

Pisa et la lecture: un point de vue de didacticien

Analyse critique de la validité didactique d'une enquête internationale sur la compréhension de l'écrit

Daniel Bain

Indépendamment des intentions des responsables PISA, il est important de se demander quelle est la validité – et donc l'intérêt – d'une telle opération du point de vue de la didactique. Laissant à d'autres le soin de défendre les aspects positifs de cette enquête, nous nous attacherons surtout à mettre en évidence ses limites sur le plan didactique, voire les obstacles que ce type d'investigation risque de susciter par rapport à certains efforts d'innovation actuels dans l'enseignement de la lecture. Nous nous demanderons en conclusion quels sont les effets secondaires de l'enquête et si le jeu en vaut la chandelle. Nos analyses, réflexions et interrogations s'appuieront essentiellement sur nos expériences en didactique du français langue maternelle, en collaboration avec des enseignants du Cycle d'orientation genevois (Secondaire I), ainsi que sur notre participation à deux enquêtes internationales (IEA, International Association for the Evaluation of Educational Achievement, Reading literacy, 1991; TIMSS, Third International Mathematics and Sciences Study, 1995).

Introduction

On pourrait nous reprocher d'emblée de détourner PISA (*Programme for International Student Assessment – Programme international pour le suivi des acquis des élèves*) de ses objectifs annoncés et explicites en lui prêtant des visées didactiques, voire simplement l'intention d'influencer directement l'enseignement. A lire ou à entendre les promoteurs de PISA, on constate qu'ils sont relativement restrictifs, prudents, voire allusifs quant à l'exploitation qu'on pourrait faire des résultats de l'enquête dans la pratique, sur le terrain. Ils en restent souvent à un objectif purement descriptif ou global: mesurer les compétences des élèves. Ils insistent sur le fait que le but de l'enquête «n'est pas de mettre en évidence des résultats par élèves, ni d'évaluer des programmes d'enseignement»; que «PISA peut seulement montrer de grandes tendances, des orientations générales dans les systèmes d'enseignement actuels» (OCDE, 1999, p. 26); que la littératie évaluée par l'enquête porte aussi sur des apprentissages faits en dehors de l'école (OCDD, 1999, p. 12).

On opposera à ces réserves plus ou moins défensives un certain nombre d'éléments qui laissent supposer qu'une exploitation des résultats pourrait bel et bien être faite pour juger de la qualité de la formation scolaire et pour chercher à l'améliorer, et ce entre autres par un réexamen des modalités de l'enseignement. Tout d'abord, cela semble effectivement une des intentions de l'opération si l'on en croit cet extrait de la préface du rapport technique sur l'enquête: «[Comparative international analyses] can also provide directions for national policy, for schools' curriculum and instructional efforts and for students' learning. Coupled with appropriate incentives, they can motivate students to learn better, teachers to teach better, and schools to be more effective» (Adam & Wu, 2002, p. 3). PISA a été clairement conçu pour fournir des *indicateurs* de performances utilisables dans le *pilotage* des systèmes de formation. Dans cette perspective, ce programme de recherche est donc censé apporter des informations pertinentes et exploitables pour la régulation du système de formation, en particulier en ce qui concerne les nécessités de revoir le *curriculum* au sens large (programmes, moyens et méthodes de formation, modalités de différenciation de l'enseignement et d'évaluation, formation des enseignants; cf. *Dictionnaire de l'évaluation et de la recherche en éducation: Curriculum*, de Landsheere, 1992). Ce type d'utilisation semble également postulé par la conclusion de Weiss (2002, p. 37) dans un article de *l'Éducateur*: «Les compétences en lecture, ainsi que ses différents niveaux de maîtrise définis dans le cadre de Pisa, pourraient bien être en effet un des apports scientifiques les plus significatifs de cette étude pour l'avenir et l'amélioration des systèmes de formation».

Par ailleurs, notons que les précautions prises par les responsables de PISA pour définir l'objet et la portée de leur enquête sont rapidement scotomisées quand il s'agit, au niveau des médias ou sur le terrain de la pratique, d'interpréter les résultats du test de lecture. C'est immédiatement le système de formation qui est mis en cause, et plus spécifiquement l'enseignement, voire les enseignants. Pour s'en persuader, il suffit de parcourir le recueil des articles sur PISA établi par l'Institut de recherche et de documentation pédagogique (Neuchâtel)¹. La conclusion, sommaire, de certains est que beaucoup trop d'élèves ne savent pas lire à 15 ans et que c'est la faute à l'école, à ses programmes ou à ses nouvelles méthodes d'enseignement du français.

Il nous paraît alors légitime d'analyser de façon critique les apports de PISA pour la didactique en insistant sur les limites de son exploitation pour l'amélioration de l'enseignement de la lecture. Dans cette perspective, nous nous demanderons notamment, dans un premier temps, quelle est la *validité* du test de compréhension de l'écrit élaboré pour PISA:

- sur le plan *conceptuel*, c'est-à-dire du *construct* auquel il se réfère: dans quelle mesure il évalue de façon adéquate à travers ses items des conduites de lecture socialement pertinentes, correspondant «à des situations les plus proches possible de la vie réelle» (Lafontaine, Baye & Matoul, 2001, p. 1);

- sur le plan *didactique*: dans quelle mesure les résultats (et notamment les scores) fournissent une image valable des compétences en lecture des élèves, un bilan, voire un diagnostic, permettant d'envisager les interventions ou remédiations nécessaires au niveau du curriculum et des modalités d'enseignement/apprentissage.

Validités conceptuelle, diagnostique et didactique se recouvrant en bonne partie (*Dictionnaire de l'évaluation et de la recherche en éducation: Validité*, de Landsheere, G., 1992), nous les prendrons généralement en considération simultanément dans la suite de nos analyses.

L'analyse critique qui suit se référera en arrière-plan à une pédagogie du texte d'inspiration socio-constructiviste développée dans le cadre du Cycle d'orientation genevois² et à un modèle du fonctionnement des discours élaboré par l'équipe de didactiques des langues de l'université de Genève (Bronckart, Bain, Schneuwly, Davaud & Pasquier, 1985; Bronckart, 1996). Ces cadres théorique et didactique mettent l'accent sur l'*activité langagière* comme moteur primordial de l'enseignement et de l'apprentissage et incitent à travailler avec les élèves dans le cadre de projets pédagogiques prenant la forme de séquences didactiques (Bain & Schneuwly, 1993; Dolz, Noverraz & Schneuwly, 2001; Schneuwly & Bain, 1993). Précisons encore que, pour la clarté du propos, nous avons adopté un point de vue spécifiquement (étroitement?) didactique et critique, sachant que d'autres se sont chargés – et se chargeront – de défendre l'intérêt de PISA pour d'autres utilisations.

Critique de la pertinence et de la validité des résultats de PISA d'un point de vue didactique

Cadre conceptuel

Ces dernières années, les responsables des enquêtes ont éprouvé la nécessité d'étudier de plus près ce concept un peu passe-partout de compétences. Ils ont en effet admis des insuffisances dans la réflexion théorique sur les dimensions mesurées par les enquêtes internationales («lack of an explicit, overarching conceptual framework based on broad theories of what skills, knowledge, and competencies are and how they relate to each other.»; OFS-OECD, 2000, p. 1). L'ambition est de définir des compétences clés, transversales, qui pourraient devenir des standards pour la formation. Le cadre conceptuel élaboré actuellement décrit un certain nombre de paramètres ou aspects (OCDE, 1999, notamment figure 1, p. 35) intervenant dans la compréhension de l'écrit et surtout dans son évaluation, mais n'aborde que très partiellement, ou par la bande, le processus de compréhension lui-même. En lisant le texte de présentation de PISA, on a l'impression d'une démarche très empirique: les données du test doivent non pas vérifier la validité d'un modèle mais faire en quelque sorte émerger celui-ci, dans

une démarche de type itératif (OECD, 1999, p. 44), à partir des «caractéristiques susceptibles d'influencer les performances des élèves, [...] composantes du processus de lecture qui feront l'objet de manipulations lors du survey de PISA sur la compréhension de l'écrit, afin de stimuler et d'évaluer la nature interactive du processus de lecture» (p. 26).

L'absence d'une théorie élaborée, organisant les différentes dimensions de la compréhension de l'écrit en tant qu'activité langagière et proposant un point de vue déterminé sur les opérations de lecture, constitue à nos yeux un premier obstacle majeur pour l'exploitation didactique des résultats, pour une contribution du programme PISA à une évaluation formative du système de formation sous l'angle pédagogique (Bain, 1995). Comme nous l'avons montré à propos de la production écrite (Bain & Schneuwly, 1993), diagnostic et remédiation impliquent une telle assise théorique. Une dissociation entre les approches du bilan (PISA) et des analyses à visée diagnostique et formative nous met dans une situation absurde de recourir à deux éclairages différents, voire divergents, sur une même réalité: les acquis (voire les difficultés) en lecture de nos élèves. Nous aboutissons à une certaine aporie qui nous semble peu problématisée par les milieux concernés.

En ce qui concerne l'opérationnalisation du cadre conceptuel dans l'évaluation, nous relèverons (comme c'était également le cas pour TIMSS) que le test est le résultat non d'un consensus sur un modèle permettant de valider les items proposés mais d'un compromis. Il s'agit en effet «de parvenir à un équilibre entre le développement des tâches qui reflètent le mieux la définition globale de la compréhension de l'écrit utilisée dans le cadre du programme OCDE/PISA, et la nécessité que ce matériel soit représentatif de la diversité linguistique et culturelle des pays participants» (OCDE, 1999, p. 28). Une telle approche empêche naturellement d'adopter une position théorique claire. Pour avoir une idée concrète des principes qui ont guidé la conception du test, il serait intéressant de disposer d'un compte rendu des décisions relatives à la sélection des items demandés aux différents groupes participant à l'élaboration de l'enquête. On notera en passant que finalement plus de 50% des items proviennent de pays anglophones (Moreau & Nidegger, à paraître, p. 1) et que le processus de sélection des items a pu être influencé fortement par le groupe d'experts chargé de superviser l'opération; mais dans quel sens?

Situations et activités de lecture

Selon la définition donnée par PISA de la *compréhension de l'écrit*, «cette capacité devrait permettre à chacun de réaliser ses objectifs, de développer ses connaissances et son potentiel, et de prendre une part active dans la société» (OCDE, 1999, p. 24). Pour mesurer de façon valide les compétences en lecture, et pour apprécier de façon fiable et crédible le niveau des élèves (notamment la proportion de médiocres lecteurs; cf. niveaux 1 et 0), il faut que «les compétences testées correspondent à des situations les plus proches possible de la vie réelle» (La-

fontaine, Baye & Matoul, 2001 p. 1). Ce n'est que très partiellement le cas, si nous nous en référons au matériel publié. En tout premier lieu, les lectures demandées ne sont pas mises en contexte; elles ne sont pas situées en tant qu'*activités langagières* effectives par rapport à une utilisation particulière dans un espace social défini, permettant de savoir plus précisément quel est l'enjeu, quelle est la tâche du lecteur et quelle signification elle peut prendre pour lui. Sur ce point, nous ne trouvons pas de traces très convaincantes des intentions des responsables de l'enquête selon lesquels: «En général, les questions et les consignes font référence à un but dont le lecteur est censé tenir compte au cours de sa lecture et de son interaction avec le texte³» (OCDE, 1999, p. 27). En fait, dans le test, «la notion de *situation* se réfère davantage à l'usage auquel l'auteur destine son texte qu'aux circonstances dans lesquelles le texte est effectivement lu» (OCDE, 1999, p. 27). Par ailleurs, la situation de test en général et plus particulièrement la succession de textes différents dans un même cahier de test rendent difficile toute mise en situation crédible ou efficace. La validité externe de l'évaluation en est affectée; il faut en tenir compte au moment d'interpréter les résultats, voire de chercher à les exploiter sur le plan didactique. On mettra en effet en contraste le type de lecture exigée par PISA et la pédagogie actuelle qui, dans la plupart des disciplines, s'efforce de donner du sens aux activités demandées aux élèves ou proposées pour leurs apprentissages. Quand elle ne peut les mettre véritablement en situation (dans des projets proches d'activités sociales effectives), cette approche didactique recourt à diverses formes de motivation ou de «fictionalisation» de l'activité à accomplir (Schneuwly & Bain, 1993).

Dans la «vie réelle», la situation authentique de lecture (il ne suffit pas que le *texte* soit authentique) est souvent bien différente de celle dans laquelle se déroule l'enquête PISA sur la compréhension de l'écrit. Prenons l'exemple du texte ou des documents à visée informative-explicative. Dans la plupart des cas, quand on lit (étudie) un document comme *Le lac Tchad*⁴ (constitué de deux figures précédées chaque fois d'un paragraphe de présentation, pseudo-légende), c'est dans une perspective particulière (historique, géographique ou zoologique) avec un objectif spécifique (saisir l'influence du climat sur l'apparition, le développement et la disparition de certaines espèces animales, par exemple) impliquant des interrogations préalables, donnant sa pertinence à telle question (raison de la disparition des hippopotames à telle époque; question 6) et écartant ou disqualifiant d'autres questions (la profondeur actuelle du lac Tchad; question 2).

En outre, la stratégie de lecture suggérée aux élèves par le matériel pour la plupart des exercices est apparemment de lire le texte d'abord, puis de répondre aux questions en se référant au texte. C'est en effet le comportement que semblent induire des consignes telles que les suivantes:

- *Servez-vous de l'éditorial intitulé «La technologie crée le besoin de nouvelles règles» de la page ci-contre [précédente] pour répondre aux questions suivantes (R236).*
- *Référez-vous au récit «Le cadeau» figurant aux trois pages qui précèdent pour répondre aux questions ci-dessous (R119).*

- Agnès Moreau, directrice du personnel d'une entreprise nommée ACOL, a rédigé à l'intention des membres du personnel de cette entreprise le communiqué qui figure aux deux pages qui précèdent. Référez-vous à ce communiqué pour répondre aux questions de cet exercice (R077).

L'élève aurait souvent avantage à lire les questions avant le texte (informatif ou narratif, par exemple) pour savoir à quoi s'intéressent particulièrement les auteurs du test et repérer les passages concernés. Les consignes semblent guider les réponses aux questions posées plutôt que la lecture du texte. De notre point de vue sur la compétence à lire un texte, il y a là plus qu'une nuance.

Certes, on peut admettre que les élèves lisent les textes proposés «en situation», mais dans une situation typiquement *scolaire*, au sens très restrictif de ce terme, avec toute sa charge d'arbitraire; impliquant également toute une part de conventions et d'habitudes propres à ce type de tests, que le testé doit savoir décoder. Il s'agit moins de montrer qu'on a compris ou apprécié le texte proposé (ce qui supposerait une perspective de lecture bien définie) que de donner les réponses attendues. Nous l'avons vu, cela implique notamment que l'on accepte ou supporte de passer d'un item à l'autre un peu comme on saute du coq à l'âne, c'est-à-dire sans qu'il y ait de lien évident entre les questions, sans qu'il y ait construction progressive de sens ou d'appréciation esthétique. Cette règle du jeu favorise les «bons élèves»: ils réussissent souvent à reconstituer «l'usage auquel l'auteur destine son texte»; d'autre part, bien au fait des stratégies de fabrication de l'excellence scolaire, ils saisissent l'enquête comme une occasion de manifester leurs capacités (Perrenoud, 1984). Sont défavorisés en revanche les adolescents moins scolaires, qui entrent plus difficilement dans l'arbitraire du jeu des lectures imposées par l'école, comme nous l'avons constaté lors d'enquêtes auprès de lecteurs précaires (Cifali, Schneuwly & Schubauer-Leoni, 1993; Pittard, 1990). Les résultats médiocres relevés dans certains de nos cantons ou dans certaines de nos filières et la proportion importante d'adolescents situés par le test aux niveaux inférieurs peuvent tenir en partie au fait que plusieurs des enquêtés ne sont pas entrés dans le jeu. En témoigne notamment dans un canton urbain comme Genève «le pourcentage élevé de non-réponses, voire de décrochage, au test (17%)» (Gabriel & Kaiser, 2002, p. 3)⁵.

Compréhension de l'écrit versus lecture de textes

On constate assez rapidement, en lisant le cadre contextuel (OCDE, 1999, chap. 1) et surtout en analysant les questions du test, que tout en parlant de *situations* et de *type de textes*, l'enquête donne *de facto* un statut très secondaire aux caractéristiques proprement discursives et textuelles. On peut comprendre les réticences ou la méfiance des auteurs à l'égard de certaines classifications de textes «créées avec une vision pratique plus que théorique». En revanche, nous avons de la peine à considérer comme théoriquement «bien plus importante» la distinction entre textes continus et non continus (ou documents) «qui fonde l'organisation de l'évaluation OCDE/PISA» (OCDE, 1999, p. 29) et qui prend le pas sur le

genre ou type de textes. Faute de repères quant à la situation de lecture (d'utilisation du texte) un «document» comme *Le lac Tchad*, paratexte séparé de son texte, est particulièrement artificiel. Comme nous l'avons constaté, les opérations de lecture sont isolées de tout contexte, de toute allusion à «l'usage auquel l'auteur destine son texte», de toute «référence aux personnes (et [...]aux objets) mis en en jeu par l'activité de lecture», contrairement à la notion de *situation* que PISA définit dans son cadre conceptuel (OCDE, 1999, p. 27). On ne peut plus parler véritablement de compréhension de *texte*; l'évaluation de la compétence de lecture semble s'autonomiser par rapport à la dimension textuelle.

Le cadre conceptuel fait certes allusion à une telle dimension en caractérisant les différentes tâches de lecture au niveau «macro», en particulier en définissant un des cinq aspects «associés à une pleine compréhension d'un texte»: *réfléchir sur la forme et la structure d'un texte* (OCDE, 1999, p. 34 et 38). Pour autant que nous puissions en juger à partir des textes et des items rendus publics, cette dimension n'est que partiellement prise en compte dans des cas où elle aurait eu toute sa pertinence, par exemple pour vérifier si l'élève peut s'aider de la structure propre au genre pour comprendre le texte (au sens étymologique: le saisir dans sa construction et son fonctionnement d'ensemble). A titre d'illustration, prenons une des questions sur la nouvelle *Le Cadeau*. Le récit se termine par «Et là, sur la terrasse, rongé jusqu'à l'os, se trouvait ce qui restait du jambon.» On demande à l'élève: *Pensez-vous que la dernière phrase du récit «Le cadeau» est une fin appropriée? Justifiez votre réponse en indiquant comment vous interprétez la relation entre cette phrase et la signification du récit.* La question nous semble pointer maladroitement un aspect effectivement crucial de ce type de nouvelle, en restant au niveau sémantique⁶. Parler dans les critères de correction de «touche finale apportée au thème du récit» nous apparaît une façon peu adéquate de faire allusion à la *chute* qui clôt le récit en le remettant en perspective, ce qui constitue une des caractéristiques formelles du genre. La connaissance ou la conscience de ce mode de fonctionnement de la nouvelle nous paraît présenter un intérêt particulier quant il s'agit de juger de la compétence à lire et à apprécier un texte. De même, pour contrôler la compréhension d'un texte explicatif (comme celui d'une encyclopédie pour les jeunes ou d'un manuel), il serait intéressant de vérifier si le lecteur peut s'aider de la structure plus ou moins conventionnelle de ce genre de texte.

De tout ce qui précède nous concluons volontiers que *compréhension de l'écrit* (PISA) et *lecture de texte* ne se recouvrent que partiellement, ce qui constitue pour nous une des limites à l'exploitation didactique de l'enquête. Comme le soulignent Moreau et Nidegger (à paraître, p. 1) le cadre conceptuel de l'enquête définit «une compétence générale de littératie dans le sens d'une compétence générale de lecture s'appliquant à toutes les situations et tous les contextes et non pas un modèle envisageant des compétences différenciées en lectures au pluriel». Prolongeant et paraphrasant une critique de Hofstetter et Schneuwly dans un article sur l'illettrisme (1993, p. 259), nous nous demandons si, dans plusieurs

exercices de PISA, nous n'avons pas affaire à un point de vue simplificateur qui réduirait la lecture à la compréhension d'énoncés. La structuration de ces énoncés en texte et le contexte discursif dans lequel ce texte se situe ne jouent souvent qu'un rôle secondaire si on considère la façon dont la majorité des questions guident la lecture des élèves.

Critères de correction

Un examen attentif des critères de correction nous amène à faire d'autres réserves: elles touchent les questions ouvertes. A la suite de critiques faites aux enquêtes antérieures (IEA, 1991, notamment), PISA a voulu leur donner plus de place (33% des items) dans le but «d'amener les élèves à mettre en œuvre des stratégies et des processus plus variés et de niveau taxonomique plus haut» (OCDE, 1999, p. 27). L'objectif est naturellement légitime, mais sa réalisation nous paraît sujette à caution dans plusieurs cas. A considérer les critères de correction, on constate qu'un crédit complet ou partiel (2 ou 1 points) peut être accordé à des réponses sommaires, allusives ou correspondant de façon discutabile à l'intention de la question et du texte. C'est le cas notamment lorsque la question porte sur la forme ou le style du texte. Par exemple:

Consigne (texte 077: Programme ACOL de vaccination volontaire contre la grippe): *Agnès a voulu donner un style amical et encourageant à ce communiqué. Pensez-vous qu'elle y a réussi? Justifiez votre réponse en vous fondant sur des détails précis de la présentation de ce communiqué, du style donné au texte [...].*

Réponse cotée 2 points: *Oui, le style de cet écrit est chaleureux et attrayant.* [Commentaire justifiant la cotation: *Utilise des termes personnels pour évaluer le style*]. Réponse cotée 1 point: *Non, c'est impossible de faire passer l'idée d'une piqûre de manière amicale et encourageante.*

Nous avons de la peine à trouver dans ces réponses la réalisation de la consigne (détails précis) et de l'objectif de la question: «Réfléchir sur la forme d'un texte: identifier les éléments qui mettent le style en accord avec le but d'un texte». Le mode de questionnement (écrit, non interactif) ne permet évidemment pas de prendre des informations complémentaires pour vérifier la qualité de compréhension que manifeste la réponse. L'option semble alors de privilégier la quantification d'une compétence de lecture (en acceptant une certaine imprécision ou un risque d'erreur d'estimation) plutôt que la pertinence de la réponse.

Modèle statistique et niveaux de lecture

Ceci nous amène à un autre obstacle d'importance, déjà relevé pour l'enquête IALS (International Adult Literacy Study; cf. Blum & Guérin-Pace, 2000): l'approche et le traitement statistiques choisis. Le modèle de Rasch (TRI: théorie de la réponse à l'item) a certes l'avantage de permettre à la fois d'estimer la difficulté des items et de produire des scores pour chaque élève. Il présente cependant un certain nombre de problèmes et d'inconvénients du point de vue qui est le nôtre. Ne serait-ce, en tout premier lieu, que le fait qu'il ne présente guère de transpa-

rence pour le non-spécialiste⁷, qui a quelque peine à estimer l'impact des contraintes et les conséquences du fonctionnement du modèle sur le choix des questions et la présentation ou l'interprétation des résultats. L'approche statistique de type psychotechnique prend comme référence primordiale les différences inter-individuelles: l'enquête a comme visée principale de comparer ou différencier les élèves et les groupes d'élèves (par pays notamment), et c'est par rapport à ce critère qu'on évalue ses qualités métrologiques. Cette approche s'oppose à une perspective éducatrice (De Landsheere, 1988; Demeuse, 2002); l'éducatrice vise en effet à mesurer plus spécifiquement des apprentissages par rapport à des objectifs pédagogiques déterminés et en référence à des résultats attendus. Le point de vue change: les objets d'étude ne sont plus les comportements des élèves; ceux-ci n'interviennent plus que comme des instruments pour réaliser des études sur le degré d'atteinte des objectifs, sur le degré de difficulté des exercices et des questions ou encore sur les progrès réalisés (Cardinet, à paraître). Dans l'enquête PISA, l'approche psychométrique (fondée sur la différenciation des élèves), l'obligation d'assurer aux items une bonne discriminativité, les contraintes de la comparaison internationale risquent bien de l'emporter sur les exigences critiques d'un cadre conceptuel définissant le contenu et la forme de tâches de lecture valides, socialement pertinentes. Pour en juger, il faudrait disposer d'un suivi des décisions prises tout au long de l'élaboration du test.

En lisant les commentaires aux résultats de PISA dans les médias, on constate qu'à côté des comparaisons internationales, les appréciations des résultats nationaux se réfèrent généralement *aux niveaux de compétences en compréhension de lecture (ou de «littératie»)* distingués par l'enquête (cf. par exemple Broi, Soussi & Wirthner, 2001, pp. 100-103). Malheureusement, dans certains articles de vulgarisation des résultats, la présentation de ces niveaux – et surtout de leur construction – est souvent sommaire, se limitant à un tableau décrivant en quelques lignes les différents niveaux. La formulation «Les élèves de ce niveau sont capables de...» donne l'impression d'avoir affaire à des catégories homogènes de compétences de lecture, clairement hiérarchisées. Mal informés, des enseignants pourraient imaginer qu'il serait possible de distinguer de façon fiable dans leurs classes à la fois des conduites et des groupes d'élèves d'un niveau déterminé. La tentation sera grande pour les responsables scolaires de transformer les niveaux les plus élevés en normes à atteindre par les élèves de telle ou telle filière à la fin de la scolarité obligatoire⁸.

Or, la constitution de ces niveaux, la définition de leurs limites sont le résultat d'une procédure psychométrique complexe, pas totalement transparente pour le commun des didacticiens ou des praticiens, et en partie arbitraire. Y interviennent à la fois des critères statistiques de nature probabiliste (probabilité de x% de réussir les items de tel niveau) et le jugement d'un groupe d'experts⁹. Comme le relèvent Lafontaine et ses collègues (2001, pp. 49-50) «Tout ce travail repose certes sur un certain arbitraire, éclairé toutefois par les connaissances

théoriques des experts. C'est ainsi que le nombre de niveaux que l'on choisit de définir et les endroits de coupure sont arbitraires: on peut définir 3, 4, ou 5 niveaux». Du fait de l'application du modèle statistique choisi, la notion de niveau atteint en lecture recouvre une certaine hétérogénéité de réussite: «Ainsi, un élève particulier se situant au niveau 3 aura par exemple réussi 65 à 70% des items de niveau 3, 80% des items de niveau 2 et 95% des items de niveau 1. En revanche, il n'aura réussi que 40% des items du niveau 4. C'est la raison pour laquelle on a considéré qu'il fallait le situer au niveau 3» (Lafontaine et al., 2001, pp. 50-51). Ce qui est confirmé par les observations de Moreau et Nidegger (à paraître, p. 3) sur un échantillon de données suisses romandes: «On constate [...] qu'un élève qui réussit un item de difficulté donnée peut échouer à des items plus faciles. En outre, un nombre important d'élèves de niveau faible (par exemple, niveau 1) réussissent des items de difficultés élevées (niveau 4 ou 5)» (Moreau & Nidegger, à paraître, p. 3). Les niveaux de compétences en compréhension de lecture sont donc loin d'avoir les propriétés d'une échelle hiérarchique comme celle de Guttman. Quelle interprétation peut-on alors leur donner sur le plan diagnostique dans la perspective d'une amélioration de l'enseignement? Certainement pas celle de certains articles de journaux qui déclarent, sur la base de cette classification en niveaux et des statistiques produites, qu'à Genève «un peu plus d'un élève sur cinq connaît de très grosses difficultés. Il lit, mais ne comprend pas vraiment le texte qui lui est proposé» (La Tribune de Genève, 01.02.2002).

D'autres analyses et critiques de Moreau et Nidegger, revisitant les résultats de PISA en Suisse romande, ne peuvent que nous alerter quant à l'exploitation des résultats de l'enquête dans le domaine didactique sur un plan *local* et *individuel*. Malgré les précautions prises, on constate par exemple des décalages parfois importants entre la calibration de certains items sur le plan international et le niveau de difficulté constaté sur l'échantillon suisse romand. D'autre part, certains constats au niveau des résultats individuels sont troublants et incitent à se demander ce que l'on mesure ou caractérise finalement avec les 5 (ou 6) niveaux de compétences. Effectivement, des analyses complémentaires sur les données suisses romandes montrent que:

des élèves avec des profils de réponses différents ont pu obtenir le même score. [...] Le score de l'élève ne dépend pour un cahier donné que du nombre d'items réussis (score brut) et non de la difficulté des items réussis. [...] Un élève peut être classé dans un niveau de lecture donné sans être confronté à des items de ce niveau. Par exemple, des élèves ont pu obtenir un niveau 5 sans être confrontés à des items de niveau 5. Tout dépend du nombre d'items réussis (Moreau et Nidegger, à paraître, p. 4).

Nous ne pouvons finalement que faire nôtres les conclusions et interrogations de Moreau et Nidegger (p. 4): «La démarche utilisée est complexe, les méthodes statistiques peuvent avoir parfois un aspect virtuel et très abstrait. La complexité des modèles utilisés rend les interprétations parfois difficiles. On peut se demander si les élèves faibles sont vraiment si faibles.» Le souci d'assurer une bonne com-

parabilité aux résultats et de garantir dans cette perspective une validité interne incontestable à la mesure nous semble avoir parfois joué au détriment de sa validité externe (mesurer les capacités de compréhension de l'écrit en situation authentique de lecture) et aux dépens de sa pertinence – donc de son exploitation didactique – sur le plan local: «Il est indispensable de se poser la question de la pertinence ou de la distance entre les instruments, conçus pour une comparaison à large échelle, et leurs exploitations locales afin d'éviter des interprétations abusives ou incohérences des résultats observés» (Moreau & Nidegger, à paraître, p. 4).

Conclusions: arrêts sur images et prolongements

Si l'on considère les contributions d'une enquête comme PISA sur le plan didactique ou plus largement pour la régulation du système de formation, on doit naturellement se poser la question: le jeu en vaut-il la chandelle? A court terme, pour caractériser les premières retombées de l'enquête, nous reprendrons deux métaphores médicales souvent utilisées dans les discussions actuelles.

Pisa comme thermomètre: un instrument utile pour le diagnostic?

On peut d'abord admettre que PISA a fonctionné à l'image d'un thermomètre en tant qu'instrument de contrôle du niveau des compétences en lecture. Mais il faut se rappeler que, selon les médecins, cet instrument ne mesure pas les températures mais les repère¹⁰ sur une échelle arbitraire. Les repères utilisés pour définir l'échelle sont précisément ceux fournis par la distribution des scores de l'ensemble des élèves participant au test (cf. la moyenne fixée à 500 et l'écart type à 100) et, dans la plupart des rapports, les résultats moyens d'autres pays. A partir de la température ainsi repérée, selon les cas, on rassurera sur la santé du patient en concluant «La Suisse au même niveau que la France» (Nidegger, 2001b, p. 36), ou au contraire on verra dans les résultats les premiers signes d'une fièvre maligne en annonçant que «les élèves genevois lisent moins bien que les Polonais» (*Le Temps*, 19.09.2002).

Ces indices de fièvre doivent-ils être pris au sérieux et susciter de nouvelles investigations à visée plus étiologique? En nous appuyant sur une approche plus clinique des problèmes de lecture et sur le témoignage des enseignants avec lesquels nous avons travaillé¹¹, nous admettons que les données corrélerent avec des difficultés que nous avons observées et que de nouvelles recherches s'imposent. Pour les aspects qui nous intéressent, nous doutons toutefois que des analyses complémentaires qui se limitent aux données de l'enquête apportent les éléments diagnostiques conduisant à des remèdes topiques et efficaces.

On peut sur ce point tirer un certain parallèle entre l'évaluation de la littératie par PISA et la mesure de l'intelligence par un test mosaïque (cf. Blum & Guérin-Pace, 2002, chap. 10: *Des tests d'intelligence à l'évaluation de l'illettrisme*). Dès que

l'on s'éloigne du facteur général en cherchant à le décomposer en sous-dimensions, on se hasarde dans des zones où les interactions sont multiples et la validité conceptuelle, voire la fidélité des données, commencent à être sujettes à caution. Déjà les trois échelles de compréhension de l'écrit *retrouver une information, développer une interprétation, réfléchir sur le texte et porter un jugement critique* ont un statut apparemment incertain: «D'après les concepteurs de PISA, il n'y aurait pas de hiérarchie du point de vue de leur difficulté. On peut toutefois constater une hiérarchie dans les résultats [...]» dans l'ordre des objectifs ci-dessus, sauf pour le canton de Genève» (Nidegger, 2002, p. 30). Autre avis en faveur d'une certaine hiérarchie, sur le plan théorique cette fois-ci: «La compréhension et l'interprétation sont des conditions nécessaires pour pouvoir réfléchir sur le contenu et la forme d'un texte et porter un jugement critique sur celui-ci» (Moser, 2001, p. 15). Pour d'autres chercheurs encore, ces «trois aspects correspondent assez largement aux notions de *sens littéral*, *sens inférentiel* et *sens personnel*» (Lafontaine et al., 2001, p. 47). Dès lors, quel intérêt didactique présente un «diagnostic» tel que «C'est dans les exercices nécessitant de la réflexion en vue de porter un jugement critique sur un texte que les élèves interrogés en Suisse ont rencontré le plus de difficultés [...]. Apparemment, l'importance accordée dans la langue d'enseignement à ces facultés complexes est moins grande en Suisse que dans d'autres pays» (Moser, 2001, p. 15)? Avons-nous vraiment là des éléments suffisants pour orienter des investigations ou des remédiations sur le plan didactique?

Des analyses sur les «types de texte» nous paraissent également peu prometteuses d'un point de vue didactique: la classification empruntée à la *grammaire de texte* de Wehrlich (1976; cf. OCDE, 1999, p. 30) ne nous fournit guère d'informations sur les activités et opérations langagières mobilisées par les différents textes, et donc sur l'origine des difficultés que peuvent rencontrer les élèves. En outre, mettre en corrélation le taux de réussite et la difficulté des textes (longueur, complexité lexicale et syntaxique) ou la formulation des questions (ouvertes ou à choix multiples) ne nous apportera que des résultats relativement évidents ou peu exploitables sur le plan didactique. Signalons en effet que, du fait des modalités de construction du test, il s'avère difficile de dissocier (étant donné leurs intercorrélations) types de textes, formats de présentation de ces textes (continus, non continus ou plus ou moins longs), modalités de questions-réponses et processus de lecture postulés (interpréter, rechercher une information, réfléchir sur le contenu et les qualités formelles d'un texte¹²) pour les mettre en relation avec les réussites ou échecs constatés dans différents ensembles ou sous-ensembles de répondants. On ne peut pas transformer en contrôle diagnostique une épreuve qui n'a pas été prévue spécifiquement à cet effet. Il nous apparaîtrait par conséquent dangereux d'orienter l'enseignement de la lecture à partir d'une épreuve qui a un tout autre but que d'établir un diagnostic.

Par ailleurs, en se référant aux *indicateurs contextuels* investigués par PISA, on risque aussi de se focaliser sur des facteurs très généraux (origine sociale et culturelle, structure scolaire, sexe), influençant la réussite scolaire en général. On se

heurtera également aux limites de validité du test signalées ci-dessus. On sera en outre limité par la fiabilité relative d'informations sur le contexte de formation reposant sur les appréciations personnelles des jeunes (Moser, 2001, p. 15) ou sur les descriptions et appréciations des directeurs des derniers établissements fréquentés par les élèves. On peut donc souhaiter qu'avant d'administrer des remèdes de cheval ou à large bande au système de formation, qu'avant d'essayer d'importer des recettes de Finlande ou d'ailleurs, on fasse dans nos écoles une étude plus clinique de la situation, de l'enseignement-apprentissage initial et continué de la lecture. Pour cela, il sera notamment nécessaire de solliciter la collaboration active des enseignants et des formateurs, d'analyser avec eux ce qui se fait – ou ne se fait pas – dans les classes en matière de lecture.

Pour ne prendre qu'un exemple au niveau secondaire 1, nous avons pu observer dans des classes de sciences l'usage qui était fait de textes dans une approche pédagogique visant à mobiliser l'activité des élèves. Nous avons alors été frappé de constater que certains d'entre eux pouvaient éviter de lire les textes proposés tout en accomplissant apparemment les tâches demandées. Les consignes de travail étaient données par écrit aux élèves (feuille + transparent affiché au rétro-projecteur)... mais lues et commentées également à haute voix par l'enseignant. A d'autres moments, dans le travail de groupe, certains élèves mettaient en œuvre des stratégies d'évitement efficaces: ils appelaient le maître pour lui dire qu'ils ne comprenaient rien à «se» textes. L'enseignant, en toute innocence, leur mâchait alors le travail de lecture en pointant et commentant dans les documents les passages pertinents. Dans tel autre groupe, une élève accomplissait à elle seule l'essentiel des tâches demandées: lisant les textes à haute voix, surlignant les informations correspondant aux questions posées, tout en assurant simultanément le secrétariat de l'équipe. Dans ces mêmes branches scientifiques, l'analyse des travaux de contrôle préparés par les enseignants montrerait probablement que peu d'entre eux se risquent à proposer aux élèves des consignes (formulations de problèmes) ou des textes dépassant la longueur de quelques lignes ou paragraphes, craignant les réactions négatives de certains élèves¹³. Des observations de ce genre devraient être entreprises pour essayer d'expliquer comment certains élèves peuvent sortir de l'école obligatoire avec des compétences de lecture très médiocres.

D'autres études encore, là où les informations nécessaires sont disponibles, devraient s'intