

QUOI DE NEUF EN LITTÉRATIE ? REGARD SUR TRENTE ANS D'ÉVALUATION DE LA LECTURE¹

Dominique Lafontaine

Introduction

Comment a évolué, en l'espace de trente ans, la manière dont sont évaluées la lecture, la compréhension ou les compétences en lecture ? Pour répondre à cette question, j'ai choisi un poste d'observation un peu singulier : les grands programmes internationaux d'évaluation de la lecture - enquêtes de l'I.E.A. (Association internationale pour l'évaluation du rendement scolaire) ou de l'Ocdé (IALS, Pisa)². Ceux-ci peuvent sembler, à première vue, bien éloignés des pratiques de classe et des préoccupations didactiques. Cet article se centrera cependant sur des questions qui sont fondamentalement les mêmes que celles que soulève toute évaluation

collective de type sommatif, utilisant le papier-crayon et comportant un certain nombre de tâches ou d'items. Il laissera de côté les interrogations portant spécifiquement sur les contraintes propres aux évaluations internationales, telles que la comparabilité des données ou la problématique de l'équivalence culturelle.

1. Cadre de référence pour l'évaluation

Préalablement à toute évaluation formelle ou structurée, un cadre d'évaluation doit être défini. Ceci suppose que soit apportée une réponse notamment aux questions suivantes :

¹ Texte paru dans L. Collès, J.-L. Dufays, G. Fabry, C. Maeder (sous la direction de) (2001), *Didactique des langues romanes. Le développement de compétences chez l'apprenant. Actes du colloque de Louvain-la-Neuve*, janvier 2000, 67-82.

² Voir tableau descriptif des études examinées en annexe. Depuis le début des années 70, l'Association internationale pour l'évaluation du rendement scolaire (IEA) a conduit plusieurs enquêtes comparatives portant sur différentes disciplines (lecture, mathématiques, sciences, éducation civique,...). Plus récemment (1998), l'Ocdé s'est lancé dans le même type d'études, avec le programme PISA.

- Quel est l'objet de l'évaluation ? Que veut-on évaluer ? Quelle définition se donne-t-on de l'objet ?
- Quel est le lien entre l'évaluation et les apprentissages menés ou le curriculum ?
- Comment se structure le domaine que l'on veut évaluer ? Quels aspects, quelles dimensions, quelles compétences l'évaluation va-t-elle prendre en compte ?
- Quelles modalités d'évaluation utiliser ? Celles-ci sont-elles cohérentes avec les choix conceptuels posés ?
- Comment présenter et interpréter les résultats de l'évaluation ? Existe-t-il une compétence représentée par un score global sur une échelle ou les compétences sont-elles plurielles et appellent-elles plusieurs scores distincts ?

C'est à la lumière de ce cadre de référence que je me propose d'examiner l'évolution de l'évaluation de la lecture dans les surveys internationaux.

2. L'objet de l'évaluation : d'une étude séparée de la lecture-compréhension et de la littérature-interprétation à une étude de la littératie au sens large

En 1971, point de départ de mon étude historique, l'I.E.A. envisage pour la première fois une étude comparative portant sur la compréhension de textes. Un des aspects intéressants à noter est que

deux études séparées sont alors prévues : l'une sera consacrée à la compréhension en lecture (*reading comprehension education*), l'autre à la littérature. L'étude sur la compréhension se focalise sur les aspects cognitifs, tandis que l'étude « littéraire » s'ouvre davantage aux aspects esthétiques et affectifs et à la dimension réflexive ou critique, à l'interprétation. D'un point de vue théorique, ce divorce entre le cognitif et le culturel mérite d'être souligné. Deux facettes ou deux composantes de la lecture (le cognitif et le culturel, la compréhension et l'interprétation) sont ainsi distinguées, que le concept étendu de littératie réunira ultérieurement¹. L'étude sur la compréhension repose sur un modèle implicite du lecteur qui puise ou restitue correctement le sens unique déposé dans le texte (*gets meaning*), alors que dans l'étude « littéraire » affleure la référence au courant « *response to literature* », la réponse étant définie comme « l'interaction entre l'individu et l'oeuvre, interaction qui peut se prolonger longtemps après la lecture » (Purves, 1973 : 36).

¹ Pour Grossman (1999), « le concept de littératie permet d'articuler les aspects sociaux et culturels aux aspects développementaux » (p. 146). Il a « le mérite de bousculer des frontières trop communément admises, parce qu'elles correspondent à des partages institutionnels entre cognitif et culturel, mais aussi disciplinaires, entre psychologie, anthropologie et linguistique. » (p. 144). Dans le même esprit, Pierre (1992) souligne que le concept de littératie « couvre l'ensemble des habiletés, des comportements et des attitudes reliés à l'écrit. » (p. 3).

Au début des années 90, deux nouvelles enquêtes relatives aux compétences en lecture voient le jour. L'enquête I.E.A. Reading Literacy (1991) et l'International Adult Literacy Study (IALS, 1994). Ces deux études font, pour la première fois dans les études internationales d'évaluation de la lecture, référence de façon majeure au concept de « **littératie** ». Alors que la référence au curriculum et à l'école est dominante dans la première étude de l'I.E.A. (1971), l'étude IEARL s'ouvre à la notion de littératie fonctionnelle (avec une référence appuyée aux problèmes d'illettrisme¹) et adopte une définition de la « *reading literacy* » qui associe la compréhension et les usages sociaux de la lecture : « *La littératie est la capacité de comprendre et d'utiliser ces formes du langage écrit requises par la société ou importantes pour l'individu.* » (Elley, 1994 : 5, c'est nous qui soulignons)². Le recours au concept de littératie s'accompagne d'une entrée en force des « documents » (textes non continus) comme supports de l'évaluation -

cartes, graphiques, tableaux, formulaires - (le genre de textes qu'il faut savoir lire pour se débrouiller dans nos sociétés, et face auxquels l'incompétence signifie l'exclusion sociale), absents des évaluations antérieures. Mais, conséquence ou non de cette référence au concept de littératie, qui charrie toujours, ne serait-ce qu'implicitement, la notion de compétences de base ou de seuil minimal, l'étude IEARL s'arrête à la compréhension et aux usages de la lecture et ne s'ouvre pas à des aspects plus critiques ou réflexifs³.

La définition de la littératie (en anglais « *literacy* » dans ce cas, et non « *reading literacy* ») adoptée dans IALS est assez proche de celle de l'étude IEARL, même si l'accent y est plus nettement mis sur l'usage (le côté fonctionnel) que sur la compréhension, et ceci n'étonnera point, puisqu'il s'agit ici d'un public d'adultes. La littératie est ici définie comme :

¹ « Un adulte sur quatre de par le monde n'accède pas à la capacité de lire (...). De surcroît, dans un monde de plus en plus complexe, un monde de l'information, les exigences de littératie s'élèvent dans toutes les nations. De nos jours, les lecteurs peu compétents encourent de sérieux désavantages dans la plupart des pays. » (Elley, Schleicher, Wagemacher, 1994 : 1) (« *The ability to read is denied to one in four of the earth's adult population. (...) Furthermore in an increasingly complex, information-ridden world, demands on literacy continue to rise in all nations. Illiterates are at serious disadvantage in most countries today.* »).

² « *Reading literacy is the ability to understand and use those written language forms that are required by society and/or valued by the individual* » (Elley, 1994 : 5).

³ Les définitions adoptées par les évaluations successives ont été regroupées en annexe 2, afin de permettre une comparaison plus aisée. Il est à noter que l'ouverture sur les « usages » ne se fait que de façon indirecte ou médiatisée, via l'introduction de nouveaux types de textes (les documents « fonctionnels » ou « authentiques »). Quelles que soient les intentions, on n'ôtera pas le fait que le seul et « authentique » usage auquel l'individu est confronté dans ce contexte est celui d'une évaluation formelle, normative, longue, où est évaluée sa capacité à se débrouiller face à l'écrit seul, sans aide ou recours possible à des outils, et hors contexte social.

« *Utiliser l'information imprimée et écrite pour fonctionner en société, réaliser ses objectifs personnels, développer ses connaissances et ses potentialités*¹ » (Murray, Kirsch and Jenkins, 1998 : 17).

Les concepteurs de l'étude entendent rompre avec une vision simpliste de la littératie, opposant d'un côté ceux qui sauraient lire (les «alphabétisés»/«*literate*s») et ceux qui ne le sauraient pas (les analphabètes/*illiterate*s). Ils insistent sur le côté multidimensionnel de la lecture en délimitant trois domaines de littératie distincts : lire de la prose, lire des documents et lire des écrits comportant des données numériques (*prose literacy, document literacy and quantitative literacy*).

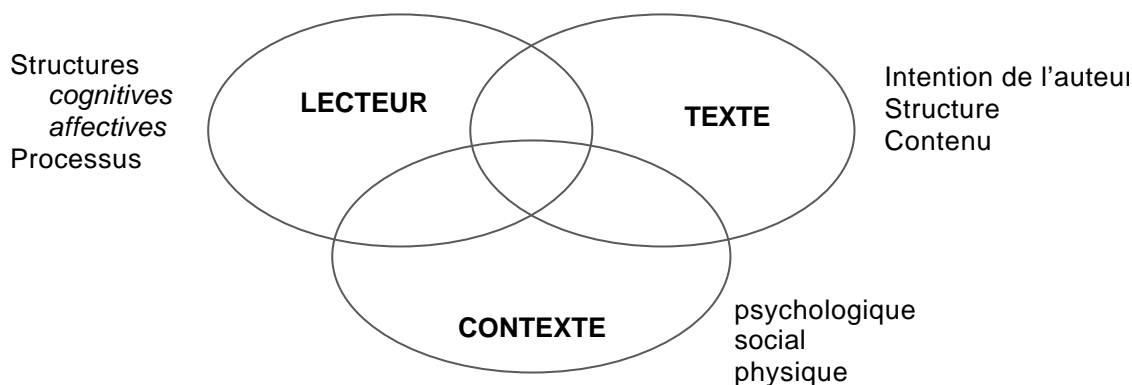
Enfin, à la fin des années 90, deux nouvelles études sont en chantier : l'étude Pisa (Programme international d'évaluation du suivi des acquis des élèves) de l'Ocdé et l'étude Pirls (*Progress in International Reading Literacy Study*) de l'I.E.A. Avec ces deux études une nouvelle étape semble franchie dans le domaine des évaluations internationales de la lecture. La conception et le développement des stimuli et des items sont désormais précédés de

l'élaboration d'un cadre conceptuel fouillé dans lequel sont détaillés les références théoriques, les choix concernant la structuration du domaine, les modalités d'évaluation, les échelles destinées à présenter les résultats, proposés par le groupe d'experts et soumis à l'examen des gestionnaires nationaux de projet (Ocdé, 1999; IEA/Pirls, 1999). Les études antérieures ont vraisemblablement élaboré de tels documents à usage interne; le fait que ces documents soient désormais destinés à être rendus publics et à faire l'objet de larges débats amène incontestablement à un effort de clarification et de planification accrus. Les études contemporaines et à venir seront ainsi davantage assises sur le plan théorique qu'elles ne l'ont été antérieurement où une approche plus empirique prédominait.

Simple coïncidence ou conséquence de cette emprise théorique plus grande, les définitions de la littératie adoptées dans Pisa et Pirls se montrent plus marquées par l'évolution des recherches fondamentales dans le domaine de la lecture, en faisant référence à une conception interactive de la lecture telle qu'illustrée dans ce schéma extrait de Giasson (1997).

¹ « *Using printed and written information to function in society, to achieve one's goals, and to develop one's knowledge and potential* » (Murray, Kirsch and Jenkins, 1998 : 17).

Modèle interactif de compréhension en lecture



Ainsi, dans Pisa, la littératie, c'est « *comprendre, utiliser des textes écrits, mais aussi réfléchir à leur propos, pour réaliser ses objectifs personnels, développer ses connaissances, ses potentialités, et participer à la vie en société*¹ » (Pisa, 1999 : 20, c'est nous qui soulignons). À côté de la compréhension et des usages, l'accent est mis sur la capacité de réflexion ou de distance critique du lecteur, qui joue un rôle actif, génératif dans l'élaboration du sens. La lecture est désormais envisagée comme un processus interactif entre un texte, un lecteur, un contexte : le lecteur, pour atteindre ses objectifs de lecture, déploie des stratégies et fait appel à ses connaissances antérieures (à sa culture, osons le dire) pour élaborer le sens et « répondre » au texte. La réconciliation entre compréhension et interprétation s'opère ainsi. L'ère de la réponse correcte et univoque semble désor-

mais révolue. Cependant, en raison des contraintes propres à une évaluation à large échelle, où la comparabilité et la standardisation des résultats importent (comparabilité entre pays et entre élèves dans les pays), l'ouverture qui se manifeste souffre certaines limites. Il est certes admis théoriquement que plusieurs interprétations et plusieurs réponses sont possibles à certaines questions, mais on s'en tiendra à des interprétations « stabilisées/stabilisables », « normées », dont la plausibilité peut être estimée et appréciée en termes de « correction ».

La définition proposée pour Pirls dans le cadre conceptuel provisoire²

¹ « *Understanding, using and reflecting on written texts, in order to achieve one's goals, to develop one's knowledge and potential and to participate in society* » (Pisa, 1999 : 20).

