Éclairages des évaluations internationales PIRLS et PISA sur les élèves français

Focusing on PIRLS and PISA international assessment of French students

Enfoques de lasvaluaciones internacionales PIRLS y PISA sobre los alumnos franceses

Erläuterung der internationalen Studien PIRLS und PISA, was französische Schüler angeht

Martine Rémond

Édition électronique

URL : <http://journals.openedition.org/rfp/433>

DOI : 10.4000/rfp.433

ISSN : 2105-2913

Éditeur

ENS Éditions

Édition imprimée

Date de publication : 1 décembre 2006

Pagination : 71-84

ISBN : 978-2-7342-1075-7

ISSN : 0556-7807

Référence électronique

Martine Rémond, « Éclairages des évaluations internationales PIRLS et PISA sur les élèves français », *Revue française de pédagogie* [En ligne], 157 | octobre-décembre 2006, mis en ligne le 01 décembre 2010, consulté le 19 avril 2019. URL : <http://journals.openedition.org/rfp/433> ; DOI : 10.4000/rfp.433

Éclairages des évaluations internationales PIRLS et PISA sur les élèves français

Martine Rémond

Dans cet article, les résultats aux deux enquêtes internationales sur la reading literacy sont analysés et rapprochés de ceux des dernières évaluations-bilans françaises. Qu'ils soient en CM1 pour PIRLS ou évalués à 15 ans dans PISA, les élèves français font preuve d'un niveau assez moyen en matière de reading literacy parmi la trentaine de pays participants. Nos analyses croisées des dispositifs d'évaluation permettent de percevoir une sorte de genèse de l'hétérogénéité révélée par PISA en 2003 et de formuler des recommandations sur les contenus et sur les stratégies d'enseignement.

Descripteurs TEE : appréciation, compréhension de texte, conduite de la classe, études internationales, évaluation normative, expérience d'enseignement, métacognition.

À l'entrée dans le nouveau millénaire, deux programmes cycliques d'évaluation de la *reading literacy* ont été engagés : en 2000, PISA (*Programme for International Student Assessment*) sous l'égide de l'OCDE, en 2001, PIRLS (*Progress in International Reading Literacy Study*) sous celle de l'IEA (1). Ils entretiennent de nombreuses similitudes tant par leur conception que par leurs méthodologies.

PIRLS prend place au moment où les élèves passent de l'apprentissage de la lecture à la lecture au service des apprentissages (quatrième année de scolarisation obligatoire), alors que PISA porte sur des élèves de quinze ans, se situant donc à une phase de transition entre l'école et le monde du travail. PIRLS

cherche « à compléter PISA en adoptant une perspective centrée sur l'acquisition » (Campbell *et al.*, 2000, p. 62) (2). Il a été prévu pour recueillir des informations sur la compréhension des élèves d'une dizaine d'années en s'intéressant, entre autres, aux facteurs qui facilitent l'acquisition de la *reading literacy* aussi bien à l'école qu'à la maison. Il tente de mettre en évidence des liens entre le niveau de compréhension des élèves, le programme d'enseignement et les pratiques pédagogiques concernant la lecture.

Ces deux enquêtes reposent sur des processus d'élaboration similaires gouvernés par des cadres conceptuels de référence pour l'évaluation (*framework*), élaborés par des groupes d'experts internationaux

spécialistes du domaine (3). Chacun des cadres de référence prévoit, dans le détail, la structuration de l'enquête depuis la définition de l'objet à évaluer jusqu'aux échelles de présentation des résultats. Rendus accessibles par leur publication (Campbell *et al.*, 2000 ; OCDE, 1999), leur fonction méthodologique va au-delà de l'enquête qu'ils organisent : ils peuvent servir à nourrir la réflexion sur les contraintes d'un protocole d'évaluation, la première étant de définir l'objet qui sera évalué.

L'ORGANISATION GÉNÉRALE DE PIRLS ET DE PISA

Les définitions

Les définitions de la *reading literacy* conditionnent l'élaboration des dimensions de la compréhension. Dans PIRLS, la *reading literacy* est définie comme « l'aptitude à comprendre et à utiliser les formes du langage écrit dont la société a besoin et/ou qui sont nécessaires à l'individu. Les jeunes lecteurs sont amenés à construire le sens 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. » (Campbell *et al.*, 2000, p. 3). Dans PISA, la *reading literacy* est définie comme « la capacité de comprendre, d'utiliser et d'analyser des textes écrits, mais aussi de réfléchir à leur propos afin de pouvoir réaliser ses objectifs, développer ses connaissances et son potentiel et jouer un rôle actif dans la société » (OCDE, 1999, p. 20)

Ces deux définitions, élaborées à un an d'intervalle, reflètent une conception interactive de la lecture où le lecteur fait appel à ses connaissances et à ses stratégies pour construire la signification du texte et utiliser l'écrit. Dans PIRLS, l'élève lit pour apprendre, mais comme dans PISA, il doit aussi participer à la vie collective, alors appelée « *communauté de lecteurs* », et répondre aux exigences de la société par le biais de la compréhension d'écrits variés.

La définition élaborée pour PISA permet d'envisager la dimension interprétative en mettant l'accent sur la capacité de « réflexion » ou de distance critique du lecteur, appelée *reflect on*. Elle ouvre aussi un horizon vers la « *préparation à la vie d'adulte* », les connaissances et les savoir-faire évalués étant supposés servir dans la vie d'adulte, d'où l'emploi de *reading literacy* (4).

Une fois donnée la définition de l'objet à évaluer, il faut fixer ce qui doit l'être. Nous reprendrons les terminologies employées dans chacun des programmes.

Que va-t-on évaluer ? Domaines de compétences évalués, caractéristiques des tâches

Le protocole d'évaluation (textes et items) de PIRLS est bâti sur la base du croisement de deux dimensions : les **processus de compréhension** (*processes of comprehension*) et les **objectifs de la lecture** (*purposes for reading*) qui « constituent la base de l'évaluation écrite de la compréhension de l'écrit » (Campbell *et al.*, 2000, p. 4). Des questionnaires sur les habitudes de lecture et les attitudes vis-à-vis de l'écrit sont soumis d'une part aux élèves, d'autre part à leurs parents. D'autres ciblant entre autres choses les pratiques pédagogiques et la formation sont renseignés par les maîtres des élèves évalués ainsi que par les directeurs d'école.

Trois dimensions ont été prises en considération pour élaborer le matériel d'évaluation de PISA : les types de tâches de lecture, les « types de textes » et le contexte.

Les processus de compréhension/les types de tâches de lecture

Les quatre processus que visait à évaluer PIRLS étaient les suivants (nous indiquons entre parenthèses la « dénomination courte » du processus qui sera employée par la suite) : « Prélever des informations explicites » (« Prélever »), « Faire des inférences directes » (« Inférer »), « Interpréter et assimiler idées et informations » (« Interpréter »), « Examiner et évaluer le contenu, la langue et les éléments textuels » (« Évaluer »). Chaque processus de compréhension est évalué pour chaque objectif défini ci-dessous.

