

Dans le cadre du cycle d'évaluations disciplinaires réalisées sur échantillon (CEDRE), la reprise, en 2009, de l'évaluation de 2003 portant sur les compétences en compréhension de l'écrit des élèves en fin d'école primaire, permet la comparaison à six ans d'intervalle.

La proportion des élèves les plus performants (groupes 4 et 5) n'a pas changé (29 %). Ces élèves sont capables d'atteindre une compréhension fine des textes lus, de justifier une argumentation, de hiérarchiser différents types d'information.

Les élèves en difficulté (groupes 0 et 1) représentent 13 % de la population en 2009.

Ils peuvent prélever dans un texte des informations si elles sont explicites mais sont en difficulté face à des tâches plus complexes.

L'évolution statistiquement significative concerne la population du groupe 3 (moyens-forts). Ces élèves sont plus nombreux en 2009 qu'en 2003 (+ 2,6 %) (cf. graphique 2).

Compréhension de l'écrit en fin d'école Évolution de 2003 à 2009

Le cycle d'évaluations disciplinaires réalisées sur échantillon (CEDRE) a débuté en 2003. L'objectif est de faire le point sur les compétences et connaissances des élèves à des moments-clés du cursus scolaire, et ceci au regard des objectifs fixés par les programmes.

Ce dispositif permet une appréciation de l'évolution des acquis des élèves qui peut être mise en relation avec les caractéristiques du système et les politiques éducatives.

Cette évaluation, qui débute un nouveau cycle en reprenant la majorité des épreuves de 2003, permet une comparaison temporelle.

Dans le cadre de cette Note, c'est une étape importante du cursus scolaire qui est ciblée : la fin de la scolarité à l'école primaire.

Recueillir des informations sur l'ensemble des compétences exigibles à l'issue de l'école au moyen d'une seule prise d'informations reviendrait à créer un outil disproportionné demandant un temps de travail trop important aux élèves. C'est pourquoi le dispositif est organisé autour d'une rotation des disciplines selon les années.

La première évaluation CEDRE en fin d'école portant sur la langue française a été réalisée au printemps 2003. La prise d'information dont il est question ici a eu lieu six ans plus tard, au printemps 2009.

Le cadre de l'évaluation

« [...] *la maîtrise de la langue française* » constitue un des « *objectifs prioritaires du CE2 et du CM* ». C'est pourquoi le cadre de cette évaluation a été établi, dès la prise d'informations de 2003, en référence aux exigences des programmes, notamment celles du cycle 3 de l'école élémentaire.

Les programmes de juin 2008 indiquent que : « *Faire accéder tous les élèves à la maîtrise de la langue française relève d'abord de l'enseignement du français [...] mais aussi de toutes les disciplines : les sciences, les mathématiques, l'histoire, la géographie, l'éducation physique et les arts. [...] L'étude de la langue française (vocabulaire, grammaire, orthographe) est conduite avec le souci de mettre en évidence ses liens avec l'expression, la compréhension et la correction rédactionnelle.* » Ainsi, les compétences à acquérir visent à construire les connaissances et les savoir-faire propres aux différents domaines disciplinaires.

Les compétences en jeu

Les élèves, pour comprendre le sens des différents documents lus (textes, tableaux, cartes, etc.), doivent mobiliser un ensemble de processus mentaux qui conduisent de la reconnaissance de mots à la compréhension

fine des textes. Certains de ces processus, en particulier la reconnaissance de mots, deviennent progressivement automatisés au cycle 3 ; d'autres, plus complexes, nécessitent une analyse précise du document lu. Afin d'évaluer le niveau de maîtrise des élèves, on a retenu, pour la conception des épreuves, des compétences graduées en fonction d'une complexité croissante des opérations mentales nécessaires pour les mettre en œuvre :

- prélever une information explicite ;
- prélever une information implicite ;
- déduire une information à partir de deux propositions tenues pour « vraies » ;
- analyser un document et tirer les éléments essentiels qui aident à la compréhension de celui-ci ;
- synthétiser un document, en trouver le thème ou choisir un résumé.

Cette hiérarchisation des opérations mentales, construite *a priori* à partir des connaissances disponibles sur ce sujet, a été ensuite mise à l'épreuve lors de l'expérimentation du protocole. Pour la restitution dans cette Note, les compétences sont regroupées en trois blocs : Prélever (compétences 1 et 2) ; Organiser (compétences 3 et 4) ; Synthétiser (compétence 5).

L'épreuve 2009

Chacune de ces cinq compétences a été envisagée dans différents champs disciplinaires et évaluée à l'aide de plusieurs supports. Ainsi, l'évaluation de la compréhension de l'écrit a été organisée autour d'une vingtaine de situations relevant de champs tels que : littérature, mathématiques, sciences, histoire, géographie, vie quotidienne. Chaque situation était composée d'un texte, littéraire (récits, descriptions, poèmes), informatif ou documentaire, voire scolaire (énoncés de problèmes, consignes, manuels) et de sept à huit questions s'y référant. Dans les cahiers destinés aux élèves, les situations d'évaluation n'étaient pas regroupées par champ disciplinaire mais s'enchaînaient en passant d'un contexte disciplinaire à un autre, celui-ci étant clairement identifiable. Les documents proposés sont d'une nature et d'une complexité adaptées à l'âge des élèves de fin de cycle 3.