² L'étude Pirls de l'IEA (testing en 2001) n'en est qu'à sa phase d'élaboration et les informations contenues dans le document provisoire (Draft framework) sont susceptibles d'être encore modifiées. Mais la probabilité est faible que des changements majeurs interviennent, qui infléchiraient nettement les commentaires relatifs à cette étude. C'est pourquoi nous avons pris le parti d'inclure cette étude dans notre revue des évaluations consacrées à la lecture.

de novembre 1999 s'inscrit dans la même ligne : « *la littératie est la capacité à comprendre et à utiliser ces formes du langage écrit requises par la société ou importantes pour l'individu. Les jeunes lecteurs sont amenés à construire le sens d'une diversité de textes. Ils apprennent à lire, à participer à des communautés de lecteurs, et pour leur plaisir¹ » (Pirls/IEA, 1999 : 2). La continuité avec l'étude IEARL de 1991 est manifeste, mais on note une nuance de taille : la référence aux théories de la lecture comme processus interactif de construction de sens est dans *PIRLS* pleinement affirmée et se traduira par l'évaluation, auprès d'élèves de 9-10 ans, de démarches complexes telles que « tirer et justifier des inférences complexes » (*draw and justify complex inferences and interpretations*) ou « examiner et évaluer le contenu, la langue et les éléments textuels » (*examine and evaluate content, language, and textual elements*). Avec ces définitions ambitieuses, on est loin d'une conception étroitement fonctionnelle de la littératie; les définitions que se donnent ces deux*

dernières études de la « *reading literacy* » embrassent de fait ce qui, en 1971, était apparu aux responsables et experts de l'IEA comme relevant de l'éducation ou de la culture « littéraires »². A l'époque, il faut le préciser, l'évaluation de la culture littéraire ne concernait pas tous les élèves, comme c'est aujourd'hui le cas dans Pisa et dans Pirls, mais une fraction d'élèves, plus ou moins sélectionnée selon les pays, qui avaient réussi à se maintenir dans les filières générales de l'enseignement ou, plus simplement, à l'école. Les exigences les plus élevées s'appliquent aujourd'hui à la masse des élèves de 9 ou 15 ans, portées par une double référence : une référence théorique, d'une part, qui voit la compréhension comme le résultat d'un processus de construction et d'une interaction entre le texte et les connaissances antérieures du lecteur, d'autre part, une référence aux exigences de la société supposées de plus en plus élevées, un **relèvement des standards de littératie** dont Pierre (1990; 1992) a bien montré toute l'ampleur³.

¹ « *Reading literacy is the ability to understand and use those written language forms required by society and valued by the individual. Young readers can construct meaning from a variety of texts. They read to learn, to participate in communities of readers and for enjoyment* » (Pirls/IEA, 1999 : 2).

² Les définitions adoptées dans Pisa et Pirls correspondent davantage à ce que d'aucuns appelleraient la « littératie cultivée, avancée ou élevée » (« *cultural, advanced or high literacy* ») (Harris and Hodges, 1995).

³ Pierre estime ainsi qu'aujourd'hui, « dans les sociétés technologisées, le moindre travail exige un niveau de littératie correspondant à une onzième année, alors qu'il y a 50 ans, en Amérique du Nord, la moyenne de scolarisation correspondait à une quatrième année primaire » (1992 : 4).

3. Quel lien entre l'évaluation et les apprentissages menés ou le curriculum ?

L'exigence de cohérence entre l'évaluation et les apprentissages prévus par les instructions officielles ou réalisés sur le terrain (curriculum théorique ou implanté) est l'une des clés de voûte de l'évaluation, de quelque nature qu'elle soit. Dans les évaluations internationales, elle se pose avec d'autant plus d'acuité qu'elle se confond avec une exigence d'équité entre pays : les pays ne souhaitent pas être évalués (en tout cas pas trop) à l'aune de standards pour lesquels il n'existerait pas un minimum de consensus international (recherche d'un dénominateur commun). C'est pourquoi les évaluations internationales envisagent généralement une étude du curriculum de la discipline concernée dans les différents pays participants. Mais, pour la lecture, ce n'est pas le cas, et cela n'a pas changé en trente ans. En 1971, on a estimé que le consensus sur les objectifs de l'enseignement de la lecture était suffisamment large pour ne pas entreprendre une étude systématique du curriculum; on a considéré qu'il existait « un accord général, au-delà des nuances nationales ou locales, sur le fait que les enfants doivent être capables de comprendre le sens d'écrits de styles variés et de contenus divers¹ ». Les

¹ « In the field of reading, there seems to be much more consensus as to the objectives of the instruction. (...) There would be general agreement that children should be able to get

études ultérieures ne varieront pas sur ce point. Et, en cela, le parallèle que l'on peut faire avec des pratiques d'évaluation plus courantes n'est pas sans intérêt. Dans les classes, quand il s'agit de lecture, on évalue aussi souvent sans tenir compte de ce qui devrait être enseigné ou de ce qui est enseigné, au point que dans certaines classes, la lecture se limite et se confond avec son évaluation. En lecture, davantage que dans d'autres disciplines comme les mathématiques, le lien entre évaluation et pratiques d'enseignement est incertain.

4. Structuration du champ et caractéristiques des tâches de lecture

Une fois défini l'objet de l'évaluation, se pose la question des aspects plus précis sur lesquels va porter l'évaluation : quel type de matériel ou de textes ? Quelles compétences ou démarches du lecteur ?

4.1. Types de textes

La plus grande constante à cet égard est sans doute la référence, omniprésente, à une variété de matériels écrits. Ainsi dans l'étude sur la compréhension de lecture de 1971, Thorndike évoque le recours à du matériel écrit de contenus et styles variés (« *written material of*

meaning efficiently from written material of a various styles and content areas », Thorndike, 1971.

various styles and content areas »). A titre d'exemples, il oppose ainsi les textes informatifs (*expository*) aux textes littéraires (*literary*), ou la prose à la poésie. Mais cette évocation ne sert que de garde-fou à la diversité. Aucun équilibre a priori entre ces différents types ou genres de textes n'est exigé pour la construction de l'épreuve. En 1991 dans IEARL, de même que dans IALS, le premier élément de structuration du champ est la référence à des « domaines » correspondant à de vastes catégories de types d'écrit : textes narratifs, textes informatifs et documents pour IEARL; prose, documents et écrits comportant des données numériques pour IALS. En 1999, dans Pisa et dans Pirls, la référence au type et au format des textes reste essentielle, même si elle emprunte parfois d'autres voies. Dans Pisa, une première distinction est posée entre textes continus et discontinus, qui se poursuit par des typologies spécifiques, en référence aux travaux théoriques sur les typologies de textes qui constituent l'un des savoirs de référence de la didactique. Pour chacun des types de textes ainsi inventoriés, une proportion d'unités et/ou d'items est définie a priori, que les développeurs de tests devront s'efforcer de respecter. Dans Pirls, la référence à des typologies de texte est absente, mais l'un des deux grands axes de structuration du champ - les buts de lecture - (*purposes of reading*) conduit à un découpage proche de celui fondé sur les types de textes : sous « lire pour

des motifs littéraires » (« *reading for literary experience* ») et « lire pour acquérir et utiliser des informations » (« *reading to acquire and use information* »), on retrouve sans trop de peine les grands types de textes évoqués plus haut. Par ailleurs, toutes les études, depuis IEARL, présentent ou envisagent de présenter les résultats sur des échelles séparées en fonction des types de textes. Au-delà des arguments théoriques relevant des typologies de textes, quelques arguments empiriques existent pour montrer que les compétences mobilisées par les textes continus, d'une part (qu'ils soient narratifs ou informatifs), et les textes non continus d'autre part (cartes, graphiques, tableaux...), sont certes étroitement corrélées, mais suffisamment distinctes pour qu'on les envisage séparément (Atash, 1995).