Dans PISA, les « types de tâche de lecture » (*questions ou aspects* en anglais) définissent la nature des capacités à évaluer et sont considérées se situer à un « macro-niveau ». Ces tâches de lecture, au nombre de cinq au départ, ont été regroupées en trois catégories : « Retrouver des informations dans des matériaux variés » (« S'informer »), « Interpréter ce que l'on lit » (« Interpréter »), « Réfléchir à ce qu'on vient de lire et y réagir » (« Réfléchir »). Trois variables guidant la construction et la correction des items ont été isolées par I. Kirsch, A. Jungleblut et P. B. Mosenthal (1998) ; contrairement à PIRLS, elles sont mentionnées dans le cadre conceptuel de PISA. Ces variables qui jouent un rôle sur la difficulté de la tâche sont les suivantes : la plausibilité des distracteurs, le type d'information

recherché et le type d'appariement à réaliser. La plausibilité des distracteurs est une variable spécifique aux QCM. Ce format d'item se caractérise par plusieurs modalités de réponse : la bonne réponse et les autres qui s'appellent distracteurs. Plus les modalités de réponses partagent de traits communs avec la réponse pertinente, plus la tâche est jugée difficile. Le type d'information à rechercher pour répondre varie sur un *continuum* allant du très concret au très abstrait. Plus l'information attendue est concrète, plus la tâche est jugée facile. L'appariement entre la question et le texte est influencé par diverses conditions qui contribuent à la difficulté de la tâche : le nombre de phrases qui doit être utilisé dans la recherche, le niveau d'inférences à effectuer pour aller de la question au texte et pour identifier l'information requise, le nombre d'informations demandées par la question et leur localisation dans la phrase... L'appariement le plus simple se fait lors de tâches de localisation qui ne nécessitent que d'établir une correspondance directe, littérale ou synonyme, entre la question et le texte.

Types de supports

PIRLS retient deux **objectifs de lecture** : « Lire pour accéder aux textes littéraires » et « Lire pour acquérir et utiliser des informations », car « ils correspondent à la majorité des lectures faites par les jeunes élèves, à la fois à l'école et en dehors, et donc aux raisons de lire à cet âge » (Campbell *et al.*, 2000, p. 15).

PISA développe la dimension « **type de textes** ». Reprenant une classification présente dans des enquêtes antérieures de l'OCDE, la distinction est faite entre les textes « continus » (descriptif, narratif, informatif, argumentatif et injonctif) et les textes « non continus » (formulaire, graphiques, cartes, etc.).

En raison de la différence d'avancement des populations respectives dans leur scolarité, il est cohérent d'utiliser une plus grande variété de supports à quinze ans. Dans PIRLS, la distinction entre les deux objectifs de lecture laisse attendre des textes de nature différente (et variés si l'on se réfère à la définition). La *framework* de PISA fixe les « types de textes » et la proportion d'items dont ils seront la cible, de même que l'utilisation pour laquelle le texte a été écrit : « le contexte » ou « la situation » qui se décline en lecture à usage personnel, public, professionnel et scolaire. La lecture à usage personnel est effectuée pour satisfaire ses besoins personnels alors que la lecture à des fins publiques est pratiquée pour prendre part à des activités sociales au sens large. La lecture à des fins professionnelles peut surprendre dans le cadre

de ce travail, il s'agit de déterminer si les élèves de quinze ans sont prêts à une entrée dans le monde du travail et s'ils peuvent réussir des tâches de « lecture pour agir ». Cette dimension « contexte » est supposée refléter des situations de lecture représentatives de celles que fréquente la population de quinze ans.

Le contenu des protocoles et le traitement des données

Une méthodologie complexe conditionne la construction et l'organisation du matériel, élaboré en anglais pour former la « version source » servant de base commune aux traductions. À la base, un pool d'items est constitué (141 pour PISA 2000), chaque élève ne verra qu'une partie du matériel (textes et items afférents) selon le principe des « cahiers tournants ». Les protocoles de PIRLS reposent sur des textes longs et intégraux (au nombre de huit) sur chacun desquels portent une douzaine de questions. PISA comme d'autres évaluations de l'OCDE procède à partir d'un grand nombre d'exercices à l'unité (*unit*) qui contrairement à ceux de PIRLS sont plutôt courts, car il faut croiser de nombreux facteurs pour obtenir la déclinaison d'items correspondant aux caractéristiques définies (voir OCDE, 1999 & 2003).

Deux formats de questions sont, pour l'essentiel (il en est d'autres, au moins dans PISA, d'importance moindre) utilisés dans les deux enquêtes : des QCM à quatre modalités et des questions nécessitant de construire des réponses selon une modalité écrite.

Le codage des réponses aux questions ouvertes est fait en fonction des instructions figurant dans un guide afin d'effectuer des corrections aussi standardisées que possible, d'un pays à l'autre, d'une langue à l'autre. Les items les plus simples sont crédités d'un seul point tandis que d'autres, plus complexes, sont pondérés en fonction de la précision de la réponse.

Au final, dans un programme comme dans l'autre, on aboutit à une échelle de scores de compréhension élaborée à partir des modèles de réponse à l'item (MRI), la moyenne internationale est fixée à 500 et l'écart-type à 100 ; les pays sont classés sur une dimension commune (cf. Vignaud, 2006). Les résultats sont exprimés sur une échelle combinée de compréhension, et pour PISA sur trois échelles dites spécifiques rendant chacune compte d'un des trois types de tâches de lecture.

La méthodologie met l'accent sur la comparaison des performances entre les pays et dans les rapports internationaux publiés par l'OCDE ou par l'IEA,

« les scores des pays sont présentés comme sur un thermomètre (échelle de compétences), et conduisent à un palmarès... » (Rémond, 2005, p. 122). Dans les analyses qui vont suivre, nous nous appuierons essentiellement sur ces rapports (Mullis *et al.*, 2003 ; OCDE, 2001 & 2003) et sur des études complémentaires que nous avons menées à la demande de la DEP (Rémond, 2005 et soumis).

LES ÉLÈVES FRANÇAIS À L'AUNE DE LA « *READING LITERACY* »

Les deux enquêtes situent à un niveau moyen les performances en compréhension de l'écrit des élèves français. En 2000 comme en 2003, la France obtient aux épreuves de compréhension de l'écrit de PISA des scores proches de la moyenne, les scores respectifs étant de 505 et 496. Pour celles de PIRLS 2001, elle obtient un score significativement supérieur à la moyenne des pays participants (525).

Il n'y a pas de corrélation systématique entre les résultats aux deux programmes : ainsi trois pays ayant des scores moyens semblables à PISA (France, Norvège et États-Unis) obtiennent des scores moyens significativement différents les uns des autres à PIRLS.

Choix d'un sous-échantillon

Certains pays ont participé à la fois à PIRLS 2001 et à PISA 2000, mais ce n'est pas la majorité. Afin de pouvoir esquisser quelques comparaisons, nous avons sélectionné, parmi ces pays, un échantillon auquel nous limiterons notre propos. La langue anglaise y est représentée par le Royaume-Uni, les États-Unis, la Nouvelle-Zélande et le Canada (État de l'Ontario), et la langue française par la France et par le Canada, province du Québec ; l'Allemagne, l'Italie, la Norvège et la Suède font également partie de cet échantillon.

Notons que dans cette configuration, les meilleures performances sont obtenues pour PIRLS par la Suède et l'Angleterre et à PISA, par le Canada et la Nouvelle-Zélande.

Penchons-nous maintenant sur les dimensions présentées en première partie qui structurent les Programmes et tout d'abord sur les objectifs de lecture-types de supports.

La structuration des performances

Un effet des types de textes

Dans PIRLS, la France obtient de meilleurs scores aux items évaluant l'objectif « Lire pour acquérir et utiliser des informations » qu'à ceux qui relèvent de « Lire pour accéder aux textes littéraires » (5) (scores moyens respectifs : 533 *contra* 518), comme le Québec. En revanche, en Italie, en Norvège et dans tous les pays anglophones, l'objectif « Littéraire » donne lieu à de meilleures performances que l'objectif « Informatif ». Enfin, en Suède et en Allemagne les deux objectifs de lecture mènent à des performances similaires.

Dans PISA, la France réussit beaucoup mieux les items portant sur les textes « non continus » que ceux qui sont relatifs aux « textes continus », c'est d'ailleurs le pays qui présente l'écart le plus fort entre ces deux « types de textes ». L'Italie représente un cas isolé en réalisant de meilleures performances aux « textes continus » dans les deux programmes. Cet effet du « format des textes » peut probablement se rattacher à des différences culturelles et à des options pédagogiques privilégiant davantage certains écrits que d'autres.

Des capacités inférentielles et interprétatives limitées

Dans les deux Programmes, les échelles de performances exprimées en termes de scores de compréhension ont été divisées en **niveaux** qui rendent compte de la progression de la difficulté de la tâche. À chaque niveau, correspondent des tâches de lecture que les élèves sont capables d'accomplir avec une certaine réussite, montrant ainsi quelles compétences ils sont capables de mobiliser

Pour PIRLS, quatre niveaux ont été ainsi définis en fonction des quartiles (Niveaux 1 à 3, correspondant aux scores de compréhension atteints par 25, 50 ou 75 % de la population générale) et d'un repérage spécifique des 10 % ayant les meilleurs scores (niveau 4).

Dans ce programme, alors qu'elle a conditionné la construction du matériel, la variable « Processus » est à peine exploitée dans les analyses ; nous avons cherché à mieux en comprendre le poids à partir des niveaux de compétences et de leurs caractéristiques que nous avons synthétisées.

Au niveau 1, les élèves savent prélever des informations explicites et établir des inférences simples. Au niveau 2, ils maîtrisent le niveau précédent et de plus,

ils sont capables d'effectuer des inférences et des interprétations simples à partir d'informations puisées dans différentes parties du texte. Ils comprennent la structure générale du texte. Au niveau 3, ils sont également capables d'effectuer des inférences qui s'appuient sur différentes caractéristiques des personnages et des événements, et ils savent les justifier. Ils interprètent en faisant appel à leurs connaissances et expériences personnelles. Ils comprennent des procédés tels que la métaphore simple. Enfin au niveau 4, ils peuvent interpréter les intentions, les sentiments, les comportements des personnages en se basant sur le texte, ils sont capables d'intégrer des idées pour dégager le thème.

Le prélèvement d'information n'est mentionné qu'au niveau 1, mais il contribue bien évidemment à tous les niveaux. On observe que du niveau 1 au niveau 3, la nature des inférences évolue, mobilisant des processus attentionnels et langagiers de plus en plus complexes. La dimension interprétative apparaît dès le niveau 2, mais ce n'est qu'au niveau 3 que les élèves manifestent qu'ils savent aller au delà du texte, et au niveau 4, qu'ils peuvent expliciter leur démarche en se basant sur le texte.

Le tableau 1 présente les profils de résultats des pays en fonction des quatre niveaux définis ainsi que leur score global à PIRLS. L'ordre des pays figurant

dans le tableau a été commandé par des critères de langue du test en vue de faire apparaître un regroupement d'anglophones et un autre très modeste de francophones.

Le pourcentage théorique de 75 % définissant le niveau 1 est largement dépassé par la plupart des pays sur lesquels porte notre comparaison. Si l'on s'intéresse maintenant aux profils de distribution des élèves selon les niveaux et les pays, on constate qu'à une exception près, la Norvège, tous les pays retenus dépassent nettement le pourcentage théorique du niveau 2. Le niveau 3 se révèle beaucoup plus discriminant, notamment pour la France, dont le pourcentage devient similaire à celui de la population générale alors qu'aux niveaux 1 et 2, elle s'en écartait résolument. Cet effet se confirme au niveau 4. Par rapport à l'Angleterre, la Suède ou encore les États-Unis, le contraste est saisissant et montre à quel point la réalisation d'inférences et la capacité interprétative varient selon les pays pour des élèves effectuant une quatrième année d'école obligatoire.

En somme, seul le quart des élèves français de CM1 testés ici dépassent le niveau 2, et disposent de bonnes capacités inférentielles et interprétatives. Seule la Norvège obtient un pourcentage inférieur. Ce pourcentage s'élève à 45 % en Grande Bretagne et à 41 % aux États-Unis, en Ontario et en Suède.

Tableau 1. – Répartition des pourcentages d'élèves selon le score moyen et le niveau de compréhension atteint à PIRLS, en fonction du pays.

	Score moyen	Niveau 4 615 et +	Niveau 3 570 et +	Niveau 2 510 et +	Niveau 1 435 et +
Pourcentages théoriques	500	10	25	50	75
Angleterre	553	24	45	72	90
États-Unis	542	19	41	68	89
Nouvelle-Zélande	529	17	35	62	84
Canada (Ontario)	548	19	40	70	92
Canada (Québec)	537	11	31	67	94
France	525	9	26	60	90
Allemagne	539	12	34	69	93
Italie	541	14	36	69	92
Norvège	499	6	19	48	80
Suède	561	20	47	80	96

Pour PISA, les performances aux trois « types de tâches de lecture » sont exprimées par les trois échelles spécifiques : « S'informer », « Interpréter », « Réfléchir », elles-mêmes divisées en cinq niveaux. Les français retrouvent plus facilement l'information qu'ils ne développent une interprétation du texte, et ils sont significativement plus nombreux à atteindre le niveau 5 à l'échelle « S'informer » que la moyenne internationale. Leurs performances chutent pour l'échelle « Réfléchir » et passent même en dessous de la moyenne internationale. Les meilleurs performances des filles sont vérifiées pour chacune de ces échelles et l'on note qu'à l'échelle « Réfléchir », les résultats des garçons diminuent davantage que ceux des filles, marquant un effet différentiel de ce type de tâche.

Chaque « type de tâches de lecture » est représenté à chacun des cinq niveaux de difficultés définis pour PISA. Les tâches regroupées sous le vocable « Interpréter » font davantage appel aux capacités inférentielles que « S'informer ». La tâche intitulée « Réfléchir » ou parfois « Réagir » évalue la capacité des élèves à faire appel à leur expérience personnelle, à leurs connaissances et à leurs idées pour les mettre au service d'un texte ; bien évidemment les capacités inférentielles et interprétatives sont convoquées, même si les caractérisations des tâches ne les mentionnent pas. Nous voyons que comme dans PIRLS, nos élèves rencontrent davantage de difficultés dès lors qu'ils doivent exercer ces capacités et que le malaise s'accroît avec « Réfléchir ». Ils manifestent, en revanche, une certaine aisance pour « Retrouver » de l'information dans PISA, mais les données disponibles ne permettent pas de le vérifier pour PIRLS.

En effet, pour PIRLS, les données concernant « les processus de compréhension », seconde dimension organisant le protocole n'ont été explorées qu'à la suite de l'insistance de pays participants désireux d'établir des comparaisons avec PISA (Mullis *et al.*, 2004). Pour des raisons méthodologiques, seules deux échelles de compréhension ont été bâties en regroupant, à chaque fois, deux catégories de « processus ; elles s'appellent « Prélever/Inférer » et « Interpréter/Apprécier ».

Un écueil perceptible dès le CM1, la « réflexion critique »

Les résultats aux échelles spécifiques de PIRLS et de PISA ont été rapprochés. Pour PISA, les pays anglophones, à l'exception de la Nouvelle-Zélande, ont des performances plus élevées à l'échelle « Réfléchir » qu'aux échelles « S'informer » et « Interpréter »,

et pour PIRLS, tous les pays anglophones ont une performance significativement meilleure à l'échelle « Interpréter-Apprécier ».

Cette constance dans la supériorité des anglophones lorsque le Processus évalué met en jeu la dimension « réflexive » (*reflect on*), mériterait d'être comprise. En nous échappant des limites de notre échantillon, signalons également un effet semblable pour cinq autres pays (Brésil, Espagne, Grèce, Mexique et Portugal), quatre d'entre eux étant hispanophones ou lusophones. L'OCDE (2003, p. 110) écrit : « certains profils de répartition semblent aller de pair avec des groupes linguistiques différents [...] Les pays où l'anglais n'est pas la langue d'enseignement présentent un tout autre profil [...] La variation de ces profils de performances peut s'expliquer par des pratiques pédagogiques différentes », avis partagé dans le cadre de PIRLS par I. V. S. Mullis *et al.* (2004) en ajoutant une perspective « liée aux programmes ». Dans une étude complémentaire sur PIRLS, M. Sainsbury *et al.* (2004) estiment que « parmi les différences internes aux pays, voire aux langues, il convient de remarquer le cas des pays anglophones... ».

L'hypothèse sous-jacente sur le poids des langues peut être émise sans encore avoir été l'objet de validation solide apportant un éclairage différent sur les performances. En revanche, quelques avancées peuvent être trouvées lorsqu'on s'intéresse aux programmes scolaires des pays participant à PIRLS. On l'a indiqué en introduction, les concepteurs de PIRLS voulaient explorer les liens potentiels entre le niveau de compréhension des élèves, le programme d'enseignement et les pratiques pédagogiques concernant la lecture. Aussi ont-ils collecté des données sur les systèmes éducatifs qu'ils ont rassemblées dans « PIRLS 2001 Encyclopaedia » (Mullis *et al.*, 2002). Le cœur du chapitre dévolu à chaque pays concerne les programmes et les politiques d'enseignement de la lecture et s'achève sur les pratiques évaluatives.

La consultation de cet ouvrage révèle que *reflect on* est inscrit dans les programmes d'enseignement de la lecture (*Reading Curriculum and Standards*) de tous les pays de langue anglaise, ainsi que de la Suède, alors qu'aucune référence n'y est faite dans ceux des autres pays de l'échantillon. Au-delà des programmes, « la réflexion critique » fait l'objet d'évaluations dans des dispositifs centralisés, dont voici deux exemples. Dans les évaluations britanniques – *National Curriculum Reading Test in England* (NCRT) –, des compétences liées aux activités « réflexives » sont évaluées tous les ans, dès l'âge de sept ans – fin du

Key-Stage (KS 1), étapes clés (6). Aux États-Unis, la « réflexion critique » est évaluée en grade 4 (vers dix ans) par la NAEP (*National Assessment of Educational Progress*).

Certains pays valorisent donc la dimension « Réfléchir » au niveau de l'enseignement scolaire, ce qui se reflète visiblement dans leurs performances aux évaluations internationales, un pôle linguistique lié à la langue maternelle anglaise étant facilement identifiable. Le pôle « Pratiques pédagogiques » a été exploré au moyen de questionnaires aux enseignants dont on rendra compte.

Les faiblesses des Français confirmées par des évaluations nationales

Les deux évaluations internationales offrent un éclairage nouveau sur les capacités de nos élèves, et les données qu'elles apportent peuvent être mises en perspective avec deux évaluations-bilans menées par la DEP en 2003 en fin de CM 2 et de classe de 3^e (cf. France, 2004a & 2004b).

Même si les deux évaluations présentent de nombreuses différences, les résultats français à PIRLS peuvent être rapprochés de ceux de l'évaluation intitulée « La maîtrise du langage et de la langue française en fin d'école primaire ». Celle-ci montre qu'un moins d'un tiers des élèves dispose des compétences attendues en fin de cycle 3 en matière de maîtrise de la langue. Ces élèves se caractérisent notamment par leurs capacités d'analyse et de synthèse, capacités sollicitées dans la réalisation d'inférences et l'interprétation des textes. PIRLS a montré que seulement 26 % des élèves de CM1 disposent de bonnes capacités inférentielles et interprétatives. Ces deux évaluations, nationale et internationale mettent donc l'accent sur la proportion limitée de nos élèves capables de mettre en oeuvre certains processus en jeu dans la compréhension de l'écrit.

Environ 35 % d'élèves participant à PISA sont scolarisés en classe de 3^e (et 50 % en classe de 2^{de}), ce qui permet de rapprocher les résultats de PISA de ceux de l'évaluation de fin de collège. Cette dernière met en évidence qu'environ la moitié des élèves encore scolarisés en classe de 3^e, ayant donc redoublé une fois au cours de leur *cursus* antérieur, n'est pas en mesure d'effectuer des inférences simples ; cette incapacité à se décentrer pour élaborer une inférence se manifeste très probablement dans PISA où une grande partie des items, à des niveaux divers, requiert ce type d'opération mentale (pour une analyse, voir Rémond, 2005).

DES COMPORTEMENTS RÉVÉLATEURS

Des relations difficiles à la production de réponses écrites

Les QCM sont nettement mieux réussis que les items demandant de rédiger une réponse, résultat assez classique et partagé par les pays participants à PIRLS ; cet effet est également constaté pour PISA mais seulement pour les pays francophones (Soussi *et al.*, 2004). Ni PIRLS, ni PISA ne visent à évaluer la production d'écrit alors que les réponses passent souvent par le canal de l'écrit, même s'il est explicitement indiqué dans PIRLS que l'écrit ne fait pas l'objet d'évaluation (Mullis *et al.*, 2003, p. 51). Plus que ceux d'autres pays, les jeunes français ne répondent pas à un certain nombre d'items pour lesquels il fallait rédiger une réponse, ce comportement s'observe dans les deux programmes.

Le pourcentage de non-réponses est spectaculairement plus élevé pour l'échelle « Réfléchir » que pour les échelles « S'informer » et « Interpréter » mais, contrairement aux autres échelles, les questions à réponse construite y sont quatre fois plus nombreuses que les QCM. Pour cette dimension, les élèves sont donc confrontés, à la fois, à un format de réponse requérant majoritairement le passage à l'écrit et à une tâche peu habituelle dans leur système scolaire.

La passation de PISA a dévoilé la réalité scolaire de la population française de quinze ans. Il s'agit d'une population dispersée depuis la classe de 4^e jusqu'à celle de 1^{re} : sans entrer dans les détails, en 2000 environ 50 % des élèves participant à l'enquête étaient scolarisés en 2^{de}, 36 % en classe de 3^e, 10 % en classe de quatrième. On remarque que la fréquence de leurs non-réponses apparaît très corrélée avec le niveau atteint dans les échelles de compétences et le retard scolaire, autrement dit, les élèves les plus en retard sont aussi ceux qui ont les performances les plus faibles et qui produisent le plus de non-réponses aux questions ouvertes.

Dans PIRLS, pour tous les processus évalués, nos élèves manifestent leurs difficultés face aux questions qui impliquent de construire une réponse consignée par écrit. Quand la tâche de rédaction est limitée et s'apparente à un prélèvement d'information, les performances deviennent supérieures à celles de la moyenne des pays, même si des opérations cognitives complexes sont en jeu. Il en est de même lorsque la question offre une « trame de réponse », en voici un exemple.

Cet item nécessite « d'interpréter et d'intégrer des idées » (texte *Le lièvre*, item H 9) :

« Les sentiments du lièvre changent pendant l'histoire.

« Complète les phrases suivantes :

- « Au début de l'histoire, le lièvre se sent...

« Parce que...

- « À la fin de l'histoire, le lièvre se sent...

