note d'
information
12.21
DÉCEMBRE

L'étude internationale PIRLS 2011 vise à mesurer les performances en lecture des élèves à la fin de leur quatrième année de scolarité obligatoire (CM1 pour la France). La France, avec un score de 520 points, se situe au delà de la moyenne internationale (500 points) mais en deçà de la moyenne européenne (534 points) et ce, quel que soit la compétence ou le type de texte considéré. Comparées aux résultats obtenus lors de PIRLS 2001 et PIRLS 2006, bien qu'en légère baisse, les performances globales de la France sont statistiquement stables.
Sur la période de dix ans (2001-2011), on observe toutefois une baisse significative des performances sur la compréhension des textes informatifs (- 13 points), de même que sur les compétences les plus complexes (- 11 points).
Par ailleurs, les élèves français sont toujours les plus nombreux à s'abstenir de répondre lorsque les réponses doivent être rédigées. Ils sont aussi les plus nombreux à ne pas terminer les épreuves.

PIRLS 2011 - Étude internationale sur la lecture des élèves au CM1

Évolution des performances à dix ans

En mai 2011, cinquante-quatre pays ou provinces[1] ont participé à l'enquête internationale PIRLS *(Progress in International Reading Literacy Study)* organisée par l'IEA[2] (Association internationale pour l'évaluation des compétences scolaires) en vue d'évaluer les performances en lecture des élèves à la fin de leur quatrième année de scolarité obligatoire. À cette étape de leur scolarité, les élèves approfondissent leurs compétences en lecture et lisent dès lors pour apprendre. Quelle que soit la matière enseignée, la lecture est essentielle à la réussite à l'école, mais elle joue un rôle important dans la réalisation de soi des élèves et permet d'augmenter leurs capacités de réflexion et de créativité. Cette étude a déjà été menée en 2001 et en 2006 ; 23 pays de 2001 et 31 pays de 2006 sont toujours présents en 2011. Ainsi, cette troisième prise d'information de la série, en plus d'apporter un constat sur les compétences en lecture des élèves français en 2011 par rapport à celles de leurs camarades des autres pays, donne lieu également à une comparaison dans le temps.

Une évaluation des élèves de CM1

Dans PIRLS, la population étudiée ne correspond pas à une génération d'élèves d'un âge défini, comme c'est le cas dans le programme PISA (Programme international pour le suivi des acquis des élèves) de l'OCDE. Le but de l'étude est de mesurer les performances de l'ensemble des élèves présents à un niveau déterminé de la scolarité, quels que soient leur âge, leur parcours et l'organisation du système éducatif (proportion d'élèves en retard, etc.).
La classe cible de l'enquête est précisément définie comme *« la classe qui représente quatre années d'enseignement. Le point de départ de ces quatre années, la première année du premier niveau de la Classification Internationale Type de l'Éducation (CITE) validée par l'UNESCO, c'est la classe où se fait le début des apprentissages systématiques de la lecture, de l'écriture et des mathématiques. »* Dans la plupart des pays, PIRLS évalue donc les élèves de quatrième année d'école obligatoire, en France les élèves de CM1. Toutefois, la moyenne d'âge des élèves ne doit pas être inférieure

1. Trois provinces canadiennes (Alberta, Ontario et Québec) ; l'État de Floride aux USA ; la Communauté autonome d'Andalousie en Espagne ; la province chinoise de Hong Kong ; l'Angleterre et l'Irlande du Nord pour le Royaume-Uni, participent à l'étude pour elles-mêmes. Ceci se justifie par le fait que les systèmes éducatifs de ces entités sont distincts.
2. L'IEA *(International Association for the Evaluation of Educational Achievement)* est une association scientifique indépendante à but non lucratif, dont la mission principale est de mettre en œuvre des études comparatives à échelle internationale sur l'évaluation et la réussite scolaire des élèves. Depuis sa création en 1958, cette association pionnière a conduit plus d'une vingtaine d'études de portée internationale dont PIRLS.

Liberté • Égalité • Fraternité
RÉPUBLIQUE FRANÇAISE
ministère éducation nationale

à 9,5 ans au moment de la passation des épreuves. C'est la raison pour laquelle l'Angleterre, Malte, la Nouvelle-Zélande, et Trinité-et-Tobago ont évalué les élèves dans leur cinquième année de scolarité.

Au moment de la passation, la moyenne d'âge internationale est de 10,2 ans ; les élèves français sont légèrement plus jeunes, 10,0 ans, comme en 2006. Les élèves français sont parmi les plus jeunes de l'étude, les plus jeunes étant les Italiens et les Norvégiens (9,7 ans) et les plus âgés les Danois et les Roumains (10,9 ans). Les politiques de chaque pays en matière d'âge de scolarisation, de promotion et de rétention de classe, conformément aux contextes économiques, sociaux et culturels ne permettent pas d'établir un lien direct entre résultats et âge moyen des élèves. Certains pays, par exemple, adoptent une politique stricte d'admission à l'école primaire sur la seule base de l'âge et utilisent ensuite le principe de la promotion automatique ; d'autres permettent aux parents une certaine latitude dans le choix du moment où leur enfant commence l'école, et peuvent autoriser, voire obliger les élèves à redoubler ou, à l'inverse, à avancer d'une classe.

Une évaluation qui vise les compétences en lecture et les objectifs donnés à la lecture

PIRLS est conçu pour évaluer la « compétence en lecture ». Celle-ci est définie comme : *« L'aptitude à comprendre et à utiliser les formes du langage écrit que requiert la société ou qui sont importants pour l'individu. Les jeunes lecteurs peuvent construire du sens à partir de textes très variés. Ils lisent pour apprendre, pour s'intégrer dans une société où la lecture joue un rôle essentiel et pour leur plaisir. »*

Le lecteur mobilise toute une gamme de capacités, de stratégies cognitives et métacognitives ainsi que ses connaissances préalables des objets abordés.

L'étude pose comme hypothèse que la compétence en lecture est directement liée aux raisons qui incitent le lecteur à lire. C'est pourquoi l'épreuve dans son ensemble a été bâtie sur le croisement de deux aspects : les compétences en lecture et les objectifs du lecteur.

