,013)
Crédibilité des tendances parallèles	Oui	Oui	Oui	Oui	Oui	Oui
Effets fixes	Non	Id élève + niveau	Non	Id élève + niveau	Non	Id élève + niveau
Cluster	UAI en 5 ^e	UAI en 5 ^e	UAI en 5 ^e	UAI en 5 ^e	UAI en 5 ^e	UAI en 5 ^e
N obs.	107 419	236 044	219 928	487 426	269 249	596 028
N élèves	56 182	56 482	115 325	115 953	141 041	141 817
N établissements	484	484	948	948	1 146	1 146

Lecture : Pour la compréhension de l’écrit, le modèle de doubles différences sans effet fixe estime un coefficient de 0,002 pour l’effet des regroupements hétérogènes : le fait d’appartenir à la cohorte d’élèves qui entre en 5^e en 2024-2025 et d’avoir été scolarisé dans un établissement qui met en place des groupes hétérogènes plutôt que dans un établissement qui ne met en place aucune forme de regroupement augmente en moyenne le score en compréhension de l’écrit de 0,002, soit 0,2 % d’un écart-type. Ce coefficient n’est pas significatif.

Note : Les modèles DiD correspondent aux modèles de doubles différences classiques (Équation 5) tandis que les modèles TWFE désignent les modèles avec effets fixes (Équation 6). Les erreurs types clusterisées au niveau de l’établissement fréquenté en 5^e sont présentées entre parenthèses. Seuils de significativité : « *** », pour un coefficient significatif à 0,1 %, « ** », pour un coefficient significatif à 1 %, « * », pour un coefficient significatif à 5 % et « . » pour un coefficient significatif à 10 %.

Champ : Élèves scolarisés en 5^e au sein des établissements de l’enquête EEC ayant adopté une forme unique de regroupement ou aucune en 5^e pour le français. Cela correspond à 1 289 collèges. Pour chaque niveau, le score en compréhension de l’écrit est calculé uniquement pour les élèves ayant un score dans chacune des compétences retenues pour ce domaine.

Sources : DEPP, Évaluations nationales en français de CP (2018), de mi-CP (2019), de CE1 (2019), de 6^e (2023) et de 4^e (2025) ; enquête EEC, questionnaires Chefs d’établissement, printemps 2025 ; Constat de rentrée 2025.

Domaine	Compréhension de l'oral					
Comparaison	Hétérogène vs pas de regroupement	Hétérogène vs pas de regroupement	Selon le profil scolaire des élèves vs pas de regroupement	Selon le profil scolaire des élèves vs pas de regroupement	Selon le profil scolaire des élèves vs hétérogène	Selon le profil scolaire des élèves vs hétérogène
Modèle	DiD	TWFE	DiD	TWFE	DiD	TWFE
Coefficient	0,007 (0,017)	0,004 (0,022)	0,007 (0,017)	-0,01 (0,021)	-0,018 (0,011)	-0,014 (0,013)
Crédibilité des tendances parallèles	Oui	Oui	Oui	Oui	Oui	Oui
Effets fixes	Non	Id élève + niveau	Non	Id élève + niveau	Non	Id élève + niveau
Cluster	UAI en 5 ^e	UAI en 5 ^e	UAI en 5 ^e	UAI en 5 ^e	UAI en 5 ^e	UAI en 5 ^e
N obs.	95 739	223 776	195 263	461 643	238 730	564 095
N élèves	54 046	56 090	110 743	115 094	135 351	140 780
N établissements	484	484	948	948	1 146	1 146

Lecture : Pour la compréhension de l'oral, le modèle de doubles différences sans effet fixe estime un coefficient de 0,007 pour l'effet des regroupements hétérogènes : le fait d'appartenir à la cohorte d'élèves qui entre en 5^e en 2024-2025 et d'avoir été scolarisé dans un établissement qui met en place des groupes hétérogènes plutôt que dans un établissement qui ne met en place aucune forme de regroupement augmente en moyenne le score en compréhension de l'oral de 0,007, soit 0,7 % d'un écart-type. Ce coefficient n'est pas significatif.

Note : Les modèles DiD correspondent aux modèles de doubles différences classiques (Équation 5) tandis que les modèles TWFE désignent les modèles avec effets fixes (Équation 6). Les erreurs types clusterisées au niveau de l'établissement fréquenté en 5^e sont présentées entre parenthèses. Seuils de significativité : « *** », pour un coefficient significatif à 0,1 %, « ** », pour un coefficient significatif à 1 %, « * », pour un coefficient significatif à 5 % et « . » pour un coefficient significatif à 10 %.

Champ : Élèves scolarisés en 5^e au sein des établissements de l'enquête EEC ayant adopté une forme unique de regroupement ou aucune en 5^e pour le français. Cela correspond à 1 289 collèges. Pour chaque niveau, le score en compréhension de l'oral est calculé uniquement pour les élèves ayant un score dans chacune des compétences retenues pour ce domaine.

Sources : DEPP, Évaluations nationales en français de CP (2018), de mi-CP (2019), de CE1 (2019), de 6^e (2023) et de 4^e (2025) ; enquête EEC, questionnaires Chefs d'établissement, printemps 2025 ; Constat de rentrée 2025.

Domaine	Fluence					
Comparaison	Hétérogène vs pas de regroupement	Hétérogène vs pas de regroupement	Selon le profil scolaire des élèves vs pas de regroupement	Selon le profil scolaire des élèves vs pas de regroupement	Selon le profil scolaire des élèves vs hétérogène	Selon le profil scolaire des élèves vs hétérogène
Modèle	DiD	TWFE	DiD	TWFE	DiD	TWFE
Coefficient	-0,024 (0,019)	0,004 (0,017)	-0,024 (0,019)	-0,009 (0,016)	-0,007 (0,013)	-0,013 (0,011)
Crédibilité des tendances parallèles	Non, écart significatif en CE1	Non, écart significatif en CE1	Non, écart significatif en CE1	Non, écart significatif en CE1	Oui	Oui
Effets fixes	Non	Id élève + niveau	Non	Id élève + niveau	Non	Id élève + niveau
Cluster	UAI en 5 ^e	UAI en 5 ^e	UAI en 5 ^e	UAI en 5 ^e	UAI en 5 ^e	UAI en 5 ^e
N obs.	102 821	190 900	210 676	392 816	258 177	481 134
N élèves	55 523	56 241	114 098	115 531	139 471	141 296
N établissements	484	484	948	948	1 146	1 146

Lecture : Pour la fluence, le modèle de doubles différences sans effet fixe estime un coefficient de -0,024 pour l'effet des regroupements hétérogènes : le fait d'appartenir à la cohorte d'élèves qui entre en 5^e en 2024-2025 et d'avoir été scolarisé dans un établissement qui met en place des groupes hétérogènes plutôt que dans un établissement qui ne met en place aucune forme de regroupement diminue en moyenne le score en fluence de 0,024, soit 2,4 % d'un écart-type. Ce coefficient n'est cependant pas significatif.

Note : Les modèles DiD correspondent aux modèles de doubles différences classiques (Équation 5) tandis que les modèles TWFE désignent les modèles avec effets fixes (Équation 6). Le score en fluence n'existe qu'à partir des évaluations de mi-CP. Les erreurs types clusterisées au niveau de l'établissement fréquenté en 5^e sont présentées entre parenthèses. Seuils de significativité : « *** », pour un coefficient significatif à 0,1 %, « ** », pour un coefficient significatif à 1 %, « * », pour un coefficient significatif à 5 % et « . » pour un coefficient significatif à 10 %.

Champ : Élèves scolarisés en 5^e au sein des établissements de l'enquête EEC ayant adopté une forme unique de regroupement ou aucune en 5^e pour le français. Cela correspond à 1 289 collèges. Pour chaque niveau, le score en fluence est calculé uniquement pour les élèves ayant un score dans chacune des compétences retenues pour ce domaine.

Sources : DEPP, Évaluations nationales en français de CP (2018), de mi-CP (2019), de CE1 (2019), de 6^e (2023) et de 4^e (2025) ; enquête EEC, questionnaires Chefs d'établissement, printemps 2025 ; Constat de rentrée 2025.

TABLEAU 15 • Coefficients de l'estimation des effets des organisations pédagogiques sur les scores par domaine aux évaluations nationales en français de 4^e

Domaine	Nombres et calculs					
Comparaison	Hétérogène vs pas de regroupement	Hétérogène vs pas de regroupement	Selon le profil scolaire des élèves vs pas de regroupement	Selon le profil scolaire des élèves vs pas de regroupement	Selon le profil scolaire des élèves vs hétérogène	Selon le profil scolaire des élèves vs hétérogène
Modèle	DiD	TWFE	DiD	TWFE	DiD	TWFE
Coefficient	-0,018 (0,019)	0,001 (0,02)	-0,018 (0,019)	-0,019 (0,019)	0,003 (0,011)	-0,02 (0,012)
Crédibilité des tendances parallèles	Oui	Oui	Oui	Oui	Non, écart significatif en CP et en CE1	Non, écart significatif en CP et en CE1
Effets fixes	Non	Id élève + niveau	Non	Id élève + niveau	Non	Id élève + niveau
Cluster	UAI en 5 ^e	UAI en 5 ^e	UAI en 5 ^e	UAI en 5 ^e	UAI en 5 ^e	UAI en 5 ^e
N obs.	104 756	229 794	224 752	492 030	271 652	595 216
N élèves	54 581	54 866	117 297	117 866	141 808	142 508
N établissements	473	473	958	958	1 151	1 151

Lecture : Pour le domaine « nombres et calculs », le modèle de doubles différences sans effet fixe estime un coefficient de -0,018 pour l'effet des regroupements hétérogènes : le fait d'appartenir à la cohorte d'élèves qui entre en 5^e en 2024-2025 et d'avoir été scolarisé dans un établissement qui met en place des groupes hétérogènes plutôt que dans un établissement qui ne met en place aucune forme de regroupement diminue en moyenne le score en nombres et calculs de 0,018, soit 1,8 % d'un écart-type. Ce coefficient n'est cependant pas significatif.

Note : Les modèles DiD correspondent aux modèles de doubles différences classiques (Équation 5) tandis que les modèles TWFE désignent les modèles avec effets fixes (Équation 6). Les erreurs types clusterisées au niveau de l'établissement fréquenté en 5^e sont présentées entre parenthèses. Seuils de significativité : « *** », pour un coefficient significatif à 0,1 %, « ** », pour un coefficient significatif à 1 %, « * », pour un coefficient significatif à 5 % et « . » pour un coefficient significatif à 10 %.

Champ : Élèves scolarisés en 5^e au sein des établissements de l'enquête EEC ayant adopté une forme unique de regroupement ou aucune en 5^e pour les mathématiques. Cela correspond à 1 291 collèves. Pour chaque niveau, le score en nombres et calculs est calculé uniquement pour les élèves ayant un score dans chacune des compétences retenues pour ce domaine.

Sources : DEPP, Évaluations nationales en mathématiques de CP (2018), de mi-CP (2019), de CE1 (2019), de 6^e (2023) et de 4^e (2025) ; enquête EEC, questionnaires Chefs d'établissement, printemps 2025 ; Constat de rentrée 2025.

Domaine	Résolution de problèmes					
Comparaison	Hétérogène vs pas de regroupement	Hétérogène vs pas de regroupement	Selon le profil scolaire des élèves vs pas de regroupement	Selon le profil scolaire des élèves vs pas de regroupement	Selon le profil scolaire des élèves vs hétérogène	Selon le profil scolaire des élèves vs hétérogène
Modèle	DiD	TWFE	DiD	TWFE	DiD	TWFE
Coefficient	-0,011 (0,015)	-0,002 (0,017)	-0,011 (0,015)	-0,014 (0,016)	-0,007 (0,009)	-0,013 (0,01)
Crédibilité des tendances parallèles	Oui	Oui	Oui	Oui	Oui	Oui
Effets fixes	Non	Id élève + niveau	Non	Id élève + niveau	Non	Id élève + niveau
Cluster	UAI en 5 ^e	UAI en 5 ^e	UAI en 5 ^e	UAI en 5 ^e	UAI en 5 ^e	UAI en 5 ^e
N obs.	104 330	232 425	223 753	497 452	270 463	601 867
N élèves	54 534	54 860	117 189	117 845	141 675	142 481
N établissements	473	473	958	958	1 151	1 151

Lecture : Pour la résolution des problèmes, le modèle de doubles différences sans effet fixe estime un coefficient de -0,011 pour l'effet des regroupements hétérogènes : le fait d'appartenir à la cohorte d'élèves qui entre en 5^e en 2024-2025 et d'avoir été scolarisé dans un établissement qui met en place des groupes hétérogènes plutôt que dans un établissement qui ne met en place aucune forme de regroupement diminue en moyenne le score en résolution des problèmes de 0,011, soit 1,1 % d'un écart-type. Ce coefficient n'est cependant pas significatif.

Note : Les modèles DiD correspondent aux modèles de doubles différences classiques (Équation 5) tandis que les modèles TWFE désignent les modèles avec effets fixes (Équation 6). Les erreurs types clusterisées au niveau de l'établissement fréquenté en 5^e sont présentées entre parenthèses. Seuils de significativité : « *** », pour un coefficient significatif à 0,1 %, « ** », pour un coefficient significatif à 1 %, « * », pour un coefficient significatif à 5 % et « . » pour un coefficient significatif à 10 %.

Champ : Élèves scolarisés en 5^e au sein des établissements de l'enquête EEC ayant adopté une forme unique de regroupement ou aucune en 5^e pour les mathématiques. Cela correspond à 1 291 collèves. Pour chaque niveau, le score en résolution des problèmes est calculé uniquement pour les élèves ayant un score dans chacune des compétences retenues pour ce domaine.

Sources : DEPP, Évaluations nationales en mathématiques de CP (2018), de mi-CP (2019), de CE1 (2019), de 6^e (2023) et de 4^e (2025) ; enquête EEC, questionnaires Chefs d'établissement, printemps 2025 ; Constat de rentrée 2025.

Retrouvez les travaux de la DEPP sur

education.gouv.fr/depp

Publications et archives

Retrouvez toutes les publications et archives de la DEPP sur

archives-statistiques-depp.education.gouv.fr

Jeux de données en open data

Retrouvez tous les jeux de données de la DEPP en open data sur

data.education.gouv.fr

Ministère de l'Éducation nationale

Direction de l'évaluation, de la prospective et de la performance (DEPP)