« Parce que... »

Dans ce cas, nos élèves parviennent à montrer leur compréhension des relations de causalité entre deux événements alors qu'en l'absence de ce guidage, ils n'apportent pas la justification demandée dans la question et se contentent de donner une réponse partielle. Lorsque les étapes du raisonnement sont présentes dans l'ébauche de réponse, ils gèrent efficacement la tâche. Certaines tâches réputées complexes peuvent donc être accomplies avec succès si elles font l'objet d'aménagements dans leur présentation ; le coût cognitif de la mise en mots est alors réduit. Cette fluctuation des performances en fonction de la structuration des items semble indiquer leur très grande dépendance à l'égard des habitudes scolaires et un manque certain d'autonomie dans le traitement des tâches sous-jacentes, qui s'explique, entre autres, par des défauts métacognitifs.

Une gestion insuffisante des tâches

Une représentation rigide des tâches de lecture

Les élèves de CM1 ont trop souvent cherché à établir une correspondance directe entre la question et le texte, ce qui a pu les conduire dans des impasses. Certains ont persévéré dans l'application de cette conduite et recopié directement des éléments du texte alors qu'il fallait mener des opérations cognitives plus complexes et aller au-delà du texte. Leur mise en correspondance, assez directe et systématique, entre des éléments de la question et du texte amène à s'interroger sur les modèles de tâches qu'ils rencontrent dans leur vie de classe. Comme cet appariement direct constitue le mode dominant de leurs activités de lecture, certains ont construit une représentation restrictive et erronée des tâches de lecture dont ils ne s'éloignent pas, contrairement à d'autres élèves. Notre hypothèse est confirmée par les réponses des enseignants aux questionnaires de PIRLS relatées dans cet article, ainsi par les constats convergents sur la pauvreté du questionnement mené en classe (Beltrami *et al.*, 2004 ; Gaonac'h &

Fayol, 2003 ; Guernier, 1999 ; Rémond, 2003 & 2005 ; Tauveron, 2002), ou encore par l'examen de manuels récents de français pour le Cycle 3 (Nora, 2006).

Un contrôle métacognitif insuffisant

La représentation erronée de l'activité de lecture a été identifiée comme étant une caractéristique des faibles lecteurs/compreneurs ; ceux-ci disposent de moins de connaissances métacognitives que les bons lecteurs/compreneurs (Rémond, 2003). Des défaillances dans le contrôle métacognitif, se manifestant par des défauts de gestion de la tâche, ont été mises en évidence par les deux programmes.

La présomption d'un traitement partiel des consignes vaut pour les deux programmes. Certains élèves semblent s'être engagés dans la tâche sans avoir pris en considération les paramètres précisant la requête, ce qui peut les avoir conduits à choisir dans un QCM un distracteur fortement plausible faute d'avoir mené un traitement suffisant de l'information présentée.

La gestion défaillante de la consigne, et donc de la tâche, se traduit aussi d'autres manières. Ainsi dans PIRLS, où la plupart des histoires de l'objectif « Littéraire » comportaient deux personnages principaux, nous avons mis en évidence des difficultés de focalisation sur « le bon personnage », celui qui fait l'objet de la question. Des élèves continuent à raisonner sur le personnage sur lequel portait la question précédente. Cette difficulté « d'inhibition » d'un personnage se traduit aussi lors de questions dans lesquelles les deux personnages sont mentionnés, mais où il faut comprendre pour lequel les deux informations doivent être données. En voici un exemple particulièrement riche : « *Vers la fin de l'histoire, le lion essaie de rassurer le lièvre. Donne **deux** exemples de la façon dont il s'y prend* ». Certains élèves répondent en prenant un exemple pour chacun des deux personnages mentionnés dans la consigne. Il n'est pas impossible que par une stratégie de distance minimale, le pronom « il » ait été interprété comme le substitut du personnage « le lièvre ». Ce traitement défectueux d'une marque anaphorique met en lumière un problème introduit par la traduction, alors que les évaluations internationales ne portent pas sur le traitement des marques anaphoriques. La version anglaise centre sans ambiguïté la question sur le lion : « *Write **two** ways in which the lion tried to make the hare feel better at the end of the story* » Cet exemple démontre remarquablement la complexité des phénomènes que la non-équivalence de la consigne fait émerger, et l'effet du non-respect, dans la version française

de la « règle de l'ordre d'importance ». Selon cette « règle », dans la question, il faut hiérarchiser les informations présentées et indiquer en premier lieu ce qui doit être fait, ceci permettant au lecteur de mieux cibler sa tâche.

De manière répétée, au travers des réponses erronées, nous avons détecté des défauts métacognitifs. Le manque d'évaluation de la pertinence de la réponse par rapport à la question et/ou au texte est récurrent.

Un manque de flexibilité

Dans les deux programmes, les jeunes français se montrent très dépendants de stéréotypes de réponses, et donc peu préparés aux diverses exigences des rencontres avec l'écrit au sens défini par PISA. Ils s'interrogent beaucoup sur leur interprétation des tâches dès que celles-ci s'écartent de leurs habitudes scolaires (par exemple, juger un point de vue, faire abstraction de son opinion). Ils vont même jusqu'à ne pas répondre lorsqu'ils doutent de leur réponse ou qu'ils ne comprennent pas suffisamment la demande. Quand ils ne reconnaissent pas leurs objets habituels d'investissement scolaire, ils essaient, coûte que coûte, d'établir des ponts avec leur culture scolaire, en particulier, celle du cours de français. Lors d'entretiens que nous avons menés après la passation des protocoles, certains nous ont signalé des tâches auxquelles ils n'ont pas répondu en raison d'une question inhabituelle leur demandant, par exemple, d'exprimer leur opinion personnelle. Faut-il chercher un lien avec ce qui pourrait être interprété comme une peur de l'erreur se manifestant très tôt (pour une discussion, voir Rémond, 2005) ?

LES PRATIQUES PÉDAGOGIQUES AUXQUELLES SONT EXPOSÉS LES ÉLÈVES DE PIRLS

Les maîtres en charge des élèves évalués par PIRLS ont rempli un questionnaire composé d'un ensemble de quarante quatre questions dont certaines ont trait à la maîtrise de la langue. Ils ont dû indiquer la fréquence à laquelle ils font pratiquer un certain nombre d'activités visant le développement des capacités et des stratégies de compréhension. Les résultats sont rapportés en prenant l'élève comme unité, et le référent retenu comme unité d'étude est « *activité pratiquée au moins une fois par semaine* » (7).

En moyenne, « *l'identification de l'idée principale* » et « *l'explication/argumentation de ce qui est*

compris » représentent les activités les plus pratiquées : 90 % des élèves sont confrontés à ces activités au moins une fois par semaine contre 80 % en France. Deux autres activités sont assez communément proposées : « *comparer : confronter ce qui a été lu à son expérience personnelle* » (73 %) et « *effectuer des inférences et opérer des généralisations* » (71 %) ; ces deux activités ne sont pratiquées que par, respectivement, 31 % et 23 % des élèves français. Les activités « *d'anticipation de la suite du texte* » sont exercées par 61 % des élèves (contra 40 % en France), « *la comparaison de ce qui a été lu avec des lectures antérieures* » par 59 % des élèves (contra 23 % des élèves français), la « *description du style ou de la structure du texte* » par 52 % des élèves (contra 34 % de nos élèves).

Le temps qui est déclaré être dévolu à l'enseignement de la lecture est bien inférieur, en France, à la moyenne internationale et contrairement à d'autres pays, nos maîtres adaptent rarement le matériel en fonction du profil de leurs élèves. Enfin, bien qu'en France, les supports liés à d'autres disciplines soient déclarés être fréquemment des vecteurs d'activités relatives à la lecture, l'enseignement de la lecture dans toutes les disciplines reste très marginal (5 % contra 21 % à l'international).