Les compétences en lecture

Quatre compétences sont définies :

– *prélever des informations explicites (Prélever)* (22 % du barème) : repérer les informations directement liées à l'objectif de la lecture ; chercher des idées précises ; chercher la définition de mots ou d'expressions ; repérer le contexte de l'histoire (époque, lieu) ; trouver l'idée principale (si elle est exprimée explicitement) ;

– *faire des inférences directes (Inférer)* (28 % du barème) : déduire que tel événement a entraîné tel autre ; déduire l'élément principal d'une série d'arguments ; déterminer le référent d'un pronom ; repérer les généralisations présentées dans le texte ; décrire la relation entre deux personnages ;

– *interpréter et assimiler idées et informations (Interpréter)* (37 % du barème) : déduire le message global ou le thème d'un texte ; envisager une alternative aux actions des personnages ; comparer des informations du texte ; saisir l'atmosphère ou le ton du récit ; trouver une application concrète aux informations contenues dans le texte ;

– *examiner et évaluer le contenu, la langue et les éléments textuels (Apprécier)* (13 % du barème) : évaluer la probabilité que les événements décrits se passent réellement ; décrire la manière dont l'auteur a amené la chute ; juger de l'intégralité ou de la clarté des informations fournies dans le texte ; décrire comment le choix des adjectifs modifie le sens.

Il est à noter que ces quatre compétences sont hiérarchisées dans leur degré de complexité et que chacun des dix textes proposés aux élèves est couvert par une combinaison de questions qui visent à évaluer l'atteinte de chacune d'entre elles.

Les objectifs donnés à la lecture

La longueur des textes utilisés dans PIRLS est en moyenne de 850 mots, elle varie de 600 à 1 000 mots selon les cas. Les textes sont présentés en général sur trois ou quatre pages contenant également des illustrations ou des documents. Pour chaque texte, entre 13 et 16 questions sont posées aux élèves. En 2011, PIRLS contient 135 questions.

Deux des objectifs les plus fréquents donnés aux élèves pour leur lecture à l'école et ailleurs structurent également l'enquête PIRLS :

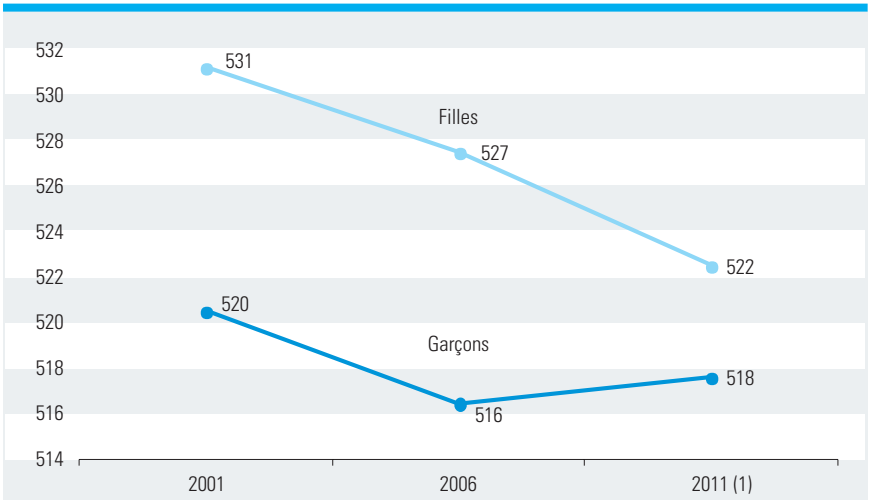
– *lire pour accéder aux textes littéraires* (cinq textes, 52 % du barème), la forme des textes utilisés principalement dans ce contexte est la fiction narrative. Le lecteur aborde le texte pour s'investir dans des événements, des situations, des actions, des conséquences, des personnages, une atmosphère, des sentiments, des idées imagées et pour apprécier la langue en soi. Les textes narratifs proposés dans PIRLS sont des histoires ou des épisodes complets appuyés par des illustrations. Ces cinq textes proposent des histoires contemporaines ou traditionnelles et couvrent une variété de contextes, on y rencontre le plus souvent deux personnages principaux et une intrigue avec un ou deux évènements centraux. Les styles utilisés sont variés et on y trouve des éléments narratifs tels que l'utilisation de la première personne, l'humour, les dialogues, voire du langage figuratif. Les thèmes des textes sont eux aussi divers ;

– *lire pour acquérir et utiliser des informations* (cinq textes, 48 % du barème), les textes informatifs prennent de nombreuses formes ; ils peuvent présenter l'information suivant ou non une logique chronologique. Au moyen de ces textes, le lecteur s'informe sur des aspects de l'univers réel pour comprendre le monde tel qu'il est, mais il peut également aller plus loin et appliquer les informations obtenues à un raisonnement ou à une action. Les cinq textes sont structurés par séquence ou par sujet, ils sont présentés avec, ou organisés autour de schémas, de cartes, d'illustrations, de photos, d'encadrés, de listes, de tableaux. Les textes sont proposés par les pays participants puis retenus après un long processus de sélection ; ils ont une syntaxe et un lexique bien adaptés à l'âge des élèves et correspondent aux prescriptions des programmes du cycle 3 de l'école élémentaire.

Résultats généraux

Sur l'ensemble de l'évaluation, la France se situe de manière statistiquement significative

GRAPHIQUE 1 – Évolution des scores des filles et des garçons français



Lecture : en France, les filles avaient un score moyen de 531 en 2001. Elles obtiennent un score de 522 en 2011.
(1) La différence filles/garçons n'est pas significative en 2011.

Source : MEN-MESR-DEPP

au-dessus du point central de l'échelle PIRLS : les élèves français obtiennent un score moyen de 520, soit un cinquième d'écart-type au dessus de la moyenne internationale fixée à 500. Le score moyen de la France est différent de manière statistiquement significative de celui de chacun des pays de l'étude, excepté la Pologne (525).

Quatre pays se partagent les meilleurs résultats : Hong Kong, avec un score de 571, la Fédération de Russie (568), la Finlande (568) et Singapour (567). Leurs scores moyens ne sont pas différents d'un point de vue statistique. Le score moyen le plus faible est obtenu par le Maroc (310).