4.2. Habiletés, compétences, démarches ou processus... (Skills, abilities, processes or aspects...)

Le premier axe de structuration du champ, on l'a vu, concerne essentiellement la nature des textes/stimuli (le pôle texte dans le trio texte/lecteur/contexte). Le second, assez logiquement, concerne la nature des compétences que l'on évalue. Sur quels processus, quelles démarches, quelles compétences va porter l'évaluation de la lecture ? Au fil du temps, la terminologie et les concepts ont certes varié, mais toutes les études envisagent, d'une manière plus ou moins explicite, une

structuration du champ en fonction des compétences du lecteur. La référence au concept spécifique de « compétence » est cependant largement absente des évaluations internationales, où le cadre de référence est à dominante anglo-saxonne¹.

En 1971, on l'a dit précédemment, l'IEA a posé d'emblée une distinction relative à cette question en prévoyant deux études séparées de la compréhension et de l'interprétation littéraire. Pour l'étude sur la compréhension, l'auteur de l'ouvrage de synthèse (Thorndike, 1973) rappelle que l'étude portera sur le contenu cognitif du texte et que l'on ne se concentrera pas sur l'appréhension du style, des sentiments, du ton ou des techniques littéraires². Ce principe une fois établi, la réflexion sur les « skills » (compétences) à évaluer pour la lecture débouche sur le choix d'évaluer (1) la compréhension de lecture (2) la vitesse de lecture (3) la connaissance du vocabulaire. L'existence de sous-compétences pour la compréhension de lecture n'est pas mentionnée dans les fondements de l'étude, même si

un examen attentif des items permet de déceler que ceux-ci font appel à des processus de nature et de complexité différentes (on trouve par exemple des tâches de lecture consistant à retrouver l'idée principale d'un passage ou nécessitant des inférences). Cette évaluation reflète bien l'état des recherches théoriques dans le domaine de la lecture de l'époque, marquées par une conception behavioriste : une centration sur des prérequis et/ou des skills associés à la compréhension (le vocabulaire et la vitesse), la compréhension représentant quant à elle une sorte de boîte noire ou de continent encore inexploré. Cette étude est en effet antérieure au formidable développement des modèles cognitifs de la lecture qui marquera les années 80 (Barr *et al.*, 1990) et dont l'objet essentiel consistera à essayer de voir précisément ce qu'il y a à l'intérieur de la boîte noire et à déterminer quelles composantes englobe la « compréhension en lecture ». A la fin des années soixante, la diversité observable dans les items saisit intuitivement une diversité de processus ou de compétences, mais les références théoriques pour penser a priori cette diversité n'existent pas encore.

L'étude IEARL de 1991 amorce un tournant sur la question des compétences à évaluer. D'après le rapport d'Elley consacré à cette étude (1994), il semble que la question des skills/processus mentaux (*skills (or mental processes)*)

¹ Le terme « *competency* » est peu utilisé en anglais et ne se distingue pas fondamentalement d'autres concepts proches tels que « *skill* » ou « *ability* ». Lorsque l'on veut faire référence à des formes d'évaluation plus authentiques, menées au cours d'activités de résolution de problèmes, ce sont les termes « *performance assessment* » qui sont utilisés en anglais.

² « *it was decided to focus upon the cognitive content of the passage and to forego most efforts to get any appraisal of style, feeling, tone or literary techniques* » (Thorndike, 1973 : 19).

n'ait pas été posée a priori¹. Elle n'apparaît en tout cas pas dans les étapes relatives à la définition des domaines et au développement de l'épreuve. En revanche, la question est évoquée quand il s'agit de décrire les tests définitifs. La spécification des processus mentaux impliqués par les différents items vient alors s'ajouter aux domaines définis préalablement, relatifs aux types de textes (textes narratifs, informatifs ou documents). La manière dont ces processus mentaux sont définis semble encore relever largement de l'intuition ou d'une démarche empirique. Aucune référence théorique à des modèles n'est présente. Le classement des items par processus a été effectué par le comité d'experts internationaux et soumis à l'appréciation de 27 juges volontaires, issus de 12 pays différents. Les processus sont différents selon le type de textes. Pour les textes suivis (narratifs et informatifs), on distingue les réponses « verbatim » et « paraphrase », les questions portant sur l'idée principale (*main idea*) et les questions d'inférence (*inference*)². Pour les textes non

suivis (documents), deux processus sont distingués : localiser l'information, d'une part, localiser et traiter l'information, d'autre part. Par rapport à 1971, l'intuition s'est précisée, même si elle ne semble pas avoir guidé véritablement l'élaboration des tests. Une hypothèse possible est que la nécessité de classer les items selon le processus impliqué se soit fait jour lorsque le choix de présenter les résultats moyens internationaux sur des échelles IRT (Item Response Theory³) a été posé⁴.

Point d'aboutissement de ce long processus qui va d'une intuition ou d'un classement empirique a posteriori à la construction d'un plan raisonné définissant a priori quels domaines on veut évaluer, dans Pisa, cinq macro-aspects⁵ à évaluer sont définis : comprendre globalement un texte, trouver de l'information, développer une interprétation, réfléchir sur le contenu d'un texte et réfléchir sur la forme d'un texte. Pour chacun

¹ C'est du moins l'impression que l'on a à la lecture de l'ouvrage de synthèse. Mais il n'est pas exclu que cet aspect ait été longuement considéré dans le comité d'experts internationaux ou dans les débats avec les représentants nationaux.

² Les items « verbatim » sont des items pour lesquels la réponse correcte correspond terme à terme à ce qui est dit dans le texte. Pour les items « paraphrase », la réponse correcte correspond à un synonyme ou une reformulation de ce qui est dit dans le texte. En cas d'inférence, le lecteur doit aller au-delà du texte, expliciter des éléments laissés implicites dans le texte.

³ Le principe de ces échelles est décrit plus loin (p. 13).

⁴ Ce dernier choix lui-même semble avoir été pris a posteriori, voire en dernière minute; il n'est en effet pas pleinement intégré à la présentation des résultats. Paradoxalement, les échelles IRT figurent dans l'introduction de Elley (1994) et non dans le chapitre consacré aux « international differences in achievement levels ».

⁵ Il ne s'agit pas à proprement parler de « processus », puisque des processus de nature différente peuvent être mobilisés pour remplir une même tâche de lecture. Si, par exemple, il s'agit de dégager l'idée principale d'un texte, celle-ci peut être exprimée dans le texte, et il suffira de la localiser; dans d'autres cas, il faudra pour cela développer une inférence simple, voire une inférence complexe.

de ces aspects, des proportions d'items à atteindre sont arrêtées avant l'élaboration des tests. Parallèlement, chacun des items de l'évaluation est classé selon le macro et parfois le micro-aspect auquel il fait appel. Ces deux critères définissent ce qui est appelé l'objectif de la question. Tel item sera par exemple considéré comme un item « retrouver une information : correspondance synonymique ». La nécessité de préciser ainsi chacun des items et de prévoir des équilibres numériques entre catégories est une question intimement liée aux échelles IRT et plus particulièrement au nombre et à la nature des échelles sur lesquelles on envisage de présenter les résultats. Il serait en effet regrettable de constater a posteriori qu'une dimension essentielle, pour laquelle on pourrait souhaiter une échelle spécifique, n'est pas représentée par un nombre d'items suffisant ou que les items couvrant cette dimension ne varient pas par leur degré de difficulté. L'arrivée des échelles IRT force ainsi, de son côté, le passage de l'intuition ou du bricolage savant, à une approche plus théorique, et plus complexe techniquement, de recherche d'équilibres a priori.

Dans leur plan d'évaluation (*Reading framework*), les experts de Pisa insistent sur le fait que les cinq aspects évalués (correspondant aux compétences du lecteur) n'obéissent pas à une hiérarchie de complexité stricte. Cette précision est capitale

pour la réflexion sur le développement des compétences chez l'apprenant et pour leur évaluation. La conviction est en effet largement répandue (il s'agit d'une représentation spontanée) que certains processus ou certaines composantes de la compétence en lecture sont intrinsèquement plus complexes que d'autres et qu'il convient donc de les aborder plus tard dans le curriculum, que ce soit à titre d'apprentissage ou comme objet d'évaluation. Or cette conception n'est que partiellement vraie. Elle n'est vraie qu'en se plaçant sur un strict plan cognitif, et en faisant abstraction de tous les autres paramètres. Or, conformément aux modèles du lecteur aujourd'hui communément admis (pour une synthèse, voir Giasson, 1990, ou Gombert et Fayol, 1992), la lecture est le résultat d'une interaction entre un texte, un lecteur et un contexte. Ainsi, il sera éventuellement plus simple, et en tout cas plus aisé, d'appliquer une démarche cognitive plus complexe, telle qu'inférer, à un texte d'un niveau de difficulté peu élevé et/ou dont le contenu est familier au lecteur que l'inverse (appliquer une démarche cognitive plus complexe à un texte difficile ou dont le contenu est peu familier au lecteur). La compétence de lecteur observée est le fruit de cette interaction singulière, il n'y a pas à sortir de là. On rappellera à ce propos cette expérience désormais classique en psychologie cognitive (Recht et Leslie, 1988 : 16-20) qui montre que face à un texte au contenu technique (portant sur une

technique sportive), des lecteurs plus faibles en général, mais familiers du contenu, obtiennent de meilleurs scores que de « bons » lecteurs peu familiers du contenu. Cette expérience illustre à merveille le fait que la lecture est bien le fruit d'une interaction. Toute réflexion sur les compétences en lecture et sur leur hiérarchisation, de même que sur la progression dans les apprentissages, doit intégrer cette donnée essentielle. Par ailleurs, sur le plan de l'évaluation, cette expérience rappelle la nécessité d'évaluer les compétences au travers d'un ensemble large de matériels écrits, afin de minimiser les « biais » (qui ne sont rien d'autre que des effets d'interaction) liés à une familiarité inégale avec le contenu ou le type de texte.

Au-delà de cette réflexion théorique, des recherches plus empiriques consacrées à IEARL et IALS se sont attachées à essayer de comprendre quels facteurs expliquent la difficulté des tâches en lecture, et plus exactement fondent la hiérarchie de difficulté observée entre les items. Kirsch et Mosenthal (1995) ont ainsi élaboré un modèle qui intègre deux types de variables, la lisibilité du texte¹ et ce qu'ils appellent des variables de processus. Parmi ces variables de processus, ils distinguent le type d'information (la question porte sur une information plus ou moins concrète), le type de correspondance (la correspondance

que le lecteur doit établir entre la question et le texte est littérale, synonyme, inférentielle, interprétative...) et le degré de plausibilité des distracteurs (certains des distracteurs proposés sont-ils proches d'éléments explicitement exprimés dans le texte). Appliqué aux données de l'étude IEA Reading Literacy, ce modèle permet de rendre compte d'une proportion considérable de la variance de la difficulté des tâches. La variable la moins influente est la lisibilité du texte; les deux autres variables ont un très haut pouvoir prédictif, plus ou moins marqué selon le type de texte ou l'âge de la population testée notamment (pour plus de détails, on se reportera à l'étude citée). L'existence d'un tel modèle permet de mieux appréhender ce qui fonde la difficulté des tâches de lecture et permet, de façon prospective, de décider sur quels paramètres agir pour couvrir un éventail de difficulté plus ou moins large, selon l'objectif que l'on se donne. Ce modèle a de la sorte été réinvesti pour construire les évaluations ultérieures (notamment Pisa).

5. Plan d'évaluation et présentation des résultats

L'utilisation des analyses IRT (Rasch, 1980; Laveault & Grégoire, 1997; Lafontaine & Schillings, 1999, Demeuse, 2000) dans les grandes enquêtes comparatives représente à n'en point douter l'une des avancées les plus importantes de ces vingt

¹ Il s'agit de mesures classiques de lisibilité (formules de Fry, 1977).

dernières années¹. On l'a vu, le recours à de telles échelles n'est sans doute pas étranger au développement d'une phase de conceptualisation théorique préalablement aux études. Par ailleurs, l'existence de telles techniques rend possible l'utilisation de plans d'évaluation complexes avec rotation de carnets (*complex matrix sampling*). Davantage d'items peuvent être utilisés, davantage de compétences évaluées sans que chaque individu testé doive pour autant subir une charge accrue. Les différents items sont répartis dans des carnets qui ont une partie commune (par exemple 25 items sur 60) et une partie variable (les 35 autres items varient d'un carnet à l'autre). La technique d'ancrage rend possible la comparaison des scores d'élèves qui n'ont pas passé exactement le même ensemble d'items.

Le principe de ces modèles IRT est, il faut le rappeler, qu'ils permettent d'exprimer, sur une même échelle, le niveau de difficulté d'un item et le niveau de compétence d'un individu. Un individu qui obtient un score d'une valeur n sur une telle échelle a une probabilité supérieure à un seuil fixé (on utilise habituellement 50 % ou 80 %) de réussir les items d'une difficulté inférieure à n et une probabilité inférieure au seuil fixé de réussir les items d'une difficulté supérieure à n . Comme l'indiquent

Kirsch, Mosenthal et Jungeblut (1998), « *les modèles de la réponse à l'item fournissent un moyen d'envisager sous un même angle les individus et les tâches*² » (p.105) Les échelles IRT permettent d'aller au-delà d'une perspective strictement normative consistant à ordonner simplement les individus ou les pays, dans les enquêtes internationales, du plus performant au moins performant. Avant les échelles IRT, il ne restait la place, au-delà du classement, que pour une analyse qualitative. Avec l'arrivée des modèles IRT, l'information se précise et devient plus diagnostique. Non seulement tel pays sait qu'il est plus ou moins performant que tel autre, mais il sait aussi en quoi : face à quel type de texte, quelle proportion des individus évalués se sont révélés capables, ou non, de mettre en oeuvre telle démarche de lecture ou d'effectuer telle tâche. Ces modèles réussissent en quelque sorte à combiner de façon élégante une approche normative (qui reste au coeur de toute étude comparative) et une approche davantage centrée sur les compétences. L'évaluation ne porte plus seulement sur une quantification de la compétence globale en lecture (score de rendement); elle vise à diagnostiquer plus précisément dans quelle mesure les individus se montrent capables de mobiliser certains types de compé-

¹ Voir exemple d'échelle IRT en annexe 3.

² « Item response theory (IRT) scaling provides a way to focus both on people and tasks together. » (Kirsch, Mosenthal et Jungeblut, 1998 : 105).

tences en lecture face à des textes et des tâches diversifiées de lecture.

6. Format des questions

L'une des questions qui a régulièrement animé les groupes d'experts ou de représentants nationaux dans les études internationales est celle du format des questions : questions ouvertes ou à choix multiple, questions fermées ou à réponse construite ? En très schématiquement résumé, les questions fermées ou les QCM ont pour elles l'avantage d'être rapidement corrigées, d'une façon fidèle et peu coûteuse; elles se prêtent mal, en revanche, à l'évaluation de certains processus (la réflexion critique notamment). Les questions ouvertes à réponse construite sont jugées plus authentiques, s'imposent pour l'évaluation de certains aspects, mais elles nécessitent l'élaboration de guides de correction standardisés dont la mise en oeuvre sera immanquablement coûteuse. La comparabilité des résultats peut dans ce dernier cas se révéler plus incertaine (problèmes de fidélité inter-correcteurs). Par ailleurs, si une question exige une réponse élaborée, l'évaluation des compétences en lecture se confond en partie avec les compétences en expression écrite. Voyons comment les évaluations successives ont tranché cette épineuse question¹.

¹ Cette question d'apparence technique véhicule celles, beaucoup plus fondamentales, du lien aux pratiques de la classe et de l'effet en retour

En 1971, pour cette grande première évaluation internationale de la lecture, tous les efforts se concentrent en vue de montrer qu'une telle évaluation comparative est faisable et les choix proposés en matière d'items sont dès lors prudents : « *le test de compréhension doit être classique : un passage est présenté à l'élève, assorti de questions portant sur ce passage*² » (Thorndike, 1973 : 20). Ce choix guidé par la prudence est de toute façon cohérent avec le modèle de la lecture que s'est donné cette évaluation. Pour toutes les questions posées, il existe en effet une et une seule réponse correcte, comme le montre la manière dont sont formulées les instructions aux élèves et rien ne s'oppose donc, sur le plan théorique, à ce que des QCM et rien que des QCM soient proposées.

Au début des années 90, l'étude IEARL, reproduit, avec peu de variations, les mêmes choix. Certes des voix se sont élevées, parmi les représentants nationaux, pour réclamer des questions ouvertes (à réponse construite), mais le comité d'experts internationaux a préféré emprunter la voie du pragmatisme, en appuyant ses choix sur une étude

(*washback effect*) que peut exercer une évaluation dont le mode de questionnement n'est pas conforme aux usages en vigueur ou aux usages que les responsables éducatifs voudraient voir adopter par les enseignants.

² « *The reading comprehension test was to be of the conventional type in which a passage is presented to the pupil together with multiple-choice questions based on that passage* » (Thorndike, 1973 : 20).

(Elley et Mangubhai, 1992) qui montre des résultats apparemment similaires avec les deux formats de questions¹. Les tests comporteront donc une majorité de questions à choix multiple (de 90 % en population A à 75 % en population B). Dix à vingt pour-cent des questions seront des questions ouvertes où il faut répondre par un mot, un chiffre ou quelques mots (celles-ci ne posent aucun problème de correction). A chaque niveau, deux questions à réponse construite seront incluses, à titre d'option internationale; les résultats de cette option sont analysés dans Kapinus et Atash (1995).

Avec IALS, c'est un pari audacieux, et exactement inverse qui est pris : toutes les questions seront des questions ouvertes. La justification du choix est plutôt laconique : « *toutes les questions étaient ouvertes parce que l'on a pensé que des adultes seraient plus intéressés à réaliser de telles tâches*² » (Murray, Kirsch and Jenkins, 1994 : 19). De prime abord, cette décision semble se fonder plus sur des raisons psychologiques que psychométriques ou théoriques.

Dans la foulée, les deux études Pisa et Pirls comporteront également une proportion importante de réponses ouvertes construites (nécessitant une correction multiple). Ce choix est dans les deux cas fondé non seulement sur le souci de rendre les tâches plus « authentiques », il est guidé aussi par des raisons théoriques. Certains des aspects ou processus évalués supposent en effet que le lecteur mette en résonance le texte et ses connaissances antérieures. Plusieurs « bonnes » réponses sont donc envisageables et il devient difficile de couler ce type de question, où il est fait appel à une part de créativité du lecteur, dans le moule des questions à choix multiple.

7. Synthèse

Ce regard porté sur trente années d'évaluations internationales de la lecture a permis de mettre au jour quelques axes d'évolution majeurs³, que l'on observerait sans doute en examinant d'autres évaluations. Ainsi, depuis les années 70 :

¹ Tout dépend évidemment des critères que l'on se donne pour évaluer la similarité. D'autres études (Kapinus et Atash, 1995) ont apporté des réponses différentes, ou en tout cas plus nuancées, à cette même question.

² « *All of the literacy tasks were open-ended rather than multiple-choice, because it was thought that adults would be more interested in performing such tasks* » (Murray, Kirsch and Jenkins, 1994 : 19).

³ Un tableau présentant les principales caractéristiques des surveys envisagés figure en annexe 1.

- **Les conceptions de ce qu'est la compréhension en lecture, puis la littératie (reading literacy), se sont considérablement étendues et enrichies**, parallèlement au développement des recherches cognitives sur les processus de lecture. D'une conception de la compréhension centrée sur la restitution d'un sens univoque déposé dans le texte (ce qui échappe à cela est renvoyé à la sphère du littéraire), on a progressivement glissé à une conception de la littératie réconciliant les versants cognitifs et culturels et où il est reconnu que le processus d'élaboration du sens d'un texte résulte d'une interaction entre ce texte et un lecteur.
- **Le travail préalable d'élaboration théorique** s'est sensiblement développé sous la double influence des modèles théoriques (processus cognitifs, typologies de textes, recherches sur les facteurs prédictifs de la difficulté des tâches) et des modèles psychométriques (échelles IRT). La démarche empirique caractéristique des premières études a progressivement cédé le pas à une approche structurée, planifiée, où les équilibres entre différentes contraintes et critères sont définis a priori.
- La proportion de questions à choix multiple a régressé tandis que se développait l'usage de **questions à réponse construite** et de guides destinés à assurer une correction aussi objective que possible de ces

réponses complexes, mieux en accord avec les conceptions contemporaines de la lecture. Un effort a été fait pour utiliser des textes ou des documents plus authentiques, que l'on n'écrit ou ne réécrit pas pour les besoins de l'évaluation.

- L'apparition de nouveaux modèles psychométriques (**échelles IRT**) a dynamisé, ou accompagné, une approche plus diagnostique, davantage centrée sur les compétences. Les enquêtes internationales ne se limitent plus aujourd'hui à produire un score global de rendement, elles visent à établir quelles proportions de lecteurs sont capables de mobiliser certains types de compétences face à des tâches de lecture diversifiées.

Les évaluations menées sur le plan international sont la résultante de compromis multiples. Celles-ci doivent composer avec les contraintes propres à toute évaluation à large échelle, les exigences liées à la comparaison des données et aux avancées des savoirs de référence (modèles cognitifs, modèles psychométriques, modèles linguistiques, psycholinguistiques et didactiques pour la lecture). Certes, la machinerie de ces évaluations est lourde, ce qui la met à l'abri des effets de mode passagers. L'évolution sur cette période de 30 ans est cependant perceptible et aura suffisamment montré, je l'espère, que ces «grosses machines»

ont réussi à s'adapter, dans un laps de temps raisonnable, et compte tenu des contraintes qui leur sont propres, à une conception de la lecture interactive qui intègre l'évaluation de compétences plus complexes que le simple repérage d'informations littérales via des questions à choix multiple.

