

L'évolution du nombre d'élèves en difficulté face à l'écrit depuis une dizaine d'années

Jeanne-Marie Daussin, Saskia Keskaik, Thierry Rocher*

Depuis une dizaine d'années, le pourcentage d'élèves en difficulté face à l'écrit a augmenté de manière significative et près d'un élève sur cinq est aujourd'hui concerné en début de 6^e. Si le niveau de compréhension de l'écrit des élèves moyens n'a pas évolué, la plupart des évaluations témoignent d'une aggravation des difficultés parmi les élèves les plus faibles. Alors que la maîtrise des mécanismes de base de la lecture reste stable, les compétences langagières (orthographe, vocabulaire, syntaxe) sont en baisse, ce qui explique l'aggravation du déficit de compréhension des textes écrits, parmi les élèves les plus faibles. En moyenne, les filles ont de meilleures performances que les garçons dans le domaine de la compréhension de l'écrit ; cet écart s'accroît dans la quasi-totalité des pays de l'OCDE depuis une dizaine d'années. En France, le statut économique, social et culturel des parents explique aujourd'hui une plus grande part de la variation des scores des élèves qu'en moyenne dans l'ensemble des pays de l'OCDE. C'est dans les collèges en zones d'éducation prioritaire que l'augmentation des difficultés est la plus marquée : près d'un tiers de ces collégiens éprouvent des difficultés face à l'écrit, contre un quart il y a dix ans. Les élèves de ZEP d'aujourd'hui ne sont peut-être pas tout à fait comparables à ceux d'hier, toutefois la composition sociale de ces collèges semble plutôt stable.

Cet article propose de faire la synthèse de résultats récents, qui concernent l'évolution du niveau des élèves dans le domaine de la compréhension de l'écrit depuis une dizaine d'années. Ces résultats sont tirés d'un ensemble d'évaluations d'élèves. Un fait saillant, commun à nombre de ces évaluations, concerne l'augmentation du pourcentage d'élèves en difficulté face à l'écrit ces dix dernières années. Ce phénomène mérite cependant d'être précisé, selon les compétences ou les moments de la scolarité qui sont évalués (fin d'école primaire ou fin de collège). Les données renseignent également sur l'évolution du degré d'équité, à travers les différences de performances observées entre filles et garçons, entre élèves d'origines sociales différentes ou entre zones de scolarisation.

Il y a une vingtaine d'années, l'ouvrage de Christian Baudelot et Roger Establet, « Le niveau monte », pointait le manque d'informations objectives sur le sujet du niveau des élèves, pour lequel la demande sociale est très forte. Depuis, des dispositifs d'évaluation des élèves se sont développés, de manière à répondre à cette demande et à envisager une évaluation globale des résultats du système éducatif (*encadré 1*). Il ne s'agit pas ici d'analyser les causes des évolutions qui peuvent être, à l'évidence, d'ordre scolaire (à travers les politiques éducatives mises en place, le temps consacré à l'apprentissage de la langue, la formation des enseignants, etc.) ou d'ordre sociétal (à travers l'évolution de la relation à l'écrit par exemple). Cet article vise plutôt à établir un diagnostic précis de la situation, indispensable pour envisager des pistes ciblées d'amélioration.

* Jeanne-Marie Daussin, Saskia Keskaik, Thierry Rocher, Depp.

Brève histoire de la question de l'évolution du niveau des élèves

Malgré une demande sociale forte et récurrente, il est frappant de constater que les évaluations visant à comparer les compétences des élèves à différentes époques sont relativement récentes dans le paysage des enquêtes statistiques.

L'intérêt pour ce type d'enquêtes trouve son origine dans le débat sur la baisse supposée du niveau scolaire des élèves, débat qui semble être particulièrement vif en France à la fin des années 1980. En 1992, dans un rapport au ministre de l'Éducation nationale, Claude Thélot, alors directeur de la Dep (Direction de l'évaluation et de la prospective, aujourd'hui Depp), s'interroge sur les raisons de ce sentiment d'inquiétude et dégage trois pistes d'explications [Thélot, 1992]. Premièrement, après une période de massification du système éducatif, l'attention est portée sur la qualité et donc le niveau de compétences des élèves. Un système de masse peut-il être performant ? Deuxièmement, dans la lignée du rapport alarmiste américain « *A Nation at Risk* » de 1983, avec le renforcement de la compétition économique au niveau international, l'élévation du niveau des élèves apparaît comme un levier indispensable. Enfin, le sentiment de déclin du système éducatif porterait moins sur les mathématiques et les sciences, disciplines valorisées socialement, que sur les lettres et les humanités. Ce distinguo traduirait une appréhension profonde quant à l'avenir du pays, sa langue et son identité. Mais ce sentiment est-il légitime ?

En 1989, Christian Baudelot et Roger Establet avaient déjà pris le contre-pied de cette idée de décadence avec leur ouvrage devenu célèbre : « *Le niveau monte* » [Baudelot, Establet, 1989]. Pour étayer leur thèse, les auteurs s'étaient appuyés sur un ensemble de sources variées (comme par exemple des séries statistiques sur les diplômes ou sur les qualifications) qui sont en réalité, pour la majorité d'entre elles, des indicateurs indirects du niveau de compétences des individus. L'utilisation de ces indicateurs pose un certain nombre de problèmes, en termes de comparabilité dans le temps. En effet, le degré d'exigence de l'institution concernant le niveau requis pour l'obtention d'un diplôme est susceptible d'évoluer, la valeur d'un diplôme sur le marché du travail également. Pour rappel, pendant les « dix glorieuses », entre 1985 et 1995, le pourcentage de bacheliers dans une génération est passé de moins de 30 % à un peu plus de 60 %. Cet effort quantitatif s'est-il fait au

détriment de la valeur du baccalauréat, c'est-à-dire du niveau exigible de connaissances et de compétences pour l'obtenir ? L'organisation française des examens ne permet pas de répondre à cette question : les épreuves changent chaque année, sans qu'aucune procédure d'ajustement ne soit mise en œuvre. Ainsi, il n'est pas possible de conclure à l'élévation du niveau des élèves sur la base d'une augmentation de la note moyenne obtenue aux épreuves du baccalauréat.

De fait, à l'époque du livre de Christian Baudelot et Roger Establet, il existait très peu d'enquêtes comparatives rigoureuses dans ce domaine. La seule mesure directe des compétences, comparable dans le temps, utilisée par ces auteurs est celle du « niveau général » des conscrits, qui fournit un indicateur fruste de l'évolution du niveau des jeunes garçons¹. La comparaison des performances des promotions 1967 et 1982 concluait à une augmentation générale de la performance à ces tests, augmentation qui n'était pas homogène, et qui concernait surtout les plus diplômés.

Face au peu de données disponibles², la Dep a réalisé dans les années 1990, une série d'enquêtes comparatives. Cependant, les méthodologies employées pouvaient présenter quelques faiblesses car il s'agissait le plus souvent de mesurer des évolutions en utilisant des données qui n'avaient pas été conçues dans cette optique de comparaison temporelle. La France a pourtant une longue expérience des évaluations standardisées des élèves, à travers la mise en place des évaluations nationales diagnostiques passées par tous les élèves de CE2 et de 6^e, à chaque rentrée scolaire, entre 1989 et 2007. Malheureusement, aucun ajustement de la difficulté des épreuves n'a été entrepris afin de distinguer ce qui relève de la difficulté des épreuves de ce qui relève du niveau des élèves (*encadré 3*). D'une part, l'objectif premier de ces évaluations, tout comme celui des examens, n'était pas de rendre compte de l'évolution du niveau des élèves dans le temps mais de servir d'outils de repérage des difficultés pour les enseignants. D'autre part, les connaissances dans le domaine de la mesure en éducation et de la psychométrie sont encore relativement peu développées et vulgarisées en France, qui accuse un retard par rapport à d'autres pays où des systèmes rigoureux d'évaluations standardisées existent depuis plusieurs décennies (USA, Royaume-Uni, Pays-Bas).

Encadré 1 (suite)

Il faut attendre le rapport de 2001 du HCcéé (Haut Conseil de l'évaluation de l'école) qui pointe à son tour le manque d'informations objectives sur ce sujet et qui recommande la mise en place d'un dispositif *ad hoc* de suivi de l'évolution des acquis des élèves dans le temps [Salines, Vrignaud, 2001]. En 2003, la Depp donne ainsi naissance au cycle des évaluations CEDRE (Cycle des Évaluations Disciplinaires Réalisées sur Échantillons) qui évalue les acquis des élèves de CM2 et de 3^e, au regard de ce qui est attendu par les programmes scolaires. Chaque année, le domaine évalué est différent et pour la première

fois avec l'évaluation de 2009, une comparaison temporelle est possible (encadré 2).

Le développement de dispositifs d'évaluations permettant d'assurer la comparabilité des résultats dans le temps est aujourd'hui encouragé par le souci de construire des indicateurs de suivi, pour le pilotage du système, avec par exemple la construction d'indicateurs annuels de résultats dans le cadre de la Lolf (Loi organique relative aux lois de finances). Parallèlement, l'essor des évaluations internationales (PISA, PIRLS), organisées en cycles, contribue également à accorder de l'importance à la dimension temporelle.

1. De 1954 à 1995, les centres militaires faisaient passer des tests aux conscrits pendant les « trois jours ». L'objectif de ces tests était orienté vers la sélection des conscrits et leur éventuel recrutement. Ils étaient censés mesurer un « niveau général » en mélangeant des tests de raisonnement, d'intelligence pratique, de vocabulaire ou encore de mathématiques.

2. Paradoxalement, les évaluations des élèves sont très présentes dans le système scolaire français, à travers les contrôles continus fréquents conduits par les enseignants. Cependant, des études docimologiques, menées depuis près d'un siècle, montrent que le jugement des élèves par les enseignants est en partie emprunt de subjectivité et peut dépendre de facteurs étrangers au niveau de compétences des élèves. La notation des élèves est ainsi susceptible de varier sensiblement selon les caractéristiques des enseignants, des contextes scolaires, ainsi que des élèves eux-mêmes. Il est donc préférable de s'appuyer sur des dispositifs d'évaluations standardisées si l'on souhaite mesurer avec fiabilité l'évolution des performances dans le temps.

Des dispositifs multiples et complémentaires, une méthodologie commune

Les résultats présentés dans cet article synthétisent les résultats publiés de six évaluations relatives à la compréhension de l'écrit : deux évaluations issues du dispositif CEDRE, cycle d'évaluations disciplinaires réalisées sur échantillons (MDL : maîtrise de la langue en CM2, CG : compétences générales en 3^e) ; deux comparaisons dites « historiques », c'est-à-dire de plus long terme, à dix et vingt ans d'intervalle en fin de primaire (LEC, « Lire, écrire, compter ») et en début de 6^e (SPEC6, étude spécifique des difficultés de lecture à l'entrée en 6^e) ; et enfin deux évaluations internationales, PISA et PIRLS (encadré 2). Les autres évaluations existantes n'ont pas été exploitées ici car les analyses comparatives pouvant être menées concernent des périodes trop courtes. Rappelons que les évolutions du système éducatif s'opèrent sur des temps longs et que des comparaisons annuelles ont peu de chances de révéler des évolutions significatives. Enfin, d'autres domaines de compétences ont également fait l'objet d'études comparatives (comme le calcul en fin de CM2) ou donneront lieu à des résultats prochainement (en particulier les langues vivantes). Cet article se concentre sur les comparaisons portant sur les compétences des élèves face à l'écrit.

Si des constats communs forts se dégagent, ces évaluations n'aboutissent pas toutes à une conclusion ferme et univoque sur « le » niveau des élèves et son évolution. Certaines différences apparaissent et renvoient à la forme et au contenu de ces évaluations. La mesure des compétences, telle qu'elle est opérée à travers des tests standardisés, est un « construit ». Les aspects méthodologiques méritent donc d'être abordés, d'autant plus qu'ils relèvent d'un champ d'application de la statistique peu développé aujourd'hui en France, celui de la psychométrie, qui concerne la mesure des dimensions psychologiques, en l'occurrence des acquis cognitifs.

Ces évaluations ont en commun de partager les mêmes principes méthodologiques, qu'il s'agisse de la sélection des échantillons, de la passation des épreuves, de la correction des réponses ou des techniques statistiques et psychométriques employées pour assurer la

comparabilité des résultats. La comparaison dans le temps repose sur la reprise à l'identique d'un ensemble d'items (c'est-à-dire de questions), d'une vague à l'autre de l'enquête¹. Ces items repris d'une enquête à l'autre servent d'« ancrage » pour estimer l'évolution du niveau de compétences des élèves dans le temps. Autrement dit, on considère que la difficulté de ces items est identique et ne dépend pas de la période à laquelle ils sont administrés. Si la réussite à des items repris diminue, c'est donc le niveau de compétences des élèves qui a évolué à la baisse.

Cependant, il convient de distinguer ce qui est observé (la performance) de ce qui est visé (la compétence). Des facteurs externes peuvent agir sur le niveau de performances, indépendamment du niveau de compétences. Il est donc possible que certains items soient moins bien réussis ou mieux réussis qu'il y a dix ou vingt ans pour des raisons diverses, sans rapport avec le niveau de compétences des élèves : par exemple si le degré de familiarité avec certains types de tâches a évolué, si des apprentissages ne sont plus exigés par les programmes, si des consignes de correction n'ont pas été appliquées à l'identique, etc. L'emploi de modèles statistiques adaptés (les modèles de réponse à l'item) permet de repérer ces items « biaisés » (encadré 3).

Encadré 2

Les dispositifs d'évaluation des élèves en France : sous le signe de la comparabilité, dans le temps ou dans l'espace

CEDRE

Le cycle CEDRE (Cycle des Évaluations Disciplinaires Réalisées sur Échantillons), évalue chaque année les connaissances et les compétences des élèves de fin de CM2 et de fin de 3^e. au regard des objectifs fixés par les programmes scolaires.

Le premier cycle a été mené entre 2003 et 2008 de la manière suivante :

- 2003 : maîtrise du langage et de la langue française (MDL) en CM2, compétences générales (CG) en 3^e ;
- 2004 : langues vivantes (anglais, allemand et espagnol) en CM2 et en 3^e ;
- 2005 : attitudes des élèves à l'égard de la vie en société en CM2 et en 3^e ;
- 2006 : histoire, géographie et éducation civique en CM2 et en 3^e ;
- 2007 : sciences expérimentales (sciences de la vie et de la Terre, sciences physiques et chimiques) en CM2 et en 3^e ;
- 2008 : mathématiques en CM2 et en 3^e.

Depuis 2009, un second cycle est relancé, les enquêtes se déroulant dans le même ordre chronologique. La reprise partielle des épreuves d'une année à l'autre, permet d'assurer la comparabilité des résultats dans le temps.

Comparaisons « historiques » (de long terme)

LEC - « Lire, écrire, compter » (1987-2007) : en 2007, la reprise de l'enquête « Lire, écrire, compter » de 1987, portant sur la lecture, le calcul et l'orthographe en fin de CM2, a permis de comparer les performances des élèves à vingt ans d'intervalle, à partir des résultats observés aux mêmes épreuves. De plus, des points intermédiaires de comparaison sont disponibles, en 1997 pour la lecture et en 1999 pour le calcul.

SPEC6 - Difficultés de lecture à l'entrée en 6^e (1997-2007) : en septembre 1997 et en septembre 2007, des échantillons d'élèves de 6^e ont été soumis à la même évaluation, portant spécifiquement sur les difficultés de lecture.

Comparaisons internationales

PIRLS (2001-2006) : l'enquête internationale PIRLS (*Progress in International Reading Literacy Study*) vise à mesurer les performances en lecture des élèves à la fin de leur quatrième année de scolarité obligatoire (CM1 pour la France). En 2006, 45 pays ou provinces ont participé à l'enquête.

PISA (2000-2009) : tous les trois ans depuis 2000, sous l'égide de l'OCDE, l'évaluation internationale

1. Par exemple, pour l'évaluation CEDRE CG de 2009, les élèves ont passé au final 184 items dont 78 provenaient de l'évaluation de 2003, date à laquelle les élèves avaient passé 212 items en tout.

Encadré 2 (suite)

PISA (Programme for International Student Assessment) mesure et compare les compétences des élèves de 15 ans dans trois domaines : compréhension de l'écrit, culture mathématique et culture scientifique.

Au-delà des comparaisons internationales, ces deux évaluations ont également pour objectif de mesurer les progrès des pays à travers l'évolution temporelle des résultats obtenus aux tests, repris en partie d'une vague d'enquête à l'autre.

De nombreux autres dispositifs d'évaluation des élèves existent mais ils ne permettent pas de répondre à la question de l'évolution du niveau des élèves dans le temps, soit parce que les séries de données comparables sont trop récentes (indicateurs de performance de la Lolf, tests de détection de l'illettrisme lors des Journées Défense et Citoyenneté, évaluations nationales CE1/CM2, etc.) soit parce qu'ils ont d'autres objectifs (les suivis de « panels » d'élèves par exemple).

Description des évaluations comparatives utilisées

Nom	Opérateur	Années	Population visée	Échantillons	Contenu	Format	Publications ¹
1^{er} degré							
LEC <i>Enquête «Lire, écrire, compter»</i>	Depp	1987, 1997, 2007 (1999 pour le calcul)	CM2 en France métropolitaine (secteur public uniquement en 1987 et 1997)	1 994 élèves en 1987 3 108 élèves en 1997 4 188 élèves en 2007	Compréhension de l'écrit, orthographe, grammaire, calcul	Textes courts, questions ouvertes, dictée, calculs posés	NI 08.38
SPEC6 <i>Test spécifique de lecture à l'entrée en 6^e</i>	Depp	1997, 2007	CM2 en France métropolitaine	2 631 élèves en 1997 9 626 élèves en 2007	Dimensions spécifiques de la lecture : automaticité, connaissances lexicales, phonologiques, morphologiques et orthographiques, compréhension d'énoncés	QCM, épreuves chronométrées	NI 99.48, EE 2010
PIRLS <i>Progress for International Reading Literacy Survey</i>	IEA (Depp en France)	2001, 2006	Classe modale des élèves de 9 ans (CM1 en France)	3 538 élèves en 2001 4 404 élèves en 2006	Compréhension de l'écrit (<i>reading literacy</i>)	Textes longs, QCM/ questions ouvertes	NI 03.22, NI 08.14
CEDRE MDL <i>Cycle des Évaluations Disciplinaires Réalisées sur Échantillons - Maîtrise du Langage</i>	Depp	2003, 2009	CM2 en France métropolitaine	6 109 élèves en 2003 4 952 élèves en 2009	Compréhension de l'écrit, outils de la langue, compréhension orale (2003), expression écrite (2009)	Textes longs, QCM/ questions ouvertes, en 2003 uniquement QCM	NE 04.10, EE 2011 à paraître
2nd degré							
PISA <i>Programme for International Student Assessment</i>	OCDE (Depp en France)	2000, 2003, 2006, 2009	Élèves de 15 ans (quel que soit leur cursus scolaire) en France	4 673 élèves en 2000 4 300 élèves en 2003 4 716 élèves en 2006 4 298 élèves en 2009	Concept de littératie : compréhension de l'écrit, culture mathématique, culture scientifique	Supports variés, QCM/ questions ouvertes	NI 01.52, NE 04.12, NI 07.42, NI 08.08, NI 10.23, NI 10.24
CEDRE CG <i>Cycle des Évaluations Disciplinaires Réalisées sur Échantillons - Compétences Générales</i>	Depp	2003, 2009	3 ^e générale de collège en France métropolitaine	18 222 élèves en 2003 5 142 élèves en 2009	Compétences générales appliquées aux différentes disciplines	Questions variées, QCM/ questions ouvertes, en 2003 uniquement QCM	NE 04.09, NI 10.22

1. Les références renvoient aux numéros des Notes d'information (NI) et Notes d'évaluation (NE) de la Depp, ou bien à l'année d'édition de la revue l'État de l'École (EE).

Source : Depp.

Assurer la comparabilité des résultats

Les modèles de réponse à l'item : des modélisations psychométriques adaptées

À partir de ses réponses aux items, il est possible d'estimer le niveau de compétences de chacun des élèves sur une échelle de scores. Le calcul des scores des élèves s'inscrit généralement dans le cadre des « modèles de réponse à l'item », couramment employés dans les évaluations nationales et internationales et particulièrement adaptés aux questions de comparabilité. En l'occurrence, toutes les évaluations présentées ici ont recours à ce type de modélisation, à l'exception du test spécifique de lecture à l'entrée en 6^e (SPEC6) dont le format est particulier (des mesures de vitesse sont associées à des mesures de performance).

Les modèles de réponse à l'item postulent que la probabilité qu'un élève réussisse un item dépend, d'une part, du niveau de compétences de l'élève, et d'autre part de certaines caractéristiques de l'item, en particulier sa difficulté.

Le modèle de réponse à l'item le plus simple peut s'écrire de la manière suivante :

$$P(Y_i^j = 1 | \theta_i, b_j) = \frac{\exp(\theta_i - b_j)}{1 + \exp(\theta_i - b_j)}$$

où Y_i^j est la réponse de l'élève i à l'item j (Y_i^j vaut 1 lorsque la réponse est juste et 0 sinon), b_j est le paramètre de difficulté de l'item j et θ_i le niveau de compétences de l'élève i .

Les élèves sont définis par leur niveau de compétences, qui correspond à leur position sur la dimension latente (ou trait latent) θ . Les items, quant à eux, sont caractérisés par leur niveau de difficulté (b_j). Très simplement, la formule ci-dessus implique notamment que la probabilité que l'élève i réussisse l'item j augmente lorsque le niveau de compétences de l'élève i augmente et diminue lorsque le niveau de difficulté de l'item j augmente.

Il est également possible d'ajouter un paramètre dit de discrimination (utilisé dans CEDRE) qui permet à la relation entre réussite à l'item et niveau de compétences d'être plus ou moins marquée selon le niveau de compétences, ou un paramètre dit de pseudo-chance (utilisé dans PIRLS) qui modélise le fait que certains items, comme les QCM, puissent être réussis au hasard.

Les valeurs des paramètres b_j et θ_i sont définies à une transformation linéaire près. Il faut donc imposer des contraintes identifiantes, en l'occur-

rence sur la moyenne et l'écart-type des θ_i (c'est-à-dire sur l'origine et l'unité de l'échelle des scores) pour la première évaluation d'une série. Dans CEDRE, par convention, la moyenne des scores a été fixée à 250 et leur écart-type à 50 en 2003 ; alors que dans PISA, la moyenne internationale a été fixée à 500 et l'écart-type à 100 en 2000. Ces valeurs n'ont pas de sens dans l'absolu. Comme pour une échelle de température en degrés Celsius ou Fahrenheit, l'objet de la mesure reste le même mais il est possible d'opérer des transformations sur l'échelle de manière arbitraire.

L'avantage de ce type de modèle est de séparer les concepts : le niveau de compétences des élèves est défini indépendamment du niveau de difficulté des items, et inversement. Dès lors, si l'on estime les paramètres de difficulté des items sur un premier échantillon, et qu'un autre échantillon passe ces items dix ans plus tard, il suffira d'appliquer les valeurs de ces paramètres, considérées comme fixes, pour en déduire le niveau de compétences des élèves dix ans plus tard.