32 pays participants obtiennent des scores moyens globaux significativement supérieurs au point central de l'échelle ; 11 pays ont même des scores moyens supérieurs à 550 (plus d'un demi-écart-type au-dessus de la moyenne). Néanmoins, dans cette partie de l'échelle, les différences de performances entre les pays sont souvent relativement faibles. À l'autre extrémité, 10 pays ont des performances globales significativement inférieures au point central de l'échelle (entre 425 et 488). Les deux derniers pays (Oman et Maroc) ont de forts pourcentages d'élèves dont les performances sont trop basses pour permettre une mesure et des comparaisons fiables (respectivement 16 et 33 %).

À chaque évaluation successive, PIRLS a montré que les filles ont des performances supérieures aux garçons dans la très grande majorité des pays. Ces résultats sont corroborés dans de nombreuses études nationales, voire dans PISA pour les élèves de 15 ans. Les résultats de 2011 ne font pas exception, les filles obtiennent des résultats significativement supérieurs à ceux des garçons dans tous les pays, sauf en Colombie, en Israël, en Espagne, en Italie et en France où cette différence n'est plus significative. Les différences de performances filles/garçons peuvent être très importantes dans certains pays (Finlande : 21 points ; Angleterre : 23 points ; Arabie saoudite : 54 points).

En France, en 2001 et en 2006, l'écart filles/garçons (11 points) était significatif mais faible ; en 2011, il est plus réduit (5 points) et ne distingue plus les deux sexes *(graphique 1)*. Cette réduction de l'écart est due en partie à une moindre performance des filles, dont le score de 2011 est significativement inférieur à celui de 2001. Les jeunes Françaises avec un score moyen de 522 ne sont qu'à 2 points au-dessus de la moyenne internationale des filles (520) ; en revanche les jeunes Français (518) s'écartent sensiblement de la moyenne internationale des garçons (504).

Le *tableau 1* présente la moyenne obtenue à PIRLS 2011 par différents groupes de pays dont la France est proche d'un point de vue géographique ou économique. Les résultats français (520) se retrouvent en deçà de la moyenne des pays voisins, des pays de l'Union européenne ou des pays de l'OCDE.

TABLEAU 1 – Regroupements des pays proches de la France

PIRLS 2011	Score moyen global (Écart-type)
Pays de l'OCDE (25 pays)	538 (70)
Union européenne (23 pays)	534 (71)
France	520 (68)

Lecture : le score moyen des 25 pays (dont la France) appartenant à l'OCDE et présents dans l'étude est de 538.
Source : MEN-MESR-DEPP

L'écart-type de la distribution est un indicateur de la dispersion. Plus il est faible, moins le système éducatif creuse d'écarts entre élèves faibles et forts ; à l'inverse, un écart-type élevé est l'indice que le système éducatif ne réussit pas à combler les inégalités des performances entre les élèves. Pour la France, l'écart-type est de 68 (71 au niveau européen) : les résultats des élèves français sont donc un peu moins dispersés que dans l'ensemble des pays européens de l'étude. Pour ce qui concerne les pays européens, on peut relever les écarts-types les plus bas aux Pays-Bas (54), en République tchèque (61) et en Autriche (63). En revanche, les plus élevés se trouvent à Malte (97), en Roumanie (91) et en Angleterre (82).

Le *graphique 2* présente les scores des pays européens.

Une surreprésentation des élèves français dans le plus faible niveau européen

Si l'on ordonne tous les élèves des pays européens ayant participé à PIRLS en fonction de leur score et que l'on découpe cet ensemble en quatre groupes de même taille, on note que les élèves français se trouvent surreprésentés dans le groupe le plus faible : 32 % contre 25 % et, à l'inverse, qu'ils sont sous-représentés dans le groupe de plus haut niveau : 17 % contre 25 %. Ces données n'ont pas évolué depuis 2006. Les scores moyens de ces différents groupes de niveau montrent une moindre réussite des élèves français du groupe le plus fort par rapport à leurs camarades européens (616 points contre 622) alors que dans le groupe le plus faible ils obtiennent un meilleur score moyen, 441 points contre 436.

Les points de référence internationaux PIRLS

Pour mieux rendre compte des résultats des élèves, PIRLS utilise son échelle internationale des performances pour caractériser ce

que les élèves sont capables de faire à différents points de référence. Quatre points sont ainsi mis en avant sur l'échelle *(tableau 2)* :

– au niveau *avancé* de l'indice de référence (625 points de score), les élèves sont capables de prendre en compte la totalité des éléments du texte pour soutenir leurs interprétations et leurs explications ;

– au niveau *élevé* de l'indice de référence (550 points), les élèves sont en mesure de distinguer les actions et les informations importantes, de faire des inférences et des interprétations à partir d'éléments du texte, d'évaluer le contenu et les éléments textuels, et de reconnaître certaines fonctionnalités de langage ;

– au niveau *intermédiaire* de l'indice de référence (475 points), les élèves peuvent prélever des informations, faire des inférences directes, utiliser certains éléments de présentation, et commencer à reconnaître des caractéristiques du langage ;

– enfin, au niveau *bas* de l'indice de référence (400 points), les élèves sont capables de prélever des informations à partir d'un texte quand celles-ci sont clairement et explicitement énoncées ou faciles à localiser.

Les élèves français faibles relativement mieux classés

Pour chaque pays européen, lorsque l'on compare les élèves les plus faibles, d'une part, et les élèves les plus performants, d'autre part, il apparaît que les 5 % d'élèves français les plus faibles se classent mieux, de manière relative (16e rang européen) que les 20 % d'élèves français les plus performants qui occupent le 20e rang européen.