Le cadre de référence utilisé ici pour analyser l'évolution dans les évaluations internationales pourrait sans doute s'appliquer utilement à d'autres types d'évaluation : évaluations externes nationales ou locales, ou encore pratiques d'évaluation bilan. Il permet en effet de s'interroger sur la cohérence interne d'une évaluation (les modalités d'évaluation sont-elles en concordance avec la définition de la lecture que l'on s'est choisie ou que les instructions officielles prônent, par exemple ?), voire de la guider. Il peut aussi permettre d'amorcer, au départ d'évaluations existantes, ou à construire, la question du lien entre apprentissage/enseignement et évaluation de la lecture. L'idée peut sembler singulière, mais comme dans nombre de classes les activités de lecture se résument hélas à son évaluation (Guernier, 1999; Giasson, 1990), elle pourrait se révéler singulièrement didactique.

8. Pour la littératie...

Au moment de conclure, je voudrais revenir un bref instant sur le concept de littératie pour tenter de cerner

quels véritables changements recouvre le recours à ce concept. On a vu, que pour les évaluations internationales, le terme apparaît à la fin des années 80 et que son usage se traduit par un élargissement à des types de textes moins littéraires ou moins savants - les documents que l'on utilise dans la vie courante. Ce lien me paraît bien établi. On a vu aussi que le recours au terme de littératie va de pair, au moins sur le plan temporel (les deux phénomènes sont quasi contemporains) avec une évolution des conceptions, des modèles de ce qu'est la lecture. Cette concomitance est, à mon avis, relativement fortuite; les modèles cognitifs de la lecture qui sous-tendent cette évolution ne font en effet pas référence au concept de littératie. Même si elle est fortuite, cette coïncidence mérite peut-être que l'on s'en saisisse. Le concept de littératie, tel qu'il a été défini par Grossman ou Pierre¹ et par beaucoup d'auteurs anglo-saxons, en réconciliant compréhension et usages, comportements et attitudes face à l'écrit, est implicitement porteur de cette conception interactive de la lecture dont on a montré toute l'importance pour une réflexion sur le développement des compétences en lecture. Malgré son caractère général, le concept a en outre le mérite de faire référence, de façon positive, à un processus continu de développement des compétences en lecture, ce que ne permet pas le

¹ Voir leurs citations plus haut.

terme d'alphabétisation¹. Enfin, alors que le concept de compréhension renvoie à des compétences d'ordre essentiellement cognitif, le concept de littératie rappelle que devenir un(e) lecteur/lectrice compétent

suppose non seulement des habiletés cognitives, mais l'élaboration d'un rapport positif de connivence ou de familiarité à l'écrit qui déborde largement la sphère du cognitif.

Références bibliographiques

Evaluations internationales de la lecture

Elley, W.B. (1994). *The IEA Study of reading literacy : Achievement and instruction in thirty-two school systems*. Oxford : Pergamon.

Lafontaine, D. (1996). *Performances en lecture et contexte éducatif. Enquête internationale menée auprès d'élèves de 9 et 14 ans*. Bruxelles : De Boeck.

IEA PIRLS (1999). *Draft framework and specifications for the PIRLS Assessment of Reading Literacy. Progress in International Reading Literacy Study*. Chestnut Hill, Ma : PIRLS International Study Center, Boston College.

Murray, T.S., Kirsch, I.S., Jenkins, L.B. (1998). *Adult literacy in OECD countries. Technical report on the First International Adult Literacy Survey*. Washington : National Center for Education Statistics, US Department of Education, Office of Educational Research and Improvement.

OCDE (1999). *Mesurer les connaissances et compétences des élèves. Un nouveau cadre d'évaluation*. Paris : Les éditions de l'OCDE.

Purves, A.C. (1973). *Literature education in ten countries. International studies in evaluation II*. Stockholm : Almqvist & Wiksell.

Thorndike, R.L. (1973). *Reading comprehension education in fifteen countries. International studies in evaluation III*. Stockholm : Almqvist & Wiksell.

¹ Le concept de littératie ne résout pas tous les problèmes, car l'adjectif correspondant, qui permettrait d'exprimer cette continuité, fait hélas défaut. Ma connaissance des nuances de la langue anglaise est insuffisante pour mesurer l'ampleur de l'écart qui sépare les «literate» des «illiterate». En français, c'est un fossé, que dis-je un monde, qui distingue les «illettrés» des «lettrés».

Autres références

- Atash, N. (1995). Assessing the dimensionality of the IEA Reading Literacy Data. In M. Binkley, K. Rust, M. Winglee (Eds), *Methodological issues in comparative educational studies. The case of the IEA Reading Literacy Study*. Washington, D.C. : National Center for Education Statistics. US Department of Education, Office of Educational Research and Improvement, 75-102.
- Barr, R., Kamil, M.L., Mosenthal, P. & Pearson, P.D. (Eds) (1990). *Handbook of reading research (vol. 2)*. New York : Longman.
- Demeuse, M. (2000). *Les échelles de mesure : Thurstone, Likert, Guttman et le modèle de Rasch*. Liège : Service de Pédagogie expérimentale de l'Université, Série "Notes techniques" - 2000/2.
- Elley, W.B. & Mangubhai, F. (1992). Multiple-choice and open-ended items in reading tests. *Studies in Educational Evaluation*, 18(2), 191-199.
- Fayol, M. (1992). Comprendre ce qu'on lit : de l'automatisme au contrôle. In M. Fayol, J.-E. Gombert, P. Lecocq, L. Spenger-Charolles & D. Zagar (Eds), *Psychologie cognitive de la lecture*. Paris : Presses Universitaires de France.
- Giasson, J. (1990). *La compréhension en lecture*. Bruxelles : De Boeck.
- Giasson, J. (1997). *La lecture : de la théorie à la pratique*. Bruxelles : De Boeck.
- Grossman, F. (1999). Littératie, compréhension et interprétation de textes. *Repères*, 19, 139-167.
- Guernier, M.-C. (1999), Lire et répondre à des questions au cycle 3. *Repères*, 19, 167-183.
- Harris, T.L. & Hodges, R.E. (Eds) (1995). *The literacy dictionary. The vocabulary of reading and writing*. Newark, Delaware : International Reading Association.
- Kapinus, B. & Atash, N. (1995). Exploring the possibilities of constructed-response items. In M. Binkley, K. Rust, M. Winglee (Eds), *Methodological issues in comparative educational studies. The case of the IEA Reading Literacy Study*. Washington, D.C. : National Center for Education Statistics. US Department of Education, Office of Educational Research and Improvement, 105-133.
- Kirsch, I.S. & Mosenthal, P.B. (1995). Interpreting the IEA Reading Literacy Scales. In M. Binkley, K. Rust, M. Winglee (Eds), *Methodological issues in comparative educational studies. The case of the IEA Reading Literacy Study*. Washington, D.C. : National Center for Education Statistics. US Department of Education, Office of Educational Research and Improvement, 135-193.
- Kirsch, I.S., Mosenthal, P.B. & Jungeblut (1998). The measurement of adult literacy. In T.S. Murray, I.S. Kirsch, L.B. Jenkins, *Adult literacy in OECD countries. Technical report on the First International Adult Literacy Survey*. Washington : National