Mais la compréhension dépend également de compétences relevant de traits distinctifs qui donnent au texte sa cohérence, comme par exemple l'usage des pronoms, les temps verbaux ou encore les champs lexicaux. Ces aspects, liés à l'étude de la langue française (grammaire, orthographe), ont aussi donné lieu à des épreuves.

Un juste équilibre de reprises et de nouveautés

En 2003, la première évaluation CEDRE a été conçue par une équipe mise en place par la DEPP, constituée d'inspecteurs de l'éducation nationale, de professeurs d'instituts universitaires de formation des maîtres, de conseillers pédagogiques, de directeurs d'écoles d'application, de maîtres formateurs et de professeurs des écoles. Pour assurer la comparabilité dans le temps, une majorité des épreuves de l'époque ont été reprises à l'identique. C'est là une condition nécessaire pour créer un « ancrage » permettant d'apprécier les évolutions éventuelles des performances des élèves. Néanmoins, il a été nécessaire d'introduire de nouvelles épreuves, soit pour remplacer celles qui, à l'époque, avaient pu être jugées moins fiables ; soit pour renouveler les thèmes proposés ; soit, encore, pour élargir le format de réponses attendues de la part des élèves, en complétant les questions à choix multiples de réponses par des questions ouvertes pour lesquelles les élèves doivent rédiger leurs réponses ; soit, enfin, pour explorer des domaines de compétences qui ne l'avaient pas été en 2003.

Introduire de nouvelles épreuves dans un dispositif de ce type doit se faire en prenant un certain nombre de précautions méthodologiques. Il a été nécessaire, notamment, de valider les nouveaux items en pratiquant une expérimentation un an avant, en mai 2008. Cette expérimentation a permis de ne retenir que les items les plus pertinents répondant, par exemple, aux critères statistiques de stabilité, de fiabilité, etc.

L'échelle de performances

Pour synthétiser le constat et en faciliter la compréhension, une échelle (*graphique 1*)

décrivant six niveaux de performances a été construite en 2003. Cette échelle sert de référence pour la comparaison des résultats à six ans d'intervalle. Les compétences décrites en 2009 à chaque niveau de l'échelle sont identiques à celles de 2003. C'est le pourcentage d'élèves à chaque niveau qui est susceptible de varier, permettant ainsi de rendre compte des évolutions temporelles de leurs performances.

La construction de cette échelle est bien entendu basée sur l'ancrage constitué par les 80 items communs aux deux périodes. L'échelle de 2003 prend en compte 84 items supplémentaires utilisés seulement à cette époque. L'échelle de 2009 contient, outre les items communs, 80 items nouveaux créés spécifiquement. Au total donc, la répartition est équilibrée avec un tiers d'items anciens, un tiers d'items nouveaux et un tiers d'items communs.

Une maîtrise différenciée des compétences en compréhension

L'examen de l'échelle montre une gradation des compétences à l'œuvre en compréhension de l'écrit depuis la moins exigeante, le prélèvement de premier niveau (information explicite), jusqu'à la plus exigeante, l'analyse. Les groupes 0 et 1 réussissent très faiblement les items relevant des compétences les plus exigeantes. Les groupes 2 et 3 peuvent réussir des items relevant de compétences d'exigences variées, par exemple, réussir une inférence à partir de la lecture du texte et échouer au repérage des idées essentielles d'un texte pour le résumer. Les groupes 4 et 5 ont des performances élevées dans chaque compétence.

Quel que soit le groupe, la hiérarchie des taux de réussite entre les compétences est la même. Les compétences de prélèvement d'information sont les mieux maîtrisées et celles de synthèse et d'analyse les moins bien maîtrisées.

L'analyse des items réussis et des items échoués par les élèves des différents groupes, permet d'identifier les points forts et les points faibles de chaque groupe.

Au sein du **groupe 0**, les élèves ne sont capables que très ponctuellement de répondre à une sollicitation. Leur meilleure performance est obtenue avec le prélèvement d'informations explicites. Ces élèves

GRAPHIQUE 1 – Évaluations en 2003 et 2009 : répartition des élèves de CM2 selon l'échelle de compréhension de l'écrit
France métropolitaine