Stabilité des résultats par rapport à 2001

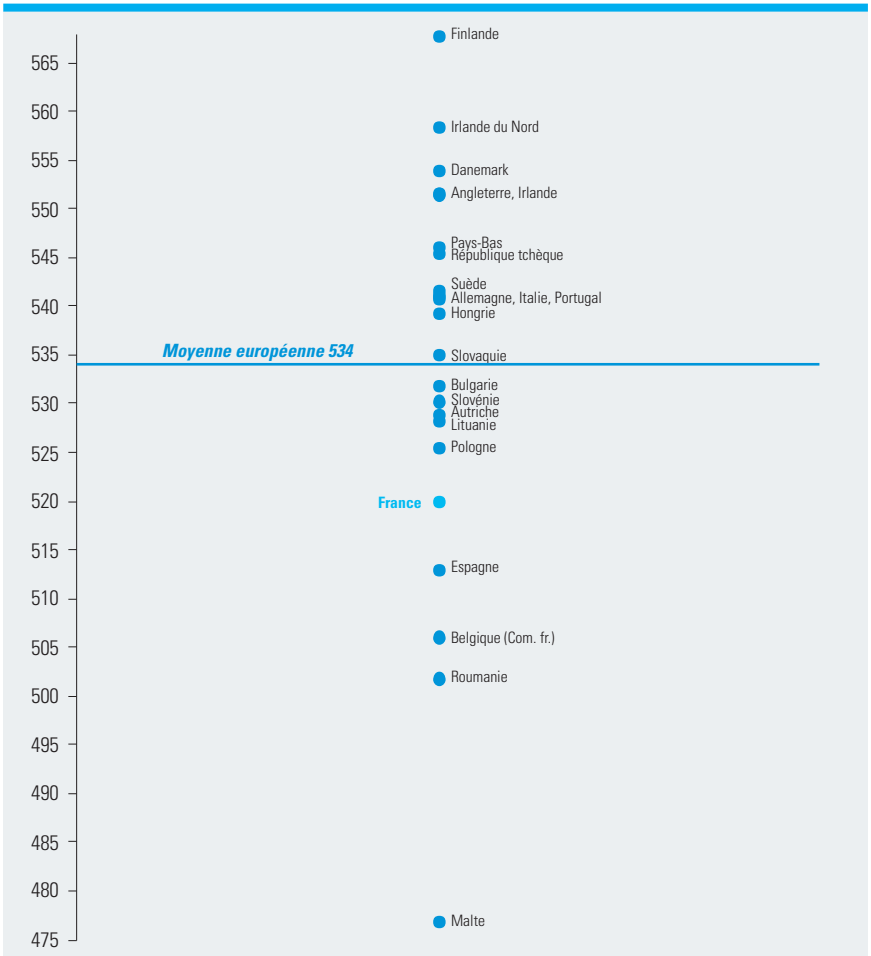
En juin 2008, de nouveaux programmes pour l'école primaire sont venus remplacer ceux de février 2002. Ces nouveaux programmes sont entrés en vigueur à la rentrée scolaire 2008-2009. Ainsi, un élève « à l'heure » dans sa scolarité, évalué dans le cadre de PIRLS, aura effectué son CP avec les programmes de 2002 et ses CE1, CE2 et CM1 régis par les nouveaux programmes de 2008.

L'évaluation 2011 reprend à l'identique quatre textes et leurs questions utilisés en 2006 ainsi que deux textes et leurs questions utilisés en 2001 et en 2006. Ainsi, 60 % des items de 2011 sont communs aux deux périodes 2006/2011, et ce fort ancrage permet une comparaison statistiquement fiable entre les prises d'informations. Ces items ont servi à élaborer l'échelle de 2011 de telle manière que les performances soient parfaitement comparables à celles observées il y a cinq et dix ans.

En 2001, le score global de la France était de 525, en 2006 de 522, et il est de 520 en 2011 *(tableau 3)*. Du point de vue statistique, ces différences, à cinq ou dix ans, ne sont pas significatives, c'est-à-dire qu'il n'est pas possible d'affirmer qu'elles seraient dues à une légère érosion des résultats.

Sur les 23 pays ayant participé à PIRLS à dix ans d'intervalle, 14 obtiennent des scores en 2011 significativement différents de ceux obtenus en 2001. Parmi eux, notons la Bulgarie (- 19 points), la Suède (- 19 points), la Lituanie (- 15 points) et les Pays-Bas (- 8 points), qui voient leurs résultats baisser, tandis que la Slovénie (+ 29 points), la République slovaque (+ 17 points) et la République tchèque (+ 9 points) augmentent leurs performances.

GRAPHIQUE 2 – Classement des pays européens participant à PIRLS 2011



Lecture : ce graphique ne prend en compte que les pays de l'Union européenne qui ont participé à PIRLS 2011. La moyenne de ce groupe de pays est de 534. Avec 520, la France se situe sous la moyenne de ce groupe.

Source : MEN-MESR-DEPP

TABLEAU 2 – Évolution des pourcentages d'élèves atteignant les points de référence internationaux de performances en lecture

		Point de référence			
		Avancé (625)	Élevé (550)	Intermédiaire (475)	Bas (400)
France	2001	7 % ⇩	37 %	77 %	95 %
	2006	5 %	35 %	76 %	96 %
	2011	5 %	35 %	75 %	95 %
Europe	2011	9 %	44 %	80 %	95 %

Lecture : en 2011, 35 % des élèves français atteignent le point de référence élevé contre 44 % des élèves européens.
⇩ : le pourcentage de 2011 est significativement inférieur à celui de 2001.
Source : MEN-MESR-DEPP

Mais les plus fortes hausses ne se font pas en Europe : Iran (+ 44 points), Hong Kong (+ 43 points), Fédération de Russie (+ 40 points) et Singapour (+ 39 points).
En France, entre 2001 et 2011, on note une légère baisse des résultats (- 10 points) dans le public hors éducation prioritaire, tandis que dans le privé et dans l'éducation prioritaire les élèves obtiennent des scores stables (respectivement + 4 et + 3 points). En 2011, tout comme c'était le cas dix ans plus tôt, les seules différences significatives concernent l'éducation prioritaire, dont les performances sont moindres *(tableau 3)*.

TABLEAU 3 – Scores moyens globaux selon le type de scolarisation

Type de scolarisation	Score moyen (Écart-type)		
	2001	2006	2011
Public hors EP	533	525	523 (66)
Privé	527	536	531 (66)
Public EP	477	478	480 (70)
France	**525 (70)**	**522 (67)**	**520 (68)**

Lecture : les élèves scolarisés dans le public en éducation prioritaire (EP) obtenaient un score de 477 en 2001 et de 478 en 2006. Ils obtiennent 480 en 2011.
Source : MEN-MESR-DEPP

Résultats détaillés de 2011

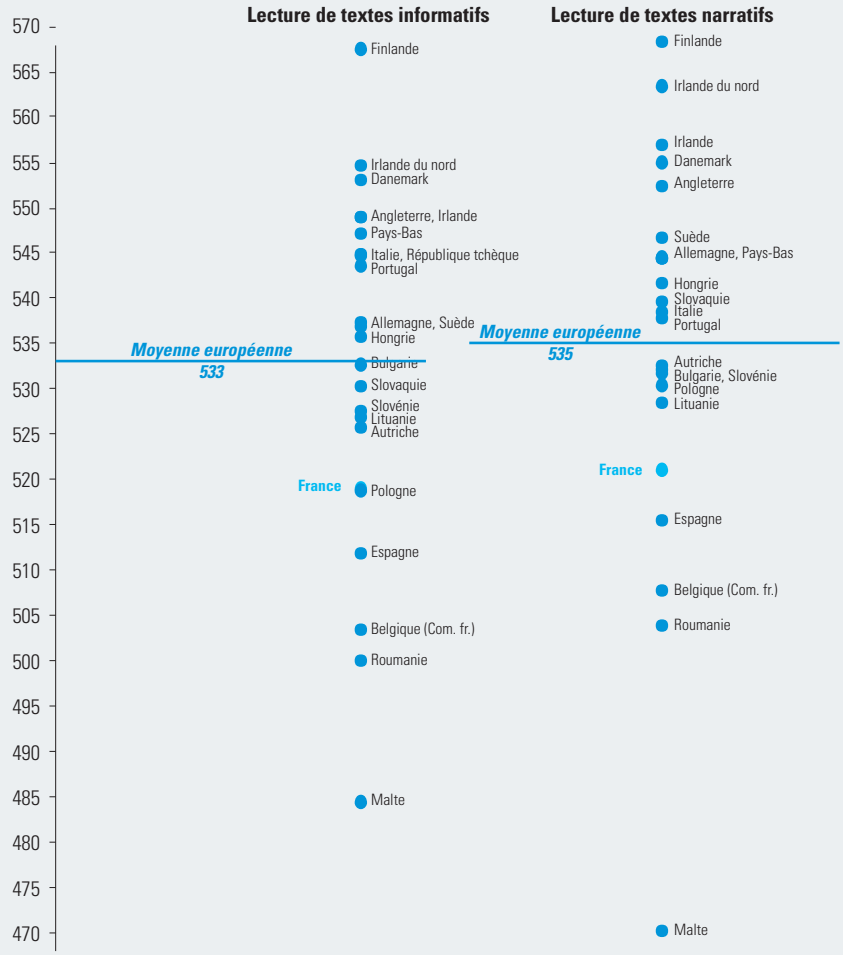
Des résultats variables selon les items

Le score moyen global de la France masque bien sûr des performances variables selon les activités proposées. L'analyse des résultats par item permet de mettre en évidence cette variabilité : en effet, sur l'ensemble des 135 items de l'épreuve, les élèves français réussissent moins bien à 69 items (soit un peu plus de 1 item sur 2) et ce, de manière significative par rapport à la moyenne européenne. À l'inverse, leurs performances sont significativement supérieures pour 17 items (soit 1 item sur 8). Au total, on note donc un comportement qui s'écarte de la moyenne près de deux fois sur trois (64 % des cas).
Parmi les 17 items pour lesquels les élèves français ont des performances significativement supérieures aux autres élèves européens, les items relevant des compétences « Prélever » et « Inférer » sont surreprésentés (88 % contre 59 % dans l'ensemble de l'épreuve), tout comme ceux qui sont en format de QCM (questions à choix multiple de réponses) (65 % contre 55 %).
Parmi les 69 items pour lesquels les élèves français ont des performances significativement inférieures aux autres élèves européens, ceux relevant des compétences « Interpréter » et « Apprécier » sont surreprésentés (57 % contre 41 %), tout comme les items qui imposent un format de réponse construite (51 % contre 45 %), de même que les items issus des textes à caractère informatif (51 % contre 47 % dans l'ensemble de l'épreuve).

Une baisse significative sur les textes informatifs

Pour chacun des deux types de textes (narratifs et informatifs), une échelle de scores a été construite avec une moyenne toujours fixée à 500. Le *graphique 3* rend compte de la distribution des seuls pays de l'Union européenne sur ces deux échelles indépendantes[3] et permet d'apprécier la performance de la France relativement aux autres pays et aux moyennes européennes.
Les élèves français se positionnent de manière significative en dessous de la moyenne européenne dans chacun des deux cas, avec un score de 519 pour les textes informatifs (Europe : 533) et 521 pour les textes narratifs (Europe : 535). Ces deux sous-scores ne s'écartent pas

GRAPHIQUE 3 – Classement des pays européens selon les types de textes



Lecture : parmi les 23 pays européens présents dans l'étude, la France occupe le 19e rang pour les textes narratifs et le 18e rang *ex aequo* (avec la Pologne) pour les textes informatifs.
Source : MEN-MESR-DEPP

3. Il faut comprendre que ces deux échelles, narrative et informative, ne sont pas directement comparables. Elles représentent des concepts différents et les items constitutifs de chaque échelle ont des niveaux de difficulté quelque peu différents. Toutefois, une mise à l'échelle MRI, en calant ces deux échelles sur celles des résultats globaux, prend en compte ces niveaux de difficultés et il devient alors possible de comparer les sous-scores par rapport au score global.

significativement du score global français (520). C'est le cas dans un peu plus de la moitié des pays européens.

Quelques pays, en revanche, ont des profils plus tranchés. L'Irlande, l'Irlande du Nord, la Pologne, la Hongrie, la République slovaque, l'Allemagne, l'Autriche, et particulièrement la Suède, montrent des écarts significatifs en faveur des textes narratifs. À l'inverse, les élèves du Portugal, de Malte et d'Italie obtiennent des scores significativement supérieurs aux textes informatifs.

En général, les augmentations ou les diminutions des scores globaux depuis 2001 et 2006 se sont traduites, pour les différents pays, par des augmentations ou des diminutions à la fois pour les textes narratifs et pour les textes informatifs. La France cependant, suit un schéma notablement différent. Alors que les performances sur la lecture des textes narratifs sont restées relativement stables, les performances qui concernent la lecture des textes de type informatif ont diminué depuis 2001 (13 points), effaçant ainsi les différences de performances observées entre les deux types de textes en 2001 et 2006 *(tableau 4)*.

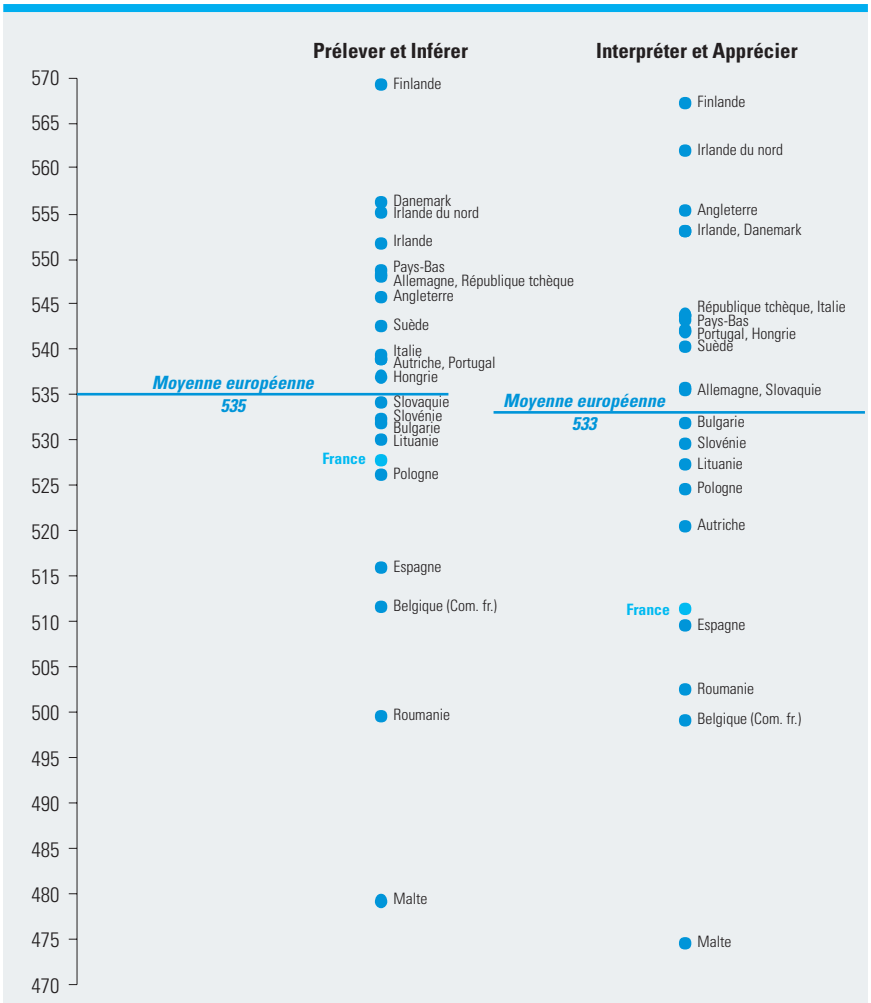
Pour conclure sur ce point, il est intéressant de noter que les filles françaises réussissent significativement mieux que les garçons les textes narratifs (526 points contre 517), alors qu'aucune différence n'est observée entre les deux sexes pour les textes informatifs (519 points dans les deux cas).

TABLEAU 4 – Évolution du score moyen de la France en fonction des types de textes

PIRLS France	Textes informatifs					Textes narratifs		
	Score moyen	Différence entre années				Score moyen	Différence entre années	
		2006		2001			2006	2001
2011	519	- 7	⇩	- 13	⇩	521	4	2
2006	526			- 6		517		- 2
2001	532					519		

Lecture : en 2011, le score moyen de la France sur les textes informatifs est de 519. C'est 13 points de moins qu'en 2001.
⇩ : la baisse est significative.
Source : MEN-MESR-DEPP

GRAPHIQUE 4 – Classement des pays européens selon les compétences



Lecture : au sein des 23 pays européens de l'étude, La France occupe le 18e rang pour les compétences « Prélever et Inférer » et le 19e rang pour les compétences « Interpréter et Apprécier ». Sur ce graphique, les compétences sont présentées dans un ordre croissant de complexité.

Source : MEN-MESR-DEPP

Une baisse significative sur les compétences les plus complexes

Les quatre compétences « Prélever », « Inférer », « Interpréter » et « Apprécier » ont été regroupées deux à deux : « Prélever et Inférer », « Interpréter et Apprécier », afin de créer deux échelles de scores indépendantes. Le *graphique 4* rend compte de ces deux échelles indépendantes[4] pour les seuls pays européens. Si par construction la moyenne internationale de ces deux échelles est de 500, la moyenne européenne est de 535 pour ce qui concerne « Prélever et Inférer », et de 533 pour « Interpréter et Apprécier ».

En général, les pays qui obtiennent les performances globales les plus élevées obtiennent aussi les performances les plus élevées sur les deux échelles des compétences. On observe cependant une tendance pour les pays les mieux classés à obtenir des performances relativement plus élevées pour « Interpréter et Apprécier » que pour « Prélever et Inférer ».

Les élèves français se positionnent de manière significative en dessous de la moyenne européenne dans chacun des deux cas avec un score de 528 pour « Prélever et Inférer » (Europe : 535) et 512 pour « Interpréter et Apprécier » (Europe : 533). Ces deux sous-scores français s'écartent significativement du score global français (520).

Cinq autres pays européens se comportent comme la France sur ce point : La Belgique (Communauté francophone), l'Espagne, l'Allemagne, l'Autriche et les Pays-Bas.

4. *Cf.* note précédente. Ici, c'est par construction que les items constitutifs de chaque échelle ont des niveaux de difficultés différents puisque les compétences sont hiérarchisées dans leur degré de complexité. Au niveau international, le taux de réussite des items est de 64 % pour « Prélever et Inférer » et de 45 % pour « Interpréter et Apprécier ».

À l'inverse, les élèves de Hongrie, d'Irlande du Nord et d'Angleterre obtiennent des scores significativement supérieurs pour « Interpréter et Apprécier ». Il est vraisemblable que ces différences de profil s'expliquent par l'accent plus ou moins marqué que met chaque système éducatif sur l'enseignement de ces compétences à ce niveau de scolarité.

Pour la France, on observe une asymétrie de l'évolution des performances sur dix ans *(tableau 5)*. Le score moyen des compétences « Prélever et Inférer », les moins exigeantes, restent relativement stables sur la décennie, alors que les compétences « Interpréter et Apprécier », plus complexes, voient leur score moyen diminuer sur la même période (11 points).