Les maîtres français font lire à haute voix bien plus souvent que leurs collègues étrangers. Une exploration des tâches confiées aux élèves après la lecture a été menée par questionnaire. Elle montre que les élèves français répondent souvent à des « *questions orales après la lecture* », mais qu'ils reçoivent rarement et moins que la moyenne internationale des questions de compréhension nécessitant une réponse exprimée par écrit. Seuls 18 % (contra 68 %) ont l'occasion d'écrire « *un texte qui fait suite ou qui répond à ce qu'ils ont lu* », 32 % (contra 65 %) discutent « *entre eux de ce qu'ils ont lu* ». Ces trois tâches appartiennent à une liste de six tâches qui comporte aussi « *faire un dessin ou réaliser un projet artistique lié à ce qui a été lu* », « *présenter un sketch ou mettre en scène ce qui a été lu* » et « *réaliser un travail de groupe sur ce qui a été lu* ». Ces trois dernières tâches sont presque inexistantes dans nos écoles. Un indicateur a été calculé pour rendre compte de la variété des tâches, il indique la proportion d'élèves par pays qui se voient proposer, chaque semaine, au moins trois tâches parmi les six énumérées ci-dessus : seuls 17 % de nos élèves de CM1 sont dans ce cas alors que 59 % de leurs pairs participant à PIRLS réalisent au moins trois activités différentes sur les textes chaque semaine.

Les données qui viennent d'être présentées émanent de connaissances déclaratives fournies par les maîtres et non d'observations des pratiques de classes. Elles doivent bien sûr être interprétées prudemment, d'autant que l'on peut objecter des effets culturels ou pédagogiques et avancer que nos élèves rencontrent d'autres tâches au moins aussi formatives en matière de lecture. Cependant, elles contribuent à éclairer certains résultats. Ainsi, on note que la Suède, l'Angleterre, les États-Unis ou encore la Nouvelle-Zélande ont fréquemment recours à du travail écrit en lien avec la lecture ; ce sont également des pays où la « réflexion critique » est inscrite dans les programmes.

La différence d'empan dans les tâches proposées s'observe aussi dans les « types de texte » utilisés au cours des activités relatives à la lecture. La « fiction » est travaillée hebdomadairement en CM1 par quatre élèves sur cinq, nos élèves étant plus en prise avec des « livres divisés en chapitres » qu'avec des fables et des contes qui, contrairement à la moyenne des pays, demeurent largement absents des supports de lecture. Le faible usage des écrits « non-fictionnels » (« *non-fiction* » en anglais) caractérise également la situation française.

LES EFFETS INTER-LANGUES

On a déjà pu faire quelques remarques au sujet des effets inter-langues, nous allons en ajouter quelques autres en nous penchant davantage sur le matériel, car celui-ci fixe et conditionne la tâche.

On s'attendait à ce que des profils similaires de réponses aux quatre modalités d'un même QCM se dessinent, sans effet de langue ; mais de temps en temps, on observe des *patterns* de réponses différents selon les langues aux quatre modalités de réponse d'un même QCM. Ce phénomène est particulièrement facile à observer dans le cas du Canada où l'on constate pour nombre d'items un écart significatif et parfois important entre les résultats des deux provinces canadiennes, les résultats de l'Ontario étant proches de ceux des pays anglophones, ceux du Québec entretenant une forte similitude avec les résultats français. La plausibilité des distracteurs peut évoluer lors du passage à une autre langue et dans les deux programmes, on a pu observer que certains distracteurs, qui ne l'étaient pas en anglais, sont devenus parfaitement plausibles en français.

Tout ceci ramène à l'éternelle question de l'équivalence des épreuves dans un contexte international.

L'information a parfois subi une structuration bien différente de la version source en anglais, qui a souvent, mais pas toujours, complexifié la tâche de nos élèves. Les traductions ne tiennent pas forcément compte des règles d'énonciation qui devraient régir les consignes des matériels d'évaluation, en particulier de celle qui pourrait être baptisée « règle de l'ordre d'importance » évoquée précédemment (Rémond, 2001 & 2005). Enfin, les consignes de codage ne gèrent pas toujours les subtilités de la langue vers laquelle elles ont été traduites. Tout particulièrement dans PISA, la structuration de certains énoncés ne correspond pas à nos habitudes rédactionnelles. La consigne peut indiquer plus précisément, dans une langue, la tâche à accomplir, ce qui explique alors une bien meilleure performance moyenne dans cette langue de test et, de manière corollaire, moins de réponses erronées et de non réponse. La question de la précision de l'information se pose parfois en termes d'équivalence du registre du lexique employé et donc de familiarité avec celui-ci.

Pour un certain nombre d'items de PIRLS, nous l'avons déjà souligné, la très nette réussite de la plupart des pays anglophones fait s'interroger sur la traduction en français, et parfois aussi dans d'autres langues.

Voici l'exemple d'un item (R 5) extrait de l'épreuve « Promenade le long de la rivière » dans sa version source en anglais : *How much would it cost a 10-year-old child to hire a bike for a full day ?* Cet item est devenu en français : « Un enfant de 10 ans veut louer un vélo pendant toute une journée : combien devra-t-il payer ? »

Ce item, qui prend ainsi la forme d'un énoncé de problème, est remarquablement réussi par les francophones et par la Suède (90 %), les anglophones atteignant une moyenne de 81 %, l'Italie et la Norvège environ 75 %.

Il faut donc accroître la vigilance lorsqu'on vérifie les traductions du matériel afin de ne pas changer la nature de ce qui est estimé d'une langue à l'autre, et donc ce qu'on évalue ; cette opinion est partagée par M. Sainsbury *et al.* (2004) à la suite d'un travail comparatif sur les profils de résultats dans plusieurs langues, déjà cité, et mettant en évidence la supériorité des anglophones à divers items.

Nous avons montré par exemple que les questions de PISA s'appuyant sur le « style » ne semblent pas

avoir le même sens pour un francophone et un anglophone, car manifestement, dans ces langues de référence, le mot « style » n'évoque pas le même concept. Dans PISA, évaluer « le style » ne revient pas à utiliser un savoir transmis par le cours de langue maternelle, mais à analyser, critiquer, s'approprier le texte.

Voici un exemple commenté portant sur le support de l'exercice « Grippe », un communiqué s'adressant à des salariés pour les inciter à se faire vacciner. La mise en forme, deux dessins et un slogan pour clore le communiqué donnent un caractère attractif à ce document qui se présentait à l'origine sous la forme d'un dépliant. Une question cible le « style ». La tâche requise nécessite d'évaluer la pertinence de la composition du texte et son adéquation au public visé. Repérer que le texte a été mis en page de manière attrayante et que sa tonalité est encourageante constituent une réponse correcte. L'essentiel relève de la présentation du communiqué (illustrations, typographie, mise en gras, sous-titres), ce qui ne relève pas du style, au moins, dans notre culture scolaire ; aussi le terme « style » employé dans la question ne peut que déstabiliser nos élèves (pour d'autres exemples, Rémond, 2005). Ces questions appartiennent à l'échelle « Réfléchir », si difficile pour nos élèves.

Enfin et pour prolonger la question de l'évaluation de la lecture littéraire, on a détecté des élèves qui produisent une réponse révélatrice d'une interprétation concurrente de celle qui est prévue par le codage international. Dans PIRLS, ces élèves se montrent très performants à l'ensemble de cette évaluation, faisant preuve d'un réel travail interprétatif, mais ils vont au delà de la programmation linéaire des questions, selon le déroulement chronologique de l'histoire. Le problème de la négociation de la marge d'interprétation a déjà été posé pour PISA (Rémond, 2004 & 2005). Plusieurs « bonnes réponses » à une même question pourraient être recevables, mais dans les deux programmes, des codages sélectifs ignorent des pans d'interprétation tout à fait plausibles ; ce fait se rattache à cette délicate question de l'évaluation de la lecture littéraire (Sainsbury & Watts, 2006 ; Sève, 2004).

COMMENT FAIRE PROGRESSER LES ÉLÈVES FRANÇAIS ?

Les deux évaluations internationales cibles de cet article offrent un éclairage nouveau sur les capaci-

tés de nos élèves en compréhension de l'écrit et sur leur pouvoir d'adaptation. Elles ouvrent des pistes de réflexion pour envisager de les faire évoluer, sans nécessairement chercher à les « conditionner » à ces types d'évaluation. Ainsi dans la mesure où les capacités inférentielles revêtent une grande importance dans le traitement de la langue orale et écrite, les constats que nous avons dressés, indiquent la grande vigilance que doit avoir l'école à leur égard.

S'interroger sur l'efficacité des pratiques

L'exploitation des questionnaires met en évidence un certain nombre d'éléments propres à la France en termes d'exposition à l'écrit : un temps consacré à l'enseignement de la lecture inférieur à la moyenne internationale, une adaptation peu fréquente du matériel en fonction du profil des élèves, beaucoup de questions orales après la lecture et peu de questions nécessitant une réponse exprimée par écrit, de rares échanges verbaux sur la compréhension des textes.

Les élèves français seraient moins exposés à des activités de nature à développer la compréhension que la moyenne de leurs pairs évalués par PIRLS : activités peu variées sur des supports eux-aussi peu variés. Ce « déficit » est particulièrement visible pour le travail de nature inférentielle et pour les activités de mise en perspective avec des lectures antérieures ou avec son expérience personnelle, désormais valorisées par les programmes français de 2002. Ces éléments contribuent à expliquer des performances révélées par les évaluations nationales ou internationales. Si les capacités inférentielles semblent faire défaut à nombre de nos élèves, ce serait parce que ceux-ci pratiquent très rarement des activités de nature à les développer. Il s'agit pourtant d'un champ crucial dans la construction de la signification du texte.

On observe que le traitement des aspects littéraires des textes estimé par les items de PIRLS est moins efficace chez les Français que pour la moyenne des pays. Ces items qui contribuent essentiellement aux niveaux de compétences les plus élevés font appel au point de vue, à la reconnaissance de techniques narratives, ou encore à l'évaluation de la plausibilité d'une histoire... Les maîtres déclarent travailler de manière prépondérante sur des textes de fiction, mais ils n'ont peut-être pas fait exercer ces types de traitements par leurs élèves.

Il faut continuer à faire porter les efforts sur la maîtrise de la langue en classe bien sûr, mais aussi dans

la formation des enseignants. Selon les questionnaires de PIRLS, certains secteurs sont beaucoup moins investis dans la formation initiale ou continue des enseignants français que dans la moyenne des pays participant à PIRLS : il s'agit, entre autres, de l'enseignement et des théories de la lecture, du développement du langage, de la psychologie.

Beaucoup de choses à l'école restent dans l'implicite alors que des travaux récents de psychologie « invitent à penser l'articulation entre apprentissage implicite et enseignement scolaire davantage en termes de complémentarité qu'en termes de substitution » (Perruchet & Pacton, 2004, p. 135).

Apprendre les élèves à comprendre

Dans certains programmes d'enseignement de la compréhension, il est conseillé de « rendre les processus transparents » au lecteur (Gaonac'h & Fayol, 2003), au lieu de laisser jouer l'implicite.

Pour que les élèves acquièrent, de manière graduelle, une certaine expertise et une efficacité face aux tâches de lecture, ils doivent apprendre à gérer leur compréhension et tout d'abord avoir conscience de l'attention conjointe qu'il faut apporter au texte et aux questions, par exemple dans les protocoles d'évaluation. La compréhension du texte est testée au travers de consignes qui n'ont pas toutes le même poids. Il faut donc leur apprendre à les lire et à les analyser, à se représenter la nature de l'information requise et à la situer par rapport au contenu du texte, qui devra nécessairement être lu. Ils doivent recevoir des questions variées pour découvrir les degrés de complexité du traitement à mener et dépasser le simple prélèvement d'informations ; ceci les amène à envisager qu'il faut produire des inférences pour relier les informations de la question et du texte. Leur conception tronçonnée des tâches les empêche de mobiliser des connaissances variées qu'ils doivent apprendre à activer (Beltrami *et al.*, 2004). Enfin, ils doivent évaluer la qualité et la pertinence de leur réponse, et éventuellement la modifier. En résumé, les élèves doivent se représenter la tâche à accomplir comme un fonctionnement en boucle qui oblige à examiner les relations réciproques entre les sommets du triangle : « question-texte-réponse ». Cette perspective est hautement fondée sur la métacognition et donc sur la régulation des apprentissages (Rémond, 2003). Ce raisonnement sur l'écrit doit être exercé le plus tôt possible pour que les élèves puissent construire la représentation mentale du contenu du texte la plus exacte possible. Les moda-

lités d'expression de la réponse doivent être elles-aussi travaillées afin de dépasser les constats dressés sur la production parcimonieuse de réponses écrites.

Les travaux de recherches soutenant ces perspectives se sont installés dans notre paysage, les derniers programmes de l'école parus en 2002 les ont pris en compte et le PIREF pour la conférence de consensus sur l'enseignement de la lecture (2003) avait demandé à Francis Grossmann d'établir une synthèse qui s'intitule : *Comment faciliter, développer et évaluer la compréhension des textes aux différentes étapes de la scolarité primaire*. Dans cette synthèse, l'auteur adopte deux points de vue, celui du linguiste et celui du didacticien. Dans une première partie ; il tente d'expliquer pourquoi les élèves ne comprennent pas ce qu'ils lisent et ce qui l'amène à définir des pistes explicatives, certaines « *pouvant être mises en relation... avec le développement des capacités interprétatives* ». Il analyse aussi les spécificités de la lecture scolaire et de « la compréhension de textes » à l'école, ou encore « les contraintes sémiographiques et la compréhension de la tâche » et « la difficulté d'articuler des objectifs différents ». Sans entrer dans une énumération inutile des entrées du document, on perçoit leur proximité avec certaines rubriques traitées dans cet article, montrant le manque de nouveauté de certains constats faits lors des évaluations internationales. La seconde partie de cette synthèse offre des propositions didactiques s'articulant sur les points de sa première analyse. Parmi celles-ci, on note : « mieux inscrire le temps de lecture dans le temps scolaire », « développer les capacités interprétatives » et « définir la nature et la place d'un entraînement à la compréhension et à l'interprétation », « mieux définir les règles de la lecture scolaire et améliorer la rédaction des questions », « diversifier et mieux intégrer l'évaluation... ».

Les missions du PIREF se sont arrêtées et ce travail n'a pas suffisamment été diffusé dans notre communauté. Il n'est pas trop tard pour en tirer profit et rendre nos élèves plus à l'aise en compréhension et en production d'écrit. Ils auront ainsi une meilleure estime d'eux-mêmes et moins peur de commettre des erreurs puisqu'ils pourront contrôler leur exécution des tâches.