Score seuil	% Population en 2009	% Population en 2003	Échelle de compréhension de l'écrit
Groupe 5 315	9,9 % 10,0 %		Les élèves sont capables de choisir les informations les plus pertinentes du texte pour atteindre une compréhension fine de celui-ci ; de hiérarchiser les différentes catégories d'informations pour reconstituer le sens explicite ou implicite d'un texte ; de justifier une argumentation.
315 Groupe 4 276	19,2 % 19,2 %		Les élèves sont capables de repérer les idées essentielles d'un texte pour le résumer. Les élèves sont capables de repérer et d'utiliser les principales règles des outils du langage. Les élèves montrent qu'ils connaissent le lexique spécifique utilisé à l'école dans les différents champs disciplinaires.
276 Groupe 3 238	31,9 % 29,3 %		Les élèves sont capables d'identifier le thème, le sujet ou l'idée principale d'un texte. Les élèves sont capables de comprendre les informations implicites d'un texte. Les élèves sont capables de lier deux informations explicites dispersées dans le texte pour construire du sens. Les élèves sont capables de prélever des informations explicites quelle que soit la nature du support. Ils maîtrisent la compétence prélever dans des textes continus ou non-continus. Les élèves connaissent les règles d'accord (sujet/verbe - nom/adjectif) dans des exercices d'application.
238 Groupe 2 199	25,9 % 26,5 %		Les élèves sont capables de prélever des informations explicites contenues dans des documents préalablement organisés sous forme de tableaux, graphiques, schémas (lecture). Ils sont capables de prélever des informations à partir d'un texte continu entendu (écoute). Ces élèves restent sur l'aspect global du texte qu'ils viennent de lire. Ils ne rentrent pas dans les détails et ne reviennent pas sur l'écrit pour confirmer leur première approche. Ils ont tendance à échouer sur les textes continus qu'ils ont à lire. Par ailleurs, lorsque la tâche se complexifie et que l'attention qui leur est demandée augmente, leurs performances diminuent.
199 Groupe 1 161	10,9 % 12,1 %		Les élèves sont capables de prélever une information explicite d'une phrase ou rapidement disponible dans le texte (1 ^{er} paragraphe). Ils n'entrent que très rarement dans la globalité d'un texte et se trouvent en difficulté sur tout support dépassant un court paragraphe.
161 Groupe 0	2,2 % 2,9 %		Bien que capables de répondre ponctuellement à quelques questions, les élèves ne maîtrisent aucune des compétences attendues en fin d'école primaire.

Lecture : l'échelle précise la description des compétences maîtrisées par les élèves de chaque groupe. La première colonne symbolise l'étendue croissante de la maîtrise des compétences du groupe 0 au groupe 5. En 2009, les élèves du groupe 3 représentent 31,9 % de l'ensemble des élèves. Ils ont entre 238 et 276 points. Ils n'étaient que 29,3 % dans ce groupe en 2003. Ils sont capables de réaliser les tâches du niveau des groupes 0, 1, 2 et 3. En revanche leur probabilité de réussir les tâches spécifiques aux groupes 4 et 5 est faible.

semblent en très forte difficulté et ne donnent une réponse correcte que dans des tâches simples qui mettent en jeu une phrase dans laquelle la réponse est explicitement présente.

Dans le **groupe 1**, les élèves montrent des caractéristiques similaires à celles du groupe 0, mais avec un niveau de performance plus élevé. Leurs performances restent faibles, bien en deçà de celles de la moyenne de l'échantillon. Ils forment avec les élèves du groupe 0 l'ensemble des élèves en grande difficulté que diverses études détectent à l'entrée en sixième.

Différentes difficultés peuvent être repérées en ce qui concerne les élèves du **groupe 2** :

- la première est celle du « survol » des documents. Les élèves semblent rester au niveau d'une compréhension globale du texte. Ils ont tendance à échouer aux questions pour lesquelles il faut avoir lu une partie précise du texte ;

- la deuxième est concomitante : les élèves ont plus de difficultés avec un texte continu qu'avec un texte non continu. Les textes continus imposent une lecture fine de l'ensemble du texte alors que les textes non continus, du fait de leur organisation, de l'apport de graphiques, de schémas, etc., facilitent l'accès à l'information ;

- enfin, dès qu'une consigne impose des prises d'indices multiples ou que la réponse requiert une grande précision, leurs performances sont moins élevées.

Les élèves du **groupe 3** réussissent les items relevant des compétences de prélèvement, que celui-ci soit explicite ou implicite. Ils savent trouver le thème d'un document, voire en sélectionner le meilleur résumé, mais ont un taux de réussite faible lorsqu'il s'agit d'analyser. Ces élèves ne maîtrisent donc pas complètement les compétences de haut niveau permettant un accès à une compréhension fine d'un document.

On observe pour les élèves des deux **groupes 4 et 5** des performances qui montrent leur capacité à atteindre une compréhension fine des textes lus. Outre les compétences déjà évoquées pour les groupes précédents, ils peuvent justifier une argumentation, hiérarchiser différents types d'information, résumer en repérant les idées essentielles...

Comparaison 2003/2009

Le score moyen obtenu par les élèves de 2003 ayant été fixé arbitrairement à 250 (cf. encadré méthodologique), on note une légère augmentation globale en 2009 avec un score moyen de 252. Cette différence n'est pas significative d'un point de vue statistique. C'est donc une très grande stabilité qui est observée globalement sur la période. L'échelle de compréhension de l'écrit (*graphique 1*) montre comment se répartissent les élèves, à ces deux moments.

Les élèves des groupes 4 et 5 ont des performances qui amènent à considérer qu'ils maîtrisent de façon satisfaisante, ou à peu près satisfaisante, les compétences attendues par les programmes en fin d'école primaire en compréhension de l'écrit. La proportion de ces élèves dans la population est parfaitement stable à six ans d'écart (29 % de la population en 2003 comme en 2009).

À l'autre extrémité de l'échelle, les groupes 0 et 1 montrent des élèves en difficulté, voire en grande difficulté (groupe 0). Ces élèves représentaient 15 % de la population de CM2 en 2003. Cette proportion reste stable en 2009 pour les deux groupes (13 %), la baisse observée n'étant pas significative.

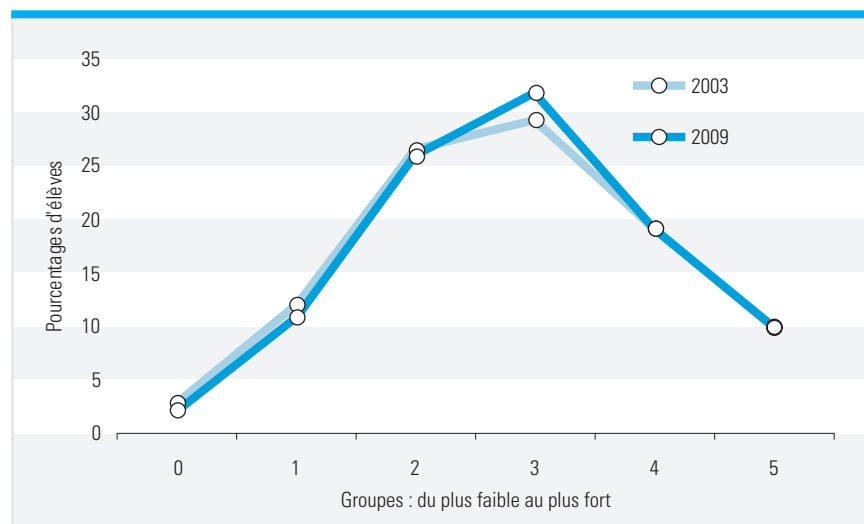
Les élèves du groupe 2 sont des élèves qui ne maîtrisent que très partiellement les différentes compétences retenues dans cette évaluation. La proportion d'élèves constituant ce groupe « moyen-faible »

reste stable elle aussi dans la population parente ; le passage de 26,5 % à 25,9 %, là non plus, n'est pas significatif. Les élèves de ce groupe méritent une attention particulière. Au regard de leurs performances, l'hypothèse selon laquelle ils n'ont pas atteint les objectifs du programme de la fin d'école primaire est vraisemblable. Il est donc pertinent de penser qu'ils risquent de rencontrer des difficultés d'apprentissage au collège.

Enfin, les élèves du groupe 3 maîtrisent des compétences comme le prélèvement d'une information, qu'elle soit explicite ou implicite, mais sont moins à l'aise lorsqu'il s'agit de compétences comme la synthèse ou l'analyse. Ce groupe d'élèves voit sa proportion augmenter au cours de la période, passant de 29,3 % à 31,9 %. Il faut noter que c'est le seul groupe de toute l'échelle où l'évolution est significative d'un point de vue statistique.

Au total donc, entre 2003 et 2009, à l'exception du groupe 3 (le groupe « moyen-fort ») les proportions d'élèves de chaque groupe restent stables pour l'ensemble de l'échelle. On peut toutefois remarquer que, même s'il s'agit d'une simple tendance puisqu'aucune significativité statistique ne se dégage, comme la proportion d'élèves des groupes 4 et 5 est parfaitement stable et que celle du groupe 3 augmente significativement, les élèves qui abondent ce groupe 3 sont issus des groupes 0, 1 et 2, c'est à dire des groupes les plus faibles.

GRAPHIQUE 2 – Performances en compréhension de l'écrit à la fin de l'école primaire (%)



Source : MENJVA - DEPP

Le graphique 3 montre que la hiérarchie dans la maîtrise des trois domaines de compétences n'a pas changé entre 2003 et 2009. Il indique une très légère hausse des réussites, répartie de manière homogène sur les trois compétences.

En 2009, la compétence « Prélever » est réussie en moyenne à 75 %. La réussite moyenne aux items communs évaluant la compétence « Organiser » s'élève à 71 %. Quant à la compétence « Synthétiser », la plus complexe, la réussite moyenne est de 68 %.