Pour la France, la différenciation par genre montre que si les filles réussissent significativement mieux que les garçons pour « Prélever et Inférer » (531 points contre 525), en revanche aucune différence n'est observée entre eux pour les compétences plus complexes « Interpréter et Apprécier » (513 points contre 510).

Une synthèse récente du nombre d'élèves en difficulté face à l'écrit[5] montre que, depuis une dizaine d'années, ce nombre a augmenté : près d'un élève sur cinq est aujourd'hui concerné en début de sixième. À partir des deux enquêtes de la DEPP, « Lire, écrire, compter » et « Difficultés de lecture en début de sixième », cette synthèse fait ressortir que la dégradation du niveau de performance face à l'écrit concerne plus particulièrement les élèves les plus en difficulté. Le niveau des élèves situés autour de la médiane n'a, quant à lui, pas évolué. C'est ce que confirment les résultats PIRLS 2011, dont le dispositif est particulièrement précis pour mesurer le niveau des élèves situés autour de la médiane.

TABLEAU 5 – Évolution du score moyen de la France en fonction des compétences

PIRLS France	« Prélever et Inférer »			« Interpréter et Apprécier »			
	Score moyen	Différence entre années		Score moyen	Différence entre années		
		2006	2001		2006	2001	
2011	528	1	- 1	512	- 4	- 11	⇩
2006	527		- 2	515		- 7	⇩
2001	529			523			

Lecture : en 2011, le score moyen de la France sur les compétences « Interpréter et Apprécier » est de 512. C'est 11 points de moins qu'en 2001.
⇩ : la baisse est significative.
Source : MEN-MESR-DEPP

Les formats de questions

L'évaluation PIRLS a recours à deux formats de questions : d'une part des questions à choix multiples (QCM) et d'autre part des questions dites à réponses construites, qui imposent à l'élève une production écrite plus ou moins longue. Les élèves français réussissent significativement mieux lorsqu'il s'agit de répondre à des QCM que lorsqu'il s'agit de rédiger des écrits, avec respectivement 67 % et 40 % de réussite. Parmi les questions dont les réponses doivent être rédigées, on note des différences selon que la réponse requise doit être longue ou brève. Plus la réponse attendue doit être élaborée, plus le score des élèves français diminue : 53 % de réussite aux questions à réponses construites brèves (un mot, un groupe de mots) ; 31 % de réussite aux questions où la réponse attendue est de l'ordre d'une phrase ; 20 % de réussite aux questions où on attend un petit paragraphe. Ce phénomène s'observe pour l'ensemble des pays et, à ce titre, la France ne se distingue pas de la moyenne européenne. C'est en Angleterre et en Irlande du Nord que l'on trouve les écarts de réussite les plus faibles d'un type de question à l'autre, c'est en revanche en Autriche que cet écart est le plus marqué.

Le format de question est lié à la compétence évaluée. Ainsi, « Prélever et Inférer » fait plus souvent appel aux questions à choix multiples (68 % des cas), alors que « Interpréter et Apprécier » demande plus souvent de développer un raisonnement et exige généralement une réponse construite et plus longue (64 % des cas). Dans ce cas, pour obtenir un bon score, des capacités en production d'écrits s'ajoutent donc à celles de la compréhension. Pour autant, un outil d'analyse statistique, la régression linéaire, rend possible la neutralisation de ce lien. Lorsqu'on neutralise par ce moyen l'effet des formats de questions, la hiérarchie des scores de réussite des élèves français au niveau des compétences demeure inchangée. Symétriquement, toutes compétences étudiées égales par ailleurs, les élèves obtiennent de meilleurs résultats quand le mode de réponse requis n'implique pas le recours à l'écrit construit.

Absences de réponse au fil de l'épreuve…

Les taux d'absence de réponse semblent liés aux formats de questions. Plus la réponse requise doit être élaborée et demande un développement écrit, plus les élèves ont tendance à s'abstenir. Comme en 2001 et en 2006, la France se distingue néanmoins des autres pays par l'amplitude du phénomène. Le taux d'absence de réponse des élèves français est de 1,9 % pour les QCM alors qu'il est six fois plus élevé lorsqu'il s'agit de productions écrites. Ce taux est même multiplié par huit lorsque l'exigence de longueur de la réponse attendue est la plus forte ; il s'agit du plus fort taux d'absence de réponse parmi les pays européens.

C'est un problème récurrent, pointé dans PIRLS depuis dix ans, c'est aussi un problème dont l'ampleur est importante : 6,4 % des réponses attendues ne sont pas fournies par les élèves.

… et abandons avant la fin de l'épreuve

Si les élèves français sont les plus nombreux au sein des pays européens à s'abstenir de répondre, ils sont aussi les plus nombreux (avec les élèves belges) à ne pas terminer les épreuves qui leur sont proposées. En moyenne, pour chacun des textes, 6,7 % des élèves français ne vont pas jusqu'au bout de la quinzaine de questions posées. C'est sept fois plus que les élèves finlandais, néerlandais ou irlandais du Nord. Ce pourcentage d'élèves qui ne terminent pas les épreuves varie en fonction du type de textes ; 4,7 % en moyenne lorsque l'épreuve porte sur un texte informatif, c'est presque deux fois plus que la moyenne européenne ; 8,7 % dans le cas d'un texte narratif, c'est là près de trois fois plus que la moyenne

5. Daussin, J.-M., Keskpaik, S. et Rocher, T. (2011), « L'évolution du nombre d'élèves en difficulté face à l'écrit depuis une dizaine d'années », *France Portrait Social*, pp. 137-152.

européenne et plus de dix fois plus que les élèves néerlandais qui sont ceux qui vont le plus souvent jusqu'au bout de ces questions. Il est possible d'estimer à 1,3 % les réponses attendues dans PIRLS qui ne sont pas fournies par les élèves français pour cette raison. Ce pourcentage s'ajoute à celui des absences de réponses au fil de l'épreuve ; c'est donc au total 7,7 % (6,4 + 1,3) des réponses attendues que les élèves français laissent vierges. Ceci a un impact mécanique sur le niveau de leurs performances.

Marc Colmant et
Marion Le Cam, DEPP B2

Pour en savoir plus

- « Les élèves de CM1 - Premiers résultats d'une évaluation internationale en lecture (PIRLS 2001) », *Note d'Information* 03.22, MEN-DEPP, avril 2003.