Martine Rémond
martine.remond@inrp.fr

IUFM de Créteil
UMR Éducation et politiques (INRP – Université Lyon 2)

NOTES

- (1) Ces dispositifs sont détaillés dans M.-T. Céard *et al.* (2003) et dans des publications de la direction de l'évaluation et de la prospective (DEP) du ministère de l'Éducation nationale qui, en France, a mis ces évaluations en œuvre. Les sites internet de l'OCDE et de l'IEA peuvent aussi être consultés : [http://www.pisa.oecd.org] et [http://www.isc.bc.edu/pirls2001.html].
- (2) Dans la première partie de cet article, nous nous référons à cette publication dont nous avons traduit de larges extraits.
- (3) Nous avons participé à ces deux dispositifs en tant que membre du groupe d'experts internationaux.
- (4) *Reading literacy* est traduit par « culture de l'écrit » ou par « compréhension de l'écrit », dans les documents PISA-OCDE.
- (5) Pour simplifier, nous parlerons désormais de l'objectif « Littéraire » et de l'objectif « Informatif ».
- (6) Pour des précisions, on peut consulter M.-T. Céard *et al.* (2003, p. 110-111) et M. Sainsbury *et al.*, (2006).
- (7) Dans cette partie, les éléments indiqués entre guillemets correspondent au libellé de modalités de réponses à renseigner. Les pourcentages en regard de ceux de la France sont ceux de l'ensemble de la population de PIRLS.

BIBLIOGRAPHIE

- BELTRAMI D. ; QUET F. ; RÉMOND M. & RUFFIER J. (2004). *Lectures pour le cycle 3. Enseigner la compréhension par le débat interprétatif*. Paris : Hatier.
- CAMPBELL J. *et al.* (2000). *Framework and Specifications for PIRLS Assessment 2001 : Progress in International Reading Literacy Study*. Chestnut Hill : Boston College : IEA.
- CÉARD M.-T. ; RÉMOND M. & VARIER M. (2003). *L'appréciation des compétences des élèves et des jeunes en lecture et écriture et l'évolution de ces compétences dans le temps*. Rapport établi à la demande du Haut Conseil de l'évaluation de l'école, n° 11, 147 p. Disponible au format PDF à l'adresse [http://cisad.adc.education.fr/hcee/publications-2003.html]; consulté le 6 décembre 2006.
- FRANCE : MINISTÈRE DE L'ÉDUCATION NATIONALE : DIRECTION DE L'ÉVALUATION ET DE LA PROSPECTIVE (2003). *Note d'information*, n° 03.22 : « Les élèves de CM1. Premiers résultats d'une évaluation internationale en lecture (PIRLS) », rédigée par M. Colmant & A. Mulliez.
- FRANCE : MINISTÈRE DE L'ÉDUCATION NATIONALE : DIRECTION DE L'ÉVALUATION ET DE LA PROSPECTIVE (2004). *Note d'information*, n° 04.09, « Les compétences générales des élèves en fin de collège », rédigée par L. Dauphin et B. Trosseille.
- FRANCE : MINISTÈRE DE L'ÉDUCATION NATIONALE : DIRECTION DE L'ÉVALUATION ET DE LA PROSPECTIVE (2004). *Note d'information*, n° 04.10, « La maîtrise de la langue en fin d'école primaire », rédigée par F. Gilbert, J.-M. Levasseur et J.-M. Pastor.
- GAONACH D. & FAYOL M. (éd.), (2003). *Aider les élèves à comprendre du texte au multimédia*. Paris : Hachette.
- GUERNIER M. C. (1999). « Interactions verbales et construction de sens. Lectures d'un texte narratif ». *Repères : recherches sur le français langue maternelle*, n° 19, p. 167-182.
- GROSSMANN F. (2003). *Comment faciliter, développer et évaluer la compréhension des textes aux différentes étapes de la scolarité primaire ?* Synthèse rédigée pour la Conférence de consensus sur l'enseignement de la lecture à l'école primaire. Paris: Programme incitatif de recherche sur l'éducation et la formation (PIREF) : Paris, 4-5 décembre.
- KIRSCH I. ; JUNGLEBLUT A. & MOSENTHAL P. B. (1998). « The Measurement of Adult Literacy ». In T. S. Murray, I. S. Kirsch & L. Jenkins, *Adult Literacy in OECD Countries : Technical Report on the First International Adult Literacy Survey*. Washington [D. C.] : US Department of Education : National Center for Education Statistics.
- MULLIS I. V. S. ; MARTIN M. O. ; GONZALES E. & KENNEDY A. (2003). *PIRLS 2001 International Report*. Boston : International Study Center.
- MULLIS I. V. S. ; MARTIN M. O. & GONZALES E. (2004). *International Achievement in the Processes of Reading Comprehension. Results from PIRLS 2001 in 35 Countries*. Chestnut Hill : Boston College.
- MULLIS I. V. S. ; MARTIN M. O. ; KENNEDY A. & FLAHERTY C. L. (2002). *PIRLS 2001 Encyclopedia. A Reference Guide to Reading Education in the Countries Participating in IEA's PIRLS*. Boston : International Study Center.
- NORA E. (2006). « Le manuel scolaire au cycle 3, un outil d'apprentissage de la compréhension ». Communication au colloque international *Les manuels scolaires* : Montréal.
- OCDE (1999). *Measuring Student Knowledge and Skills : A New Framework for Assessment*. Paris : OCDE.
- OCDE (2003). *La lecture, moteur de changement : performances et engagement d'un pays à l'autre : résultats de PISA 2000*. Paris : OCDE.
- PERRUCHET P. & PACTON S. (2004). « Qu'apportent à la pédagogie les travaux de laboratoire sur l'apprentissage implicite ? » *L'Année psychologique*, vol. 104, p. 121-146.
- RÉMOND M. (2001). « Adapter n'est pas traduire : adaptation dans différents contextes culturels d'épreuves d'évaluation de la littéracie ». In C. Sabatier & P. Dasen (éd.), *Cultures, développement et éducation : autres enfants, autres écoles*. Paris : L'Harmattan, p. 171-184.
- RÉMOND M. (2003). « Enseigner à comprendre : les entraînements métacognitifs ». In D. Gaonach'h & M. Fayol (éd.), *Aider les élèves à comprendre du texte au multimédia*. Paris : Hachette, p. 205-232.
- RÉMOND M. (2005). « Regards croisés sur les évaluations institutionnelles », *Repères : recherches en didactique du français langue maternelle*, n° 32, p. 113-140.
- RÉMOND M. (soumis). « Que nous apprend PIRLS sur la compréhension des élèves français de 10 ans ? ». *Repères : recherches sur la didactique du français langue maternelle*.
- SAINSBURY M. ; SCHAGEN I. & HAMMOND P. (2004). *What did PIRLS Tell us About Reading Skills ?* Conférence de

- l'International Association for Educational Assessment* : Philadelphia, June.
- SAINSBURY M. & WATTS A. (2006). « Validity Challenges in a High-Stakes Context : National Curriculum Tests in England ». In M. Sainsbury, A. Watts, & C. Harrison (éd.), *Assessing Reading-Defining the Construct : From Theory to Classrooms*. The Mere : NFER.
- SÈVE P. (2004). « Peut-on évaluer la lecture littéraire ? » In *La lecture et la culture littéraire au cycle des approfondissements*. Versailles : CRDP-Scéren de l'académie de Versailles, p. 51-66.
- SOUSSI A. ; BROI A.-M. ; MOREAU J. & WIRTHNER M. (2004). *La littéracie dans quatre pays francophones : les résultats des jeunes de 15 ans en compréhension de l'écrit*. Neuchâtel : IRDP.
- TAUVERON C. [éd.] (2002). *Lire la littérature à l'école : pourquoi et comment conduire cet apprentissage spécifique ? De la GS au CM*. Paris : Hatier.
- VRIGNAUD P. (2006). « La mesure de la littéracie dans PISA : la méthodologie est la réponse, mais quelle était la question ? », *Revue française de pédagogie*, n° 157, p. 27-41.