Des performances différenciées selon les caractéristiques des élèves

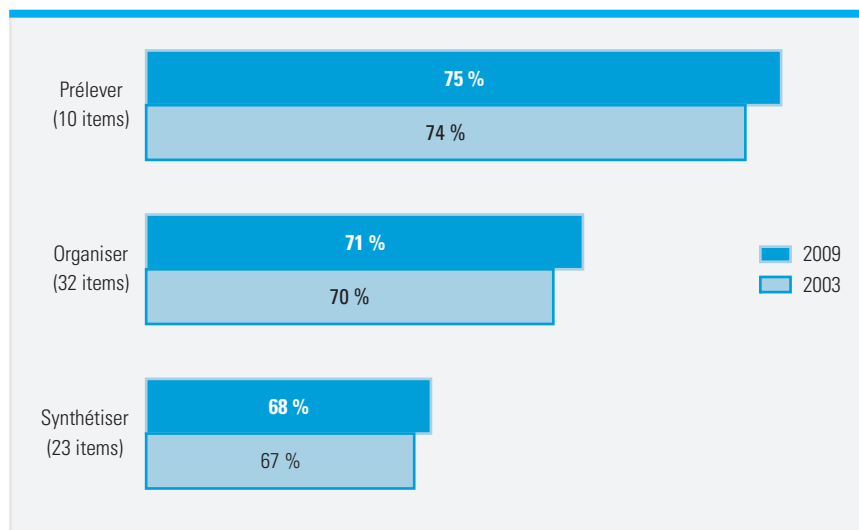
Les résultats mettent en évidence une évolution différenciée en fonction des caractéristiques sociodémographiques et scolaires des élèves.

Des performances qui varient selon le sexe des élèves...

Les filles obtiennent des performances supérieures à celles des garçons, tant en 2003 qu'en 2009, et ces différences sont significatives les deux années. Cette information n'étonnera pas le lecteur averti puisqu'elle a déjà été mise en évidence dans de nombreuses études portant sur les compétences de la langue ; néanmoins, il est intéressant de noter la stabilité des écarts de points.

Le tableau 1 montre la manière dont filles et garçons se répartissent dans l'échelle de compréhension au cours des deux prises d'informations. On remarquera que les filles, tout comme les garçons, sont légèrement moins nombreux dans les bas niveaux (groupes 0 et 1) en 2009 qu'en 2003 et que la proportion de filles progresse un

GRAPHIQUE 3 – Évolution de la réussite aux items communs selon les compétences entre 2003 et 2009
France métropolitaine



Source : MENJVA - DEPP

peu dans l'ensemble des niveaux supérieurs (groupes 4 et 5), celle des garçons ayant plutôt tendance à se tasser. Pour autant, la seule différence statistiquement significative concerne les filles du groupe 1 : elles sont 9,0 % en 2009 versus 11,4 % en 2003.

De cette répartition, il ressort toutefois qu'entre 2003 et 2009, la proportion de garçons parmi les élèves en difficulté a augmenté : en 2009, 60 % des élèves en difficulté (groupes 0 et 1) sont des garçons alors qu'ils ne représentaient que 54 % de ces groupes en 2003. À l'inverse, pour ce qui concerne les élèves aux performances supérieures (groupes 4 et 5), le pourcentage de filles est supérieur à celui des garçons en 2009 comme en 2003 (environ 55 %).

... selon le secteur scolaire et l'appartenance à l'éducation prioritaire...

Dans l'échantillon 2009, 16 % des élèves de CM2 étaient scolarisés dans une école privée sous contrat, 12 % dans une école publique en éducation prioritaire (EP) et

72 % dans une école publique hors EP. Cette répartition a très peu évolué puisque, en 2003, les pourcentages d'élèves des différents secteurs étaient respectivement de 16 %, 13 % et 71 %.

Le tableau 2 présente les scores moyens des élèves selon leur type de scolarisation. En 2003, les différences significatives ne concernaient que les établissements de l'éducation prioritaire (EP), dont les résultats étaient significativement inférieurs à ceux des deux autres secteurs, entre lesquels on n'observait pas de différence. Ces moindres performances de l'EP s'observent à nouveau en 2009. D'autre part, contrairement à 2003, l'écart de 7 points de score observé en 2009 entre les performances des élèves du privé et de ceux du public hors EP, au détriment de ces derniers, est statistiquement significatif. En ce qui concerne la période 2003/2009, on ne note aucune différence de score moyen au sein de chaque secteur, les hausses de 10 points pour l'éducation prioritaire et de 5 points pour le secteur privé n'étant pas significatives.

TABLEAU 1 – Répartition et scores moyens selon le sexe des élèves et répartition selon les groupes de niveaux en 2003 et en 2009
France métropolitaine

		Répartition (en %)	Score moyen	Groupe 0	Groupe 1	Groupe 2	Groupe 3	Groupe 4	Groupe 5
Ensemble	2003	100	250	2,9	12,1	26,5	29,3	19,2	10,0
	2009	100	252	2,2	10,9	25,9	31,9	19,2	9,9
Garçons	2003	49	247	3,6	12,8	27,2	28,9	18,7	8,8
	2009	50	249	2,8	12,7	26,3	32,0	17,4	8,9
Filles	2003	51	253	2,2	11,4	25,9	29,7	19,6	11,1
	2009	50	256	1,6	9,0	25,6	31,9	21,0	10,9

Lecture : les garçons représentent 49 % des élèves enquêtés en 2003 et 50,2 % en 2009. Leur score est passé de 247 à 249 ; 2,8 % d'entre eux appartiennent au groupe de niveau 0 en 2009 contre 3,6 % en 2003. (Les évolutions significatives sont marquées en gras et en bleu). Les valeurs sont arrondies à la décimale la plus proche, de manière à obtenir des sommes égales à 100 %.

Source : MENJVA - DEPP

TABLEAU 2 – Répartition et scores moyens selon les conditions de scolarisation des élèves et répartition selon les groupes de niveaux en 2003 et en 2009
France métropolitaine

		Répartition (en %)	Score moyen	Groupe 0	Groupe 1	Groupe 2	Groupe 3	Groupe 4	Groupe 5
Ensemble	2003	100	250	2,9	12,1	26,5	29,3	19,2	10,0
	2009	100	252	2,2	10,9	25,9	31,9	19,2	9,9
Élèves scolarisés dans le secteur public hors éducation prioritaire	2003	71	254	2,5	10,7	24,9	30,2	20,5	11,3
	2009	72	254	2,0	10,6	25,2	32,2	19,6	10,5
Élèves scolarisés dans des écoles de l'éducation prioritaire (EP)	2003	13	222	7,5	24,4	33,8	20,6	10,7	3,0
	2009	12	232	4,8	18,4	33,5	28,7	11,0	3,6
Élèves scolarisés dans le secteur privé	2003	16	256	0,8	8,0	27,8	33,1	20,2	10,1
	2009	16	261	1,4	6,4	23,5	33,0	23,8	12,0

Lecture : les élèves inscrits dans des écoles de l'éducation prioritaire représentent 13 % des élèves enquêtés en 2003 et 12 % en 2009. Leur score est passé de 222 à 232 ; 4,8 % d'entre eux appartiennent au groupe de niveau 0 en 2009 contre 7,5 en 2003. (Les évolutions significatives sont marquées en gras et en bleu). Les valeurs sont arrondies à la décimale la plus proche, de manière à obtenir des sommes égales à 100 %.

Source : MENJVA - DEPP

Pour ce qui est de l'évolution des pourcentages d'élèves, groupe par groupe, de 2003 à 2009, en fonction des conditions de scolarisation, seul le groupe 3 « moyen-fort » en EP évolue de manière significative, à la hausse. Cette augmentation est d'ailleurs importante puisque le pourcentage d'élèves de ce groupe passe de 20,6 % en 2003 à 28,7 % en 2009.

Une évolution notable apparaît par ailleurs, toujours pour l'EP : près d'un tiers des élèves (31,9 %) se situaient dans l'un des deux groupes faibles (0 et 1) en 2003 alors qu'ils en constituent moins d'un quart (23,2 %) en 2009.

Si l'on considère l'ensemble des élèves aux faibles performances (groupes 0 et 1), environ 9 % d'entre eux sont scolarisés dans le privé, en 2003 comme en 2009 ; en EP, leur proportion décroît de 25 % à 20 % entre 2003 et 2009 ; enfin, dans le public hors EP, elle passe de 66 % à 71 % entre les deux années.

... et selon le cursus scolaire

Comme en 2003, le fait d'être en retard ou « à l'heure » dans sa scolarité a un impact très important sur les scores moyens obtenus par les élèves. L'écart de performances est considérable : il représente près

d'un écart-type entre ces deux groupes, c'est-à-dire 45 à 50 points selon l'année. Vu leur importance, ces écarts, tant en 2003 qu'en 2009, sont bien entendu statistiquement significatifs ; en revanche, on n'observe pas d'évolution significative des scores de chaque groupe d'élèves d'une année à l'autre (tableau 3).

Tout comme pour les autres caractéristiques décrites ci-dessus, on ne constate pas d'évolution significative des proportions d'élèves, groupe par groupe, de 2003 à 2009, en fonction du cursus scolaire, à l'exception, à nouveau, du groupe 3 où les élèves « à l'heure dans leur scolarité » sont en proportion significativement plus importante en 2009 (34,4 %) qu'en 2003 (31,6 %).

Si l'on considère l'ensemble des élèves en difficulté (groupes 0 et 1), il faut noter qu'une proportion très importante d'entre eux (40 % en 2009) correspond à des élèves qui ont pris du retard dans le déroulement de leur scolarité. Cette proportion était encore plus importante (50 %) en 2003.

En revanche, dans les groupes d'élèves aux performances élevées et dont on considère qu'ils maîtrisent de façon satisfaisante les compétences attendues (groupes 4 et 5), la proportion d'élèves en retard est très faible (3 % en 2003 et 4 % en 2009).

Remarques

Les élèves français semblent plus à l'aise dans le prélèvement d'informations (compétence considérée comme de bas niveau) alors qu'ils éprouvent des difficultés dans l'analyse d'un document (compétence considérée comme de haut niveau). Ceci corrobore les résultats de l'évaluation internationale PIRLS. Cette évaluation a montré, pour des élèves de fin de CM1, que les Français se classent mieux à un niveau international lorsqu'il s'agit de mettre en œuvre, pour comprendre leur lecture, les compétences les plus simples (prélever des informations explicites ; faire des inférences directes) que lorsqu'il s'agit de mobiliser des compétences plus sophistiquées (interpréter et assimiler idées et informations ; apprécier le contenu, la langue et les éléments textuels). L'échelle construite et présentée dans cette Note propose un échelonnement des réussites des élèves. Elle illustre surtout leur hétérogénéité dans la maîtrise des compétences attendues en fin d'école primaire. Cette hétérogénéité est grande pour ces élèves qui arrivent en sixième et elle impose à l'école élémentaire une réflexion sur les difficultés qui se manifestent pour ces élèves en amont, dès le cycle 2.

TABLEAU 3 – Répartition et scores moyens selon le cursus scolaire des élèves et répartition selon les groupes de niveaux en 2003 et en 2009
France métropolitaine

		Répartition (en %)	Score moyen	Groupe 0	Groupe 1	Groupe 2	Groupe 3	Groupe 4	Groupe 5
Ensemble	2003	100	250	2,9	12,1	26,5	29,3	19,2	10,0
	2009	100	252	2,2	10,9	25,9	31,9	19,2	9,9
Élèves « à l'heure »	2003	81	260	1,1	7,9	24,4	31,6	22,8	12,2
	2009	85	259	1,1	8,0	23,9	34,4	21,5	11,3
Élèves en retard	2003	19	211	10,5	29,3	35,1	20,0	4,1	1,0
	2009	15	215	8,8	27,6	37,8	17,8	6,1	2,0

Lecture : les élèves « à l'heure » dans leur scolarité représentent 81 % des élèves enquêtés en 2003 et 85 % en 2009. Leur score est passé de 260 à 259 ; 8,0 % d'entre eux appartiennent au groupe de niveau 1 en 2009 contre 7,9 % en 2003. (Les évolutions significatives sont marquées en gras et en bleu). Les valeurs sont arrondies à la décimale la plus proche, et de manière à obtenir des sommes égales à 100 %.

Source : MENJVA - DEPP

Dans ce cadre, les outils d'évaluation diagnostique mis à la disposition des enseignants doivent permettre une meilleure détection des difficultés que rencontrent ces élèves et donc une meilleure prise en compte de celles-ci.

Marc Colmant,
Jeanne-Marie Daussin et
Pascal Bessonneau, DEPP B2

Pour en savoir plus

« Évolution des performances en lecture des élèves de CM1 - Résultats de l'étude internationale PIRLS », *Note d'Information 08.14*, MEN-DEPP, mars 2008.

« Compréhension à l'écrit et à l'oral des élèves en fin d'école primaire - Évaluation 2003 », *Les dossiers Enseignement scolaire n° 185*, MEN-DEPP, juillet 2007.

« La maîtrise du langage et de la langue française en fin d'école primaire », *Note Évaluation 04.10*, MENESR-DEP, octobre 2004.

www.education.gouv.fr
depp.documentation@education.gouv.fr

Méthodologie

Le CEDRE

Le cycle des évaluations disciplinaires réalisées sur échantillon (CEDRE) établit des bilans nationaux des acquis des élèves en fin d'école et en fin de collège. Il couvre les compétences des élèves dans la plupart des domaines disciplinaires en référence aux programmes. La présentation des résultats permet de situer les performances des élèves sur des échelles de niveau allant de la maîtrise pratiquement complète de ces compétences à une maîtrise bien moins assurée, voire très faible, de celles-ci. Renouvelées tous les six ans, ces évaluations permettent de répondre à la question de l'évolution du « niveau des élèves » au fil du temps.

Le calendrier des évaluations CEDRE

2003 : maîtrise de la langue (*Note Évaluation 04.10*).

2004 : langues étrangères (*Note Évaluation 05.06 et 05.07*).

2005 : attitudes à l'égard de la vie en société (*Note Évaluation 06.02*).

2006 : histoire, géographie et éducation civique (*Note d'Information 07.44*).

2007 : sciences (*Note d'Information 11.05*).

2008 : mathématiques (*Note d'Information 10.17*).

2009 : reprise du cycle, maîtrise de la langue.

L'échantillonnage

La population visée est celle des élèves de CM2 des écoles publiques et privées sous contrat de France métropolitaine.

— L'échantillon 2003

En 2003, 350 écoles ont été sélectionnées parmi les écoles de plus de six élèves de CM2.

Le tirage de l'échantillon a tenu compte de la taille des écoles avec une strate pour les écoles comprenant moins de quinze élèves de CM2 et une autre pour les écoles en comprenant au moins quinze. Dans chaque strate, une deuxième stratification a été effectuée sur le secteur de l'établissement : public hors éducation prioritaire, éducation prioritaire, et privé sous contrat. Tous les élèves de CM2 de ces écoles ont été évalués, ce qui représente environ 9 000 élèves. Les réponses de 6 109 élèves ont pu être analysées dans les 278 écoles ayant répondu.

— L'échantillon 2009

En 2009, 215 écoles ont été sélectionnées parmi les écoles de plus de six élèves de CM2, en vue d'une représentativité nationale.

Une stratification a d'abord été effectuée sur le secteur de l'établissement : public hors éducation prioritaire, éducation prioritaire, et privé sous contrat. Dans chaque strate, le tirage de l'échantillon a tenu compte de la taille des écoles en utilisant une strate pour les écoles comprenant moins de quinze élèves de CM2 et une autre pour les écoles en comprenant au moins quinze.

Dans chaque école, tous les élèves ont été évalués, soit environ 6 000 élèves. Les réponses de 4 952 élèves ont pu être analysées dans les 189 écoles répondantes.

Pour tenir compte de la non-réponse, les échantillons ont été redressés.

Contraintes de la situation d'évaluation

Pour l'enquête de 2003, en raison de contraintes techniques fortes (la taille de l'échantillon de départ interdisant une possibilité de correction par des experts), les 164 items sont du format « questions à choix multiples » (QCM). Cette forme de questionnement a pour avantage d'éviter un travail de correction et d'automatiser la saisie des réponses des élèves, mais a pour inconvénient de ne pas permettre l'évaluation des compétences des élèves en production d'écrit.

En 2009, des questions ouvertes, c'est-à-dire appelant une réponse rédigée, ont été introduites. Le test comportait 12 questions ouvertes et 148 QCM.

Par ailleurs, si chaque élève avait dû passer l'ensemble des situations proposées, au moins six heures d'évaluation par élève auraient été nécessaires. Pour limiter la passation à deux heures pour chaque élève,

les items ont été regroupés en douze « blocs » répartis ensuite dans douze cahiers différents (2003) ou en treize « blocs » répartis dans treize cahiers (2009), selon un plan expérimental organisant un « tuilage » des blocs. Ce dispositif, couramment utilisé dans les évaluations internationales, permet d'estimer la probabilité de réussite à chaque item sans que chaque élève ne passe l'ensemble des items.

La construction de l'échelle de performance

L'échelle de performance a été élaborée en utilisant les modèles de réponse à l'item. Le score moyen de maîtrise de la langue, correspondant à la performance moyenne des élèves de l'échantillon de 2003, a été fixé par construction à 250 et l'écart-type à 50. Cela implique qu'environ deux tiers des élèves ont un score compris entre 200 et 300. Mais cette échelle, comme celle de l'enquête internationale PIRLS par exemple, n'a aucune valeur normative et, en particulier, la moyenne de 250 ne constitue en rien un seuil qui correspondrait à des compétences minimales à atteindre.

Sur la base de constats fréquemment établis dans les différentes évaluations antérieures de la DEPP, qui montrent que 15 % des élèves peuvent être considérés en difficulté en fin d'école primaire, la partie la plus basse de l'échelle est constituée des scores obtenus par les 15 % d'élèves ayant les résultats les plus faibles. À l'opposé, la partie supérieure, constituée des scores les plus élevés, rassemble 10 % des élèves. Entre ces deux niveaux, l'échelle a été scindée en trois parties d'amplitudes de scores égales, correspondant à trois groupes intermédiaires.

Les modèles de réponse à l'item ont l'avantage de positionner sur la même échelle les scores des élèves et les difficultés des items. Cette correspondance permet de caractériser les compétences maîtrisées par chacun des groupes d'élèves.

La comparabilité 2003/2009

Afin de pouvoir comparer les résultats des enquêtes 2003 et 2009, 80 items d'ancrage, *i.e.* des items de 2003, ont été repris à l'identique dans l'évaluation de 2009.

Lors de l'analyse des résultats de 2009, les modèles de réponse à l'item ont été réutilisés et appliqués cette fois à l'ensemble des résultats 2003 et 2009.

L'estimation conjointe des modèles de réponse à l'item à partir des données de 2003 et de 2009, ainsi que la présence d'items communs aux deux évaluations, permet la comparaison directe à la fois des scores des individus et des difficultés des items entre les deux passations.

Il est à noter qu'un soin particulier a été apporté à l'analyse de ces items communs à 2003 et 2009. Ainsi, les items retenus dans l'analyse finale devaient, d'une part, ne pas présenter de fonctionnement différentiel (par exemple un écart de taux de réussite entre les deux passations anormalement élevé), d'autre part, devaient avoir des propriétés psychométriques satisfaisantes.

Les données de 2003 ont donc été réanalysées dans cette perspective de comparaison. C'est pourquoi certaines différences – notamment sur la valeur des scores-seuils – peuvent apparaître par rapport à la publication initiale de 2004.

Significativité

Les évaluations du CEDRE sont réalisées sur échantillons. De ce fait, les résultats sont soumis à une variabilité qui dépend des erreurs d'échantillonnage. Il est possible d'estimer statistiquement ces erreurs d'échantillonnage et de produire des intervalles de confiance. À titre d'exemple, le score moyen des élèves sur l'échelle de maîtrise de la langue est de 252 en 2009, mais le vrai score, tel qu'il serait calculé pour l'ensemble des élèves de CM2, se situe, avec une probabilité de 95 %, entre 249 et 255 (c'est-à-dire ± 3 points). Par conséquent, le score moyen des élèves de 2009 n'est pas significativement différent de celui des élèves de 2003, qui est de 250 (± 3 points).