- « Évolution des performances en lecture des élèves de CM1. Résultats de l'étude internationale PIRLS 2006 », *Note d'Information* 08.14, mars 2008.

- Site Internet de l'étude : http://timss.bc.edu/

Il est possible d'y télécharger notamment les documents suivants :

PIRLS 2011 Assessment Framework, août 2009 (présentation du cadre de l'étude) ;

PIRLS 2011 International Report, décembre 2012 (rapport international) ;

PIRLS 2011 Encyclopedia, septembre 2012 (guide de référence des systèmes éducatifs des pays participants).

Tous les documents des PIRLS 2001 et 2006 sont également disponibles à la même adresse.

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

Méthodologie

Cadre

PIRLS *(Progress in International Reading Literacy Study)* est pilotée par l'IEA (Association internationale pour l'évaluation du rendement scolaire) et mise en œuvre par l'*International Study Center* du *Boston College*. Les 45 pays sont les suivants : Allemagne, Angleterre, Arabie saoudite, Australie, Autriche, Azerbaïdjan, Belgique (Communauté francophone), Bulgarie, Canada, Colombie, Croatie, Danemark, Émirats arabes unis, Espagne, États-Unis, Fédération de Russie, Finlande, France, Géorgie, Hongrie, Indonésie, Iran, Irlande, Irlande du Nord, Israël, Italie, Lituanie, Malte, Maroc, Norvège, Nouvelle-Zélande, Oman, Pays-Bas, Pologne, Portugal, Qatar, Roumanie, Singapour, République slovaque, République tchèque, Slovénie, Suède, Taïwan, Trinité-et-Tobago. En outre, certaines provinces ont participé pour elles-mêmes : Alberta, Ontario, Québec (Canada) ; Floride (USA) ; Andalousie (Espagne), Hong Kong, Dubaï, Abu Dhabi (EAU).

L'échantillon

En France, l'enquête porte sur un échantillon de 174 écoles primaires et élémentaires. Le tirage de l'échantillon tient compte de la taille des écoles et de quelques caractéristiques (public/privé ; ZEP/hors ZEP). Au sein de chaque école, et suivant sa taille, une ou deux classes entières ont été retenues. L'échantillon français regroupe 4 438 élèves effectivement évalués, répartis au sein de 277 classes.
Au niveau international, l'échantillon porte sur 10 539 écoles et 291 709 élèves.

Comparaison 2001-2006-2011

Pour assurer la comparabilité dans le temps, l'évaluation PIRLS a réutilisé une partie des textes et leurs questions afférentes des évaluations de 2001 et de 2006 pour celle de 2011. Ainsi, en 2011, l'ensemble du protocole est composé de dix blocs (textes et questions) dont quatre (deux informatifs et deux narratifs) sont issus de l'épreuve de 2006 et deux (un informatif et un narratif) de l'épreuve de 2001. Ces six blocs, qui avaient été gardés sous embargo, ont ainsi servi d'ancrage pour les échelles de comparaison dans le temps. Ainsi, quatre blocs nouveaux ont été introduits en 2011.

Les procédures de passation

La mise en œuvre de l'enquête est basée sur des procédures standardisées afin de garantir la comparabilité des résultats : désignation de responsables de l'enquête dans chaque établissement, respect des consignes de passation, procédures de contrôle, etc. L'évaluation, de type « papier-crayon », est fondée sur la technique dite des « cahiers tournants ». Les textes et leurs questions ont été répartis dans treize cahiers d'évaluation différents, à raison de deux textes par cahier, en ménageant des parties communes. Ces treize cahiers ont été distribués aléatoirement aux élèves (à raison d'un seul cahier par élève), qui avaient deux fois 40 minutes pour y répondre. Ce procédé a ainsi permis de tester l'équivalent de près de 7 heures d'évaluation sans allonger le temps de passation individuel.
En ce qui concerne les questionnaires de contexte, ils sont au nombre de quatre et ont dû être renseignés par les élèves eux-mêmes ; leurs parents ; leurs enseignants ; ainsi que les directeurs des écoles concernées. Ces questionnaires de contexte permettront de cerner l'environnement des élèves et d'apporter des informations supplémentaires sur leur rapport à la lecture, éclairant ainsi certains résultats.

La construction des scores

Une échelle de score a été élaborée en utilisant des modèles statistiques particuliers (modèles de réponse à l'item : MRI). Cette approche permet de distribuer les performances des élèves sur une même échelle même si, individuellement, ils ont répondu à des tests différents *(cf. PIRLS 2011 Technical Report)*. En 2001, la moyenne internationale a été fixée à 500 et l'écart-type à 100. Pour permettre les comparaisons temporelles, les résultats de 2006 et 2011 ont été distribués sur cette échelle.
Le choix d'une telle échelle est arbitraire, mais son intérêt est de permettre de classer les pays sur une dimension commune. On comprend que cette méthodologie met l'accent sur une évaluation « relative » des performances (comparaison de pays à pays), même si l'aspect évaluation « absolue » du niveau des performances n'est pas totalement absent.

Les corrections

Les questions à réponses construites ont été corrigées par un groupe d'enseignants formés à cette tâche suivant des consignes internationales strictes par le biais d'une interface informatique. Dans chaque pays, un lot de cahiers a subi une double correction permettant de vérifier le degré d'accord des correcteurs. D'autre part, pour vérifier la fiabilité des corrections entre chacun des pays participants, un lot de copies a été corrigé par les correcteurs de tous les pays. Enfin, pour maîtriser une éventuelle évolution dans la sévérité des correcteurs d'un cycle à l'autre de l'étude, une étude de fiabilité a été réalisée en proposant aux correcteurs de 2011 un lot de copies de 2001 et 2006.

DIRECTION DE L'ÉVALUATION,
DE LA PROSPECTIVE ET DE LA PERFORMANCE
Directeur de la publication : Michel Quéré

Secrétaire de rédaction : Marc Saillard
Maquettiste : Frédéric Voiret
Impression : DEPP/DVE

DEPP, Département de la valorisation et de l'édition
61-65 rue Dutot – 75732 Paris Cedex 15
depp.diffusion@education.gouv.fr
Reproduction autorisée à la condition expresse de mentionner la source
ISSN 1286-9392