Cette méthode n'est cependant valide que si l'hypothèse d'unidimensionnalité est respectée. En effet, d'après le modèle, un élève est uniquement caractérisé par son niveau sur le continuum θ . Autrement dit, à niveau égal sur ce continuum, il ne devrait pas y avoir de différences de probabilité de réussite entre les élèves, quelle que soit la cohorte évaluée. Or, il peut arriver qu'un item ne soit pas réussi de la même manière selon les cohortes, après avoir contrôlé du niveau de compétences. On dira que cet item est biaisé : à niveau de compétences égal, un groupe d'élève est avantagé ou désavantagé par rapport à un autre groupe d'élèves. Cet item sera généralement éliminé lors de l'analyse.

Le degré de précision des évaluations

Dans le cadre d'un modèle de réponse à l'item, le niveau de compétences de chaque élève fait l'objet d'une estimation à partir de ses réponses aux items. Il est possible de quantifier le degré de précision de cette estimation pour chaque élève, c'est-à-dire de déterminer avec quel degré de certitude il est possible de se prononcer sur son niveau de compétences sur le continuum. Cette précision va dépendre de certaines caractéristiques du test, en particulier de la difficulté des items qui le composent. Par exemple, un test composé d'items faciles permettra de mesurer avec précision le niveau de compétences des élèves les plus faibles.

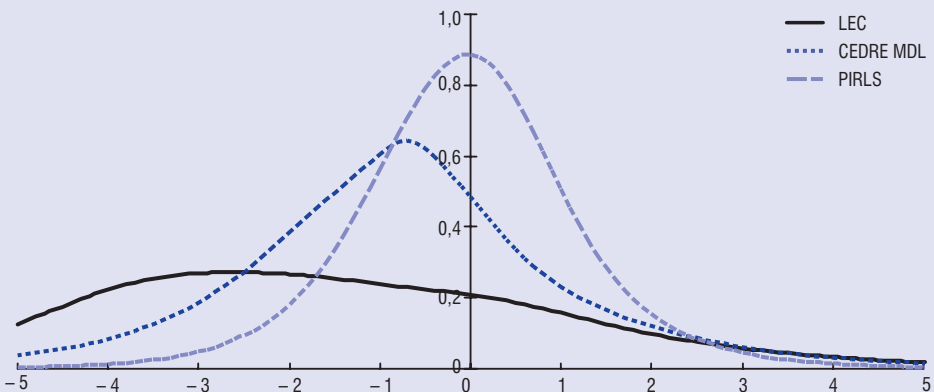
Si ces évaluations reposent sur les mêmes fondements méthodologiques, elles se distinguent néanmoins sur certains aspects qui compliquent l'exercice de la synthèse. Tout d'abord, les dates des points de comparaison diffèrent d'un dispositif à l'autre. Plus important, les populations visées ne sont pas les mêmes : en primaire, il peut s'agir des élèves de CM1, de CM2 ou de début de 6^e ; dans le second degré, le champ peut concerner un niveau scolaire (élèves de 3^e) ou un âge (élèves de 15 ans)². Surtout, les contenus des évaluations sont différents, ce qui conduit à sortir d'une vision univoque du niveau des élèves : certaines dimensions peuvent évoluer, d'autres pas. Ainsi, les évaluations CEDRE portent sur les compétences et les connaissances attendues, telles que fixées par les programmes nationaux. Elles sont par conséquent fortement ancrées dans le cadre scolaire. En l'occurrence, l'évaluation CEDRE CG vise des compétences générales ou procédurales (savoir-faire) attendues en fin de collège,

Encadré 3 (suite)

Une statistique permet de mesurer la précision du test selon le niveau de compétences : il s'agit de l'information du test. La figure ci-dessous représente l'information de trois tests retenus ici pour le premier degré : PIRLS 2006, CEDRE MDL 2009 et LEC 2007 (le test SPEC6 ne figure pas car il n'a pas été traité selon les modèles de réponse à l'item). En abscisse figure le niveau de compétences des élèves sur le continuum θ , dont l'échelle a été standardisée à 0 avec un écart-type de 1 pour chacune des évaluations. En ordonnée, il s'agit de la valeur de l'information (c'est-à-dire de la précision) du test. Il ressort que l'évaluation PIRLS est très précise pour les élèves moyens, situés au centre du continuum, mais peu précise pour les élèves les plus faibles et les élèves les meilleurs. L'évaluation CEDRE MDL est plus précise

pour les élèves situés un peu au-dessous de la moyenne. L'évaluation LEC est, quant à elle, la plus précise des trois pour les élèves les plus faibles mais elle est très peu précise pour les élèves moyens. Or, les évaluations PIRLS et CEDRE MDL montrent une grande stabilité des résultats dans le temps, alors que l'évaluation LEC montre une aggravation des difficultés des élèves les plus faibles. Cette apparente contradiction pourrait donc être expliquée par le degré de précision de ces évaluations qui diffère selon le niveau de compétences étudié. Autrement dit, le niveau des élèves situés dans la moyenne est stable, ce que confirme PIRLS avec une grande précision, tandis que le niveau des élèves les plus faibles tend à baisser, ce qu'indique LEC avec une plus grande précision.

Information des tests



Champ : voir figure encadré 2 selon les évaluations.
Sources : Depp.

2. La population elle-même a changé en raison de la diminution des taux de redoublement. À niveau scolaire donné, les élèves sont moins âgés aujourd'hui qu'il y a dix ou vingt ans. Cependant, une mise en relation directe de ce phénomène avec l'évolution des performances des élèves serait hasardeuse : en guise d'illustration, dans l'enquête LEC, le taux d'élèves en retard a fortement diminué de 1987 à 1997, passant de 33 % à 19 %, tandis que le niveau de performances des élèves est resté très stable sur cette période ; de 1997 à 2007, le taux de retard a reculé de 19 % à 16 % tandis que les performances des plus faibles ont baissé.

dans des situations clairement disciplinaires. En revanche, le contenu de l'évaluation internationale PISA dépasse le cadre strictement scolaire et s'intéresse plus particulièrement à la capacité des élèves à utiliser leurs connaissances dans des situations de la vie quotidienne. C'est le concept de littératie qui est visé, c'est-à-dire « la capacité à utiliser et à réagir à des textes afin de réaliser ses objectifs, de développer ses connaissances et son potentiel et de participer à la société » [OCDE, 2010]. Le premier type d'évaluation permet d'évaluer l'atteinte des objectifs fixés par l'école et le second porte un regard externe sur des attendus plus généraux. Si ces deux approches sont complémentaires, elles peuvent conduire à des divergences d'appréciation concernant les tendances.

Au-delà des objectifs, les caractéristiques de ces évaluations sont sensiblement différentes. Le type et la longueur des textes, le format du questionnement (ouvert ou fermé de type question à choix multiples, QCM) sont des éléments essentiels d'une évaluation de la compréhension de l'écrit et peuvent conduire à des performances différentes. Ainsi, comme nous le verrons plus loin, les élèves français ont plus de difficultés face à un questionnement appelant une réponse écrite que face à des QCM.

Enfin, le degré de précision de ces évaluations n'est pas le même selon le niveau des élèves sur le continuum de performances. Pour simplifier, un test composé de questions faciles sera plus précis pour mesurer le niveau des élèves en difficulté, qu'un test composé de questions difficiles (de la même manière qu'une balance de salle de bains n'est pas adaptée pour peser une fourmi ou un éléphant). En l'occurrence, une analyse de la précision des différents tests de compréhension de l'écrit en fin de primaire fait ressortir que les évaluations PIRLS et CEDRE MDL sont très précises pour les élèves médians, c'est-à-dire situés au milieu du continuum de l'échelle de compétences (*encadré 3*). Ce n'est pas le cas du test de lecture-compréhension de l'évaluation LEC qui est plus précis pour les élèves situés dans les bas niveaux de compétences. Ces propriétés différentes sont susceptibles d'expliquer certaines différences d'interprétation des tendances selon les évaluations.

Stabilité pour les élèves moyens ...

Le niveau de performances des élèves proches de la médiane n'a pas évolué depuis une dizaine d'années, que ce soit à l'école ou au collège³.

À l'école primaire, l'évaluation CEDRE MDL affiche une grande stabilité des résultats, stabilité qui concerne d'ailleurs tous les élèves en fin de CM2, quel que soit leur niveau dans la distribution des résultats. Comme on l'a vu, cette évaluation s'avère plus précise pour les élèves médians, ce qui permet de conforter la stabilité du niveau de ces élèves en maîtrise de la langue. L'évaluation internationale PIRLS révèle aussi une très grande stabilité des résultats des élèves de CM1 entre 2001 et 2006, bien que la place de la France soit en deçà de la moyenne des pays de l'Union européenne ayant participé à PIRLS. Comme dans CEDRE MDL, cette stabilité concerne l'ensemble de la distribution des élèves, des élèves les plus faibles aux élèves les plus performants. Le constat d'une stabilité du niveau des élèves situés dans la médiane est cependant renforcé, étant donné le degré de précision très élevé de l'évaluation PIRLS pour ces élèves.

Au collège, en fin de 3^e, l'évaluation CEDRE CG montre également une grande stabilité des résultats des élèves situés dans les niveaux de performances médians entre 2003 et 2009. La part des élèves situés dans les niveaux les plus élevés a cependant diminué de manière statistiquement significative, passant de 10,0 % à 7,1 %. L'évaluation CEDRE CG pointe aussi une augmentation du nombre d'élèves dans les niveaux les plus faibles (*cf. infra*). Les résultats

3. Il s'agit d'une comparaison des résultats des élèves proches de la médiane, d'une vague d'enquête à l'autre, la médiane étant recalculée lors de chaque vague d'enquête.

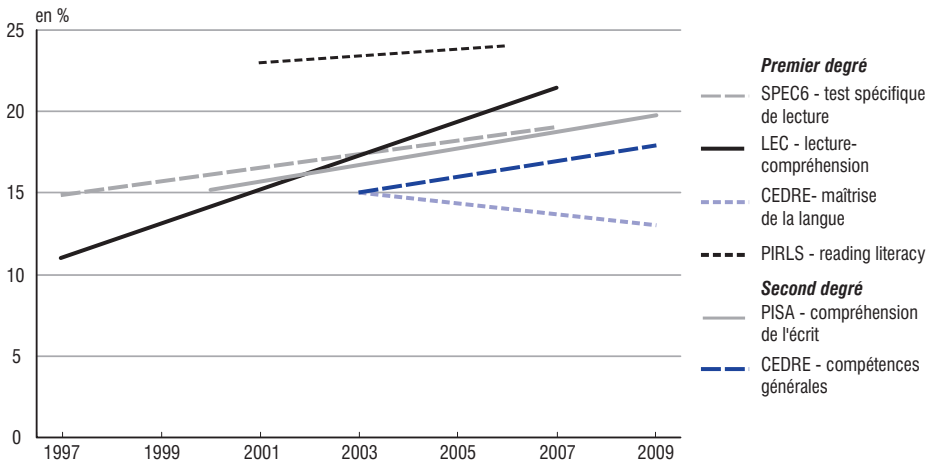
de la France à l'évaluation internationale PISA s'agissant des élèves de 15 ans montrent que les élèves situés dans les niveaux de performances intermédiaires affichent des résultats très comparables en 2000 et 2009. Contrairement à CEDRE CG, cette stabilité concerne également les élèves des niveaux supérieurs. Globalement, l'évolution du score moyen des élèves n'est pas significative dans PISA, et la France demeure dans la moyenne des pays de l'OCDE. En revanche, PISA indique une baisse significative des résultats pour les élèves les plus faibles (cf. *infra*).

... mais la part des élèves en difficulté face à l'écrit a augmenté

Cependant, depuis une dizaine d'années, la proportion d'élèves pouvant être considérés en difficulté face à l'écrit a augmenté, et ce phénomène concerne aujourd'hui près d'un élève sur cinq (*figure 1*).

En fin d'école, le pourcentage d'élèves faibles en compréhension de l'écrit a presque doublé de 1997 à 2007, passant de 11,0 % à 21,4 %, selon les résultats de l'évaluation LEC. En début de 6^e, la part des élèves en difficulté de lecture augmente également, mais de manière moins marquée d'après l'enquête SPEC6 (« test spécifique de lecture à l'entrée en 6^e ») : de 14,9 % en 1997 à 19,0 % en 2007. Ces deux évaluations interrogent cependant des dimensions différentes de la lecture, qui ne sont pas toutes sujettes à variation, comme nous le verrons plus loin avec le détail des résultats de l'enquête SPEC6. Ces deux dernières enquêtes sont, rappelons-le, très adaptées pour mesurer les compétences des élèves les plus faibles. De leur côté, les évaluations comparatives portant sur l'école primaire, PIRLS et CEDRE MDL, montrent une stabilité des résultats, même pour les élèves les plus en difficulté. Cependant, ces deux évaluations sont peu adaptées pour évaluer les élèves situés dans les faibles niveaux de performances, ce qui est susceptible d'expliquer l'apparente contradiction de résultats obtenus en regard de ceux des évaluations LEC et SPEC6.

1. Évolution de la part des élèves en difficulté



Champ : voir *figure encadré 2* selon les évaluations.

Lecture : selon l'évaluation PISA, en 2000, environ 15 % des élèves de 15 ans se situaient dans les groupes de performances inférieurs au niveau 2.

Note : courbes en trait plein = évolutions significatives au seuil de 1 % ; courbes en pointillés longs = évolutions significatives au seuil de 5 % ; courbes en pointillés courts = évolutions non significatives au seuil de 10 %. « Élèves en difficultés » = SPEC6 : élèves de début de 6^e en difficulté de lecture ; LEC : élèves de CM2 dans les faibles niveaux (référence : 1^{er} décile en 1987) ; CEDRE MDL : élèves de fin de CM2 dans les groupes inférieurs au niveau 2 ; PIRLS : élèves de CM1 sous l'« Intermediate Benchmark » ; PISA : élèves de 15 ans dans les groupes inférieurs au niveau 2 ; CEDRE CG = élèves de fin de 3^e générale dans les groupes inférieurs au niveau 2.

Source : Depp.

Dans le second degré, les conclusions des évaluations PISA et CEDRE CG sont convergentes : la part des élèves les plus en difficulté a augmenté de manière significative : de 15,2 % à 19,8 % pour la première, et de 15,0 % à 17,9 % pour la seconde. Les contenus de ces évaluations sont cependant différents, PISA étant orienté vers la littérature alors que CEDRE CG repose sur les programmes scolaires. Mais concernant les élèves les plus faibles repérés par l'une ou l'autre de ces évaluations, il s'agit d'élèves ayant des compétences très limitées dans le traitement de l'information écrite, compétences qui seront un obstacle à la poursuite d'études mais aussi dans l'adaptation à la vie quotidienne. En outre, ce sont chez les plus faibles d'entre eux que la baisse est la plus importante : les groupes les moins performants à PISA représentaient 4,2 % des élèves en 2000, ils sont 7,9 % en 2009.

Ces différentes évaluations montrent également des taux de non-réponse des élèves importants et en hausse lorsqu'il s'agit de produire une réponse rédigée de manière construite. Ce phénomène est bien mis en exergue par les évaluations internationales : le taux de non-réponse observé aux questions ouvertes de PISA 2009 est de 16 % en France contre 11 % dans l'ensemble des pays de l'OCDE. Plutôt que de risquer de répondre faux, les élèves français préfèrent s'abstenir, un comportement qui a été interprété au regard du statut de l'erreur dans le système éducatif français. Elle serait plus souvent qu'ailleurs envisagée comme une faute et non comme faisant naturellement partie du processus d'apprentissage.

Des automatismes de lecture au même niveau mais des difficultés langagières plus marquées

S'il semble avéré que le niveau global de compréhension de l'écrit est en baisse pour les élèves les plus faibles, ces difficultés méritent d'être précisées. Les évaluations LEC et SPEC6 apportent des éclairages complémentaires sur les dimensions concernées par l'augmentation des difficultés.

Il ressort clairement que l'orthographe est moins bien maîtrisée aujourd'hui qu'il y a vingt ans. La même dictée a été proposée aux élèves de CM2 en 1987 et en 2007, à partir d'un texte d'une dizaine de lignes. Le nombre moyen d'erreurs est passé de 10,7 en 1987 à 14,7 en 2007 (figure 2). Le pourcentage d'élèves qui faisaient plus de 15 erreurs était de 26 % en 1987, il est aujourd'hui de 46 %. Ce sont principalement les erreurs grammaticales qui ont augmenté : de 7 en moyenne en 1987 à 11 en 2007. Par exemple, 87 % des élèves conjuguait correctement « tombait » dans la phrase « Le soir tombait. » ; aujourd'hui, ils ne sont plus que 63 %. En revanche, sur des conjugaisons difficiles pour les élèves de CM2, comme l'accord avec l'auxiliaire « avoir », le pourcentage de réussite n'évo-lue pas : environ 30 % des élèves, que ce soit en 1987 ou en 2007, écrivent correctement « vus » dans la phrase « Elle les a peut-être vus ! ».

2. Orthographe et grammaire en CM2

	1987	2007
Orthographe - dictée, nombre moyens d'erreurs		
Erreurs lexicales	2,1	2,6
Erreurs grammaticales	7,1	10,8
Erreurs de ponctuation	1,1	0,9
Autres (oubli de mots, ...)	0,3	0,4
Ensemble	10,7	14,7
Exercices de grammaire, réussites moyennes (en %)		
Transformations pronominales	60,6	49,5
Conjugaisons	51,9	37,4

Champ : France métropolitaine, secteur public élèves de fin de CM2.

Lecture : en 1987, les élèves de CM2 de l'échantillon ont fait en moyenne 10,7 erreurs à la dictée ; parmi les transformations pronominales proposées, ils en ont réussi 60,6 %.

Source : Depp, évaluation LEC « Lire, écrire, compter » 1987-2007.

L'évaluation SPEC6, qui portait spécifiquement sur les difficultés de lecture à l'entrée en 6^e, fait ressortir des contrastes intéressants en termes d'évolution temporelle (figure 3). Les mécanismes de base, c'est-à-dire les automatismes impliqués dans l'identification des mots, restent stables entre 1997 et 2007 tandis que les compétences langagières se dégradent. En effet, parmi les mécanismes de base de la lecture, le degré de connaissances phonologiques (exercices de prononciation⁴), morphologiques (reconnaître des mots de la même famille), les connaissances lexicales sur mots fréquents (reconnaître les mots qui existent parmi une liste de vrais mots et de non-mots), ainsi que la rapidité des élèves à identifier des mots écrits sont globalement constants entre 1997 et 2007 : les pourcentages d'élèves éprouvant des difficultés dans ces dimensions restent stables sur dix ans. En revanche, les élèves en début de 6^e en 2007 ont une maîtrise de la langue française moins bonne qu'en 1997 : un niveau de vocabulaire plus pauvre, une moins grande maîtrise orthographique et un plus faible niveau de compréhension d'énoncés écrits. Or, ce sont ces dimensions liées au langage qui sont les plus corrélées avec le niveau des élèves en français⁵. L'augmentation du nombre d'élèves faibles en compréhension de l'écrit est donc à rapprocher de l'appauvrissement de ces compétences langagières, plutôt que de problèmes « fonctionnels » de lecture.

3. Proportion d'élèves de début de 6^e en difficulté de lecture, par type de compétence évaluée

	1997	2007	Significativité
Rapidité de traitement	22,5	24,1	n.s.
Connaissances lexicales sur mots rares	24,2	30,7	***
Connaissances lexicales sur mots fréquents	19,6	19,9	n.s.
Connaissances phonologiques	21,7	21,6	n.s.
Connaissances morphologiques	17,0	17,6	n.s.
Connaissances orthographiques	23,1	33,6	***
Compréhension d'énoncés	28,4	33,6	**
Global - difficulté de lecture	14,9	19,0	**

Champ : France métropolitaine, élèves de début de 6^e.

Lecture : parmi les élèves de début de 6^e en 1997, 22,5 % ont des difficultés de rapidité de traitement des mots écrits contre 24,1 % en 2007.

Note : * évolution significative au seuil de 10 %, ** évolution significative au seuil de 5 %, *** évolution significative au seuil de 1 %, n.s. évolution non significative.

Source : Depp, SPEC6 « Test spécifique des difficultés de lecture en début de 6^e » 1997-2007.

Les filles creusent l'écart en compréhension de l'écrit : un phénomène général dans les pays de l'OCDE

Les filles ont de meilleurs résultats que les garçons en lecture et en compréhension de l'écrit, et cela dès l'enseignement primaire. Ce constat national s'élargit au niveau international : dans tous les pays participants à l'évaluation PISA 2009, les filles obtiennent un score moyen en compréhension de l'écrit significativement supérieur à celui des garçons.

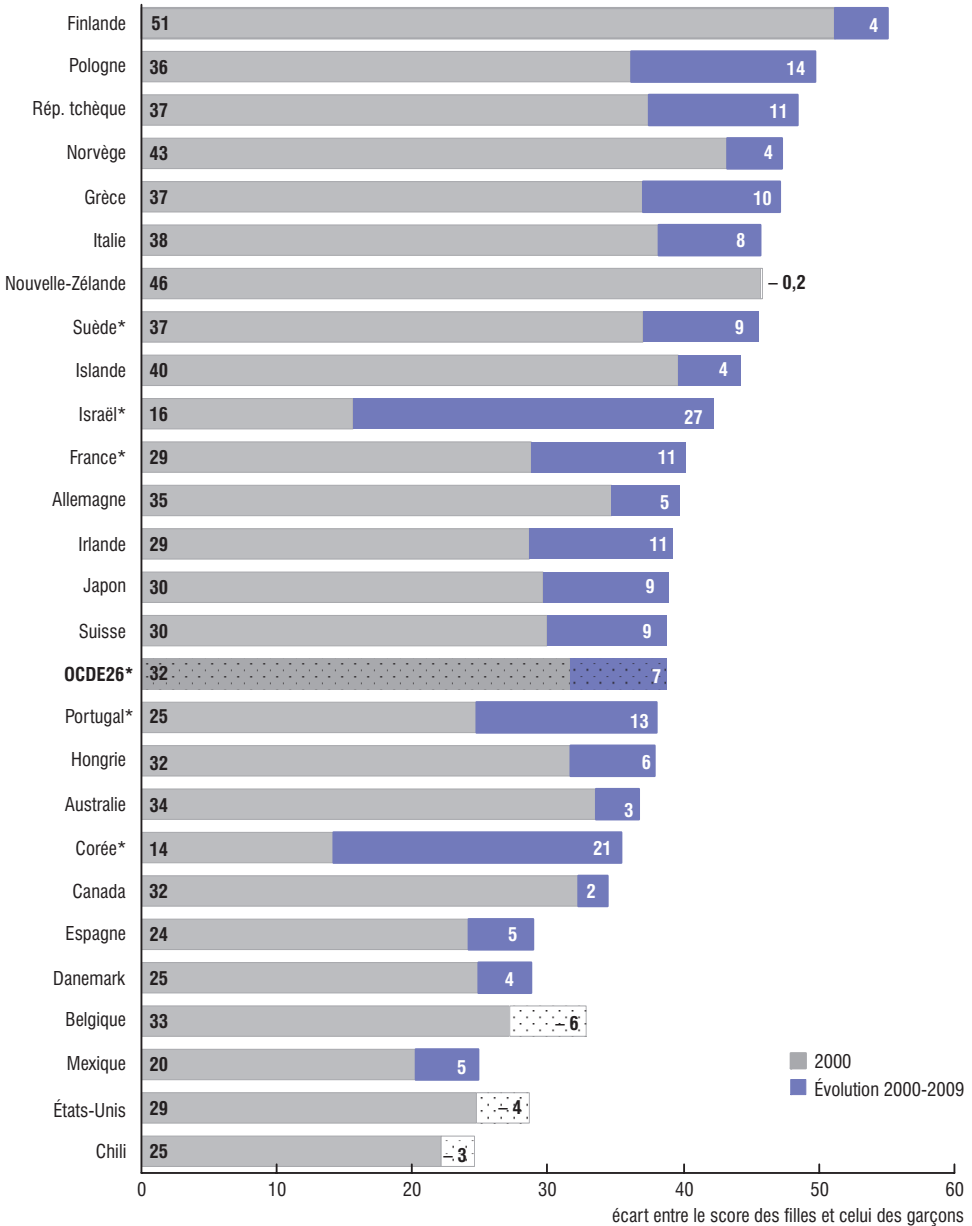
En outre, les comparaisons temporelles montrent que le différentiel filles-garçons tend à augmenter. En moyenne dans les 26 pays de l'OCDE ayant participé à l'opération PISA en 2000 et en 2009, la différence entre le score moyen des filles et celui des garçons augmente de manière significative de 7 points⁶, passant de 32 points à 39 points en faveur des filles sur cette période (figure 4). L'écart entre filles et garçons semble se creuser dans la majorité des pays de l'OCDE, mais surtout en Israël, en Corée, au Portugal, en France et en Suède où cette évolution est statistiquement significative. L'écart entre les filles et les garçons reste considérable, même dans les pays qui tendent vers une plus grande homogénéité des résultats (Belgique, États-Unis, Chili, Nouvelle-Zélande).

4. Prononce-t-on de la même manière : « danger » et « danjer », « doctour » et « docteur », « maizon » et « maison », « image » et « imache », etc. ?

5. Le niveau des élèves en français a été mesuré grâce à l'évaluation diagnostique de début de 6^e en 2007.

6. La moyenne des scores au niveau international a été fixée à 500 et l'écart-type à 100 en 2000 (encadré 3).

4. Différence entre 2000 et 2009 des scores moyens des garçons et des filles en compréhension de l'écrit à PISA



Champ : élèves de 15 ans.

Lecture : en France, les filles ont obtenu un score moyen supérieur de 29 points à celui des garçons en 2000. L'évolution de la différence entre le score moyen des filles et celui des garçons entre 2000 et 2009 est de 11 points. Cette différence est statistiquement significative.

Note : la figure représente les 26 pays de l'OCDE ayant participé à l'évaluation PISA en 2000 et en 2009. Les pays où l'évolution de la différence entre le score moyen des filles et celui des garçons est statistiquement significative, sont indiqués avec un astérisque *. La moyenne des scores au niveau international a été fixée à 500 et l'écart-type à 100 en 2000.

Sources : Depp ; OCDE, évaluations PISA 2000 et PISA 2009.

En France, la différence entre les filles et les garçons augmente de 11 points entre les deux cycles d'évaluation, l'écart étant désormais de 40 points. À titre d'illustration, selon l'évaluation PISA, le progrès fait par les élèves lors d'une année d'études correspond en moyenne à 39 points sur l'échelle de compréhension de l'écrit⁷. En France, l'écart entre les filles et les garçons est comparable à celui que l'on observe entre les élèves de 15 ans inscrits en classe de 2^{nde} générale et technologique et les élèves de 3^e de 14 ans, à l'heure⁸. Les filles devancent ainsi les garçons comme si elles avaient bénéficié d'une année d'enseignement supplémentaire.

Même si l'augmentation de la part des élèves en difficulté concerne aussi les filles, les garçons sont toujours plus nombreux à se retrouver aux plus bas niveaux de compétences en compréhension de l'écrit : 11,5 % d'entre eux sont dans le niveau de performances le moins élevé de PISA en 2009, contre 4,6 % des filles.

Le statut économique, social et culturel (SESC) a aujourd'hui un pouvoir explicatif sur les variations de performances plus élevé en France que dans les autres pays de l'OCDE

Au-delà du niveau de performances, la comparaison des disparités sociales, observées à différentes périodes, renseigne sur l'évolution du degré d'équité du système éducatif. Ainsi, l'enquête PISA permet non seulement de situer les élèves français sur une échelle internationale mais elle offre également des éléments pour analyser l'équité du système éducatif français par rapport à ceux des autres pays de l'OCDE, à travers des indicateurs mesurant le lien entre les performances des élèves et leur milieu social. En France, ce lien se renforce depuis une dizaine d'années, traduisant un accroissement des disparités sociales.

L'écart de scores moyens entre les élèves socialement défavorisés et les élèves socialement favorisés augmente en France entre 2000 et 2009 et se situe désormais à 96 points, significativement au-dessus de la moyenne de l'OCDE qui n'a pas évolué. Plus précisément, les performances en compréhension de l'écrit sont comparées pour deux groupes d'élèves : le quart des élèves aux indices socioéconomiques de statut professionnel⁹ (HISEI) les plus élevés (dernier quartile), et le quart des élèves aux indices HISEI les plus faibles (premier quartile). En 2000, l'écart entre les scores moyens de ces deux groupes était de 83 points en France, comparable à celui de 80 points qui concerne l'ensemble de 26 pays de l'OCDE ayant participé à PISA 2000 et 2009 (figure 5). Il faut noter que l'accroissement de cet écart s'explique principalement par la baisse des résultats des élèves les moins favorisés en termes de statut professionnel des parents : le score moyen des élèves appartenant au premier quartile sur l'échelle de l'indice a baissé de 14 points entre les deux cycles d'évaluation tandis que celui des élèves du dernier quartile est resté stable.

Le lien entre origine sociale et performances est donc renforcé. En France, le statut économique, social et culturel (SESC) des parents explique une plus grande part de la variation des scores des élèves qu'en moyenne dans l'ensemble des pays de l'OCDE. Cet indice est significativement plus corrélé aux performances des élèves que dans l'ensemble des pays de l'OCDE. Ainsi, l'écart de score associé à la variation d'une unité de l'indice SESC est de 51 points en France contre 38 points en moyenne dans l'OCDE.

7. Cette référence a été établie en calculant la différence de score moyen entre des élèves de deux niveaux scolaires (« grades »).

8. C'est-à-dire les élèves qui ont l'âge de ceux n'ayant jamais redoublé ni sauté une classe.

9. Le questionnaire de contexte de l'enquête PISA permet de recueillir des informations très variées concernant, entre autres, le statut professionnel, le niveau de formation des parents d'élèves et les ressources de diverse nature (matérielles, culturelles, éducatives...) que possède le foyer. Ces renseignements, qui se basent sur les déclarations d'élèves, ont été intégrés dans différents indices. Par exemple, l'indice HISEI est l'indice socioéconomique le plus élevé parmi les indices des professions des deux parents, codés selon la nomenclature internationale de métiers CITP (classification internationale type des professions) puis transformés en indice continu.

5. Relation entre les performances des élèves en compréhension de l'écrit à PISA et les caractéristiques du milieu social

	2000	2009	Évolution
Score moyen des élèves appartenant au premier quartile sur l'échelle de l'indice HISEI¹			
France	470	455	- 14
OCDE	462	463	1
Score moyen des élèves appartenant au dernier quartile sur l'échelle de l'indice HISEI¹			
France	552	551	- 2
OCDE	543	542	0
Écart entre le score moyen du premier quartile et celui du dernier quartile sur l'échelle de l'indice HISEI¹			
France	83	96	13
OCDE	80	80	1
Écart de score associé à la variation d'une unité de l'indice SESC²			
France	44	51	7
OCDE	39	38	- 1

1. L'indice socioéconomique de statut professionnel le plus élevé (HISEI) est l'indice le plus élevé parmi les indices de professions des deux parents sur l'échelle internationale de statut économique (SEI).

2. L'indice de statut économique, social et culturel (SESC) intègre également des éléments sur le niveau d'éducation des parents et les possessions du foyer. Champ : élèves de 15 ans.

Lecture : les élèves français se situant dans le premier quartile sur l'échelle de l'indice HISEI ont obtenu un score moyen de 470 points en 2000 et de 455 points en 2009. En 2009, la variation d'une unité de l'indice SESC conduit à un écart de score de 51 points en France, ce qui est significativement au-dessus de la moyenne de l'OCDE (38 points). (Les différences significatives entre la France et la moyenne des 26 pays de l'OCDE sont indiquées **en gras**).

Note : les indices sont construits à partir des réponses des élèves à des questions portant sur les caractéristiques de leur environnement social et familial. La moyenne des scores au niveau international a été fixée à 500 et l'écart-type à 100 en 2000.

Source : Depp ; OCDE, évaluations PISA 2000 et PISA 2009.

Les écarts de performances concernent particulièrement les collèges en zones d'éducation prioritaire

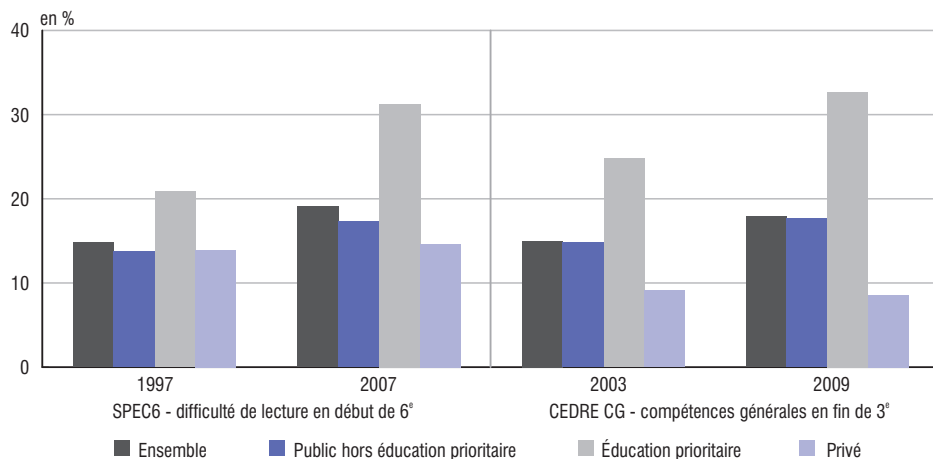
Si le pourcentage d'élèves en difficulté de lecture à l'entrée en 6^e a augmenté de 1997 à 2007 (de 14,9 % à 19,0 %, cf. *supra*), une analyse plus détaillée montre que cet accroissement touche particulièrement les collèges en zones d'éducation prioritaire (ZEP) où le pourcentage d'élèves en difficulté de lecture a augmenté de 20,9 % à 31,3 % (figure 6). Pour les collèges publics hors éducation prioritaire et les collèges privés, si la proportion globale d'élèves en difficulté n'a pas varié de manière significative, les compétences langagières se sont néanmoins dégradées, mais pas les mécanismes de base de la lecture. En revanche, dans les zones d'éducation prioritaire, toutes les dimensions de la lecture sont affectées, y compris les mécanismes de base, ce qui explique l'augmentation significative de la proportion globale d'élèves en difficulté.

En fin de collège, la comparaison des résultats de l'évaluation CEDRE CG entre 2003 et 2009 montre également que la situation des collèges appartenant au secteur de l'éducation prioritaire est particulièrement interrogée : le score moyen des élèves qui y sont inscrits a baissé de 12 points sur la période observée, contre une baisse de 5 points dans l'ensemble de la population des élèves de 3^e. La proportion d'élèves dans les niveaux de performances les plus faibles (inférieurs au niveau 2) dans le secteur de l'éducation prioritaire a augmenté de 7,7 points entre les deux cycles d'évaluation, passant de 24,9 % en 2003 à 32,6 % en 2009. En comparaison, la part des élèves faibles n'a augmenté que de 2,9 points dans les collèges du secteur public hors éducation prioritaire et elle est restée stable dans le secteur privé. En 2009, près d'un tiers des élèves de fin de 3^e sont ainsi en difficulté dans le secteur de l'éducation prioritaire, contre 17,7 % dans les collèges publics hors éducation prioritaire et 8,5 % dans les établissements privés.

Ces résultats affinent le diagnostic en précisant les zones où se concentre la baisse des performances face à l'écrit. Toutefois, ils ne permettent pas d'évaluer la politique des ZEP car on ne sait pas quelle aurait été l'évolution des compétences des élèves dans ces collèges sans

sa mise en place. Par ailleurs, si le champ des collèges relevant du secteur de l'éducation prioritaire est resté relativement stable sur la période (93 % des établissements qui étaient en ZEP en 2003 le sont également en 2009 ; 98 % des établissements qui sont en ZEP en 2009 l'étaient déjà en 2003), les caractéristiques des élèves qui vont dans ces collèges ont pu évoluer. On note toutefois que la composition sociale des collèges de ZEP ne semble pas avoir changé de manière significative entre 2003 et 2009, si l'on observe la répartition des élèves selon les grands groupes de professions et catégories sociales (PCS) de leurs parents. ■

6. Évolution des difficultés selon le secteur et les zones d'éducation



Champ : voir figure encadré 2 selon les évaluations.

Lecture : le pourcentage d'élèves de début de 6^e en ZEP en difficulté de lecture passe de 20,9 % à 31,3 % entre 1997 et 2007.

Note : évolutions significatives au seuil de 5 % pour l'ensemble, 1 % pour l'éducation prioritaire, 10 % pour le public hors éducation prioritaire pour CEDRE CG seulement ; non significatives pour le public hors ZEP pour SPEC6 ou le privé.

Source : Depp, SPEC6 « Test spécifique des difficultés de lecture en début de 6^e » 1997-2007, CEDRE CG « Cycle des évaluations disciplinaires réalisées sur échantillons - Compétences générales » 2003-2009.

Pour en savoir plus

Baudelot C., Establet R., « Le niveau monte – Réfutation d’une vieille idée concernant la prétendue décadence de nos écoles », Paris : Seuil, 1989.

Bourny G., Bessonneau P., Daussin J.-M., Kesksaik S., « L’évolution des compétences générales des élèves en fin de collège de 2003 à 2009 », *Note d’information* n° 10.22, Depp, 2010.

Colmant M., Dos Santos S., « Évolution des performances en lecture des élèves de CM1. Résultats de l’étude internationale PIRLS », *Note d’information* n° 08.14, Depp, 2008.

Fumel S., Kesksaik S., Girard J., « L’évolution des acquis des élèves de 15 ans en compréhension de l’écrit. Premiers résultats de l’évaluation internationale PISA 2009. », *Note d’information* n° 10.24, Depp, 2010.

Rocher T., « Lire, écrire, compter : les performances des élèves de CM2 à vingt ans d’intervalle 1987-2007 », *Note d’information* n° 08.38, Depp, 2008.

Salines M., Vrignaud P., « Apprécier et certifier les acquis des élèves en fin de collège : diplôme et évaluations-bilans », *Rapport établi à la demande du Haut Conseil de l’Évaluation de l’École*, n° 2, 2001.

Thélot C., « Que sait-on des connaissances des élèves ? », *Rapport à Monsieur le Ministre d’État, Ministre de l’Éducation Nationale et de la Culture*, Depp, 1992.

CEDRE : site compagnon de la collection Depp-CNDP http://www.cndp.fr/collection/collection-detail/Evaluations-eleves-151282.html?tx_cndpcollection_pi1

PIRLS : site <http://timss.bc.edu>

PISA : site <http://pisa.oecd.org>

« PISA 2009 Results : What Students Know and Can Do : Student Performance in Reading, Mathematics and Science (Volume I) », OCDE, 2010.