- Center for Education Statistics, US Department of Education, Office of Educational Research and Improvement, 105-134.
- Lafontaine, D. (1996). *Performances en lecture et contexte éducatif. Enquête internationale menée auprès d'élèves de 9 et 14 ans*. Bruxelles : De Boeck.
- Lafontaine, D. & Schillings, P. (1999). Mesurer les progrès accomplis grâce au modèle de la réponse à l'item: l'évolution des compétences en lecture à 14-15 ans en Communauté française de Belgique entre 1991 et 1998. *Scientia Paedagogica Experimentalis*, XXXVI, 2, 165-178.
- Laveault, D., Gregoire, J. (1997). *Introduction aux théories des tests en sciences humaines*. Bruxelles : De Boeck, coll. Méthodes en sciences humaines.
- Martin, M.O., Rust, K., Adams, R.J. (Eds) (1999). *Technical Standards for IEA Studies*. Delft, Netherlands : Eburon Publishers.
- Pierre, R. (1990). De l'alphabétisation à la littéracie : pour une réforme en profondeur de l'enseignement. *Scientia Paedagogica Experimentalis*, XXVIII, 2, 151-186.
- Pierre, R. (1992). La compréhension de textes écrits face au rehaussement des standards de littératie. *Scientia Paedagogica Experimentalis*, XXIX, 1, 3-21.
- Rasch, G. (1980). *Probabilistic models for some intelligence and attainment tests*. Chicago : University of Chicago.
- Recht et Leslie (1988). Effect of prior knowledge on good and poor readers's memory of text. *Journal of Educational Psychology*, 80, 16-20.
- Rosenblatt, L.M. (1938/1976). *Literature as exploration*. New York : Modern Language Association.

Liste des sigles utilisés

- IALS : International Adult Literacy Study.
 IEA : Association internationale pour l'évaluation du rendement scolaire.
 IEARL : Etude de l'IEA sur la «Reading Literacy».
- IRT (échelle) : Item Response Theory.
 PIRLS : Progress in International Reading Literacy Study.
 PISA : Programme international d'évaluation du suivi des acquis des élèves.

ANNEXES

ANNEXE 1

Etudes/ Commanditaire	Objet	Population	Domaines	Habilités/ Processus	Echelles	Plan d'évaluation	Format des questions
IEA Six Subject Study (1971)	Compréhension de la lecture	10-11 ans 14-15 ans Fin du secondaire supérieur	Pas de domaines spécifiés (« variété de contenus et de styles ») Tous les textes sont continus (narratifs et informatifs)	Compréhension de la lecture/ vitesse de lecture/ vocabulaire	-	Pas de rotation. Une seule forme de test	100 % QCM
IEARL (1991)	Littératie/ Compréhension de l'écrit (Reading Literacy)	9-10 ans 14-15 ans	Textes narratifs / Textes informatifs/ Documents	Verbatim/ Paraphrase/Idee principale/ Inférer/ Repérer/ Repérer et traiter l'information	3 échelles : textes narratifs/ textes informatifs/ documents	Pas de rotation. Une seule forme de test.	75 % à 90 % de QCM + quelques items ouverts à réponse « fermée ».
IALS (1994)/ Océ et Statistique Canada	Littératie	Adultes	Prose/ Documents/ Données chiffrées	-	3 échelles : prose/ documents/ données chiffrées	Plan complexe avec rotation de carnets	100 % de questions ouvertes

Etudes/ Commanditaire	Objet	Population	Domaines	Habiletés/ Processus	Echelles	Plan d'évaluation	Format des questions
PISA (2000)/ Ocdé	Littératie/ compréhension de l'écrit (Reading Literacy)	15 ans	Textes continus vs non continus	Retrouver de l'information/ Compréhension globale/ Tirer une interprétation/ Réfléchir sur le contenu/ Réfléchir sur la forme	Pas encore déterminé	Plan complexe avec rotation de carnets	55 % QCM, 45 % de questions ouvertes à réponse construite
PIRLS (2001)/ IEA	Littératie/ compréhension de l'écrit (Reading Literacy)		Buts de lecture (lire pour des motifs littéraires vs lire pour acquérir et utiliser de l'information)	Retrouver une information précise et explicite dans le texte/Faire une inférence directe/ Faire et justifier des inférences complexes et des interprétations/ Examiner et évaluer le contenu, la langue et les aspects textuels.	2 échelles : buts de lecture (littéraire vs information)	Plan complexe avec rotation de carnets	70 % de questions ouvertes à réponse construite, 30 % de QCM

Annexe 2
Définitions de la lecture ou de la littératie dans les évaluations internationales

Surveys	Définitions
IEA (1971)	Compréhension de la lecture, sans plus de précision.
IEARL (1991)	<p>« La compréhension de l'écrit (reading literacy) est la capacité de <u>comprendre</u> et d'<u>utiliser</u> les formes du langage écrit qui sont nécessaires à la vie en société et importantes pour l'individu ».</p> <p><i>« Reading literacy is the ability to <u>understand</u> and <u>use</u> those written language forms that are required by society and/or valued by the individual. »</i></p>
IALS (1994)	<p>« La littératie, c'est <u>utiliser</u> l'information imprimée et écrite pour fonctionner en société, atteindre ses objectifs personnels, développer ses connaissances et ses potentialités ».</p> <p><i>« <u>Using</u> printed and written information to function in society, to achieve one's goals, and to develop one's knowledge and potential. »</i></p>
PISA (2000)	<p>« Comprendre l'écrit, c'est non seulement <u>comprendre</u> les textes écrits, mais aussi les <u>utiliser</u> et <u>réfléchir</u> à leur propos pour atteindre ses objectifs personnels, développer ses connaissances, ses potentialités et participer à la vie en société ».</p> <p><i>« <u>understanding, using and reflecting</u> on written texts, in order to achieve one's goals, to develop one's knowledge and potential and to participate in society. »</i></p>
PIRLS (2001)	<p>« La compréhension de l'écrit (reading literacy) est la capacité de <u>comprendre</u> et d'<u>utiliser</u> les formes du langage écrit qui sont nécessaires à la vie en société et importantes pour l'individu. Les jeunes lecteurs sont amenés à construire le sens d'une diversité de textes. Ils lisent pour apprendre, pour participer à des communautés de lecteurs et pour le plaisir ».</p> <p><i>« Reading literacy is the ability to <u>understand and use</u> those written language forms required by society and valued by the individual. Young readers can <u>construct meaning</u> from a variety of texts. They read to learn, to participate in communities of readers and for enjoyment. »</i></p>

Annexe 3

Echelle IRT de l'enquête IEA Reading Literacy (Lafontaine, 1996).

		Population A distribution		Population B distribution		Niveau		
5		3 %	32 %					
6		97 %	68 %					
7								
8		81 %	55 %					
9								
10								
11		68 %	43 %					
12								
13								
14		50 %	25 %					
15								
16								
17		37 %	4 %					
18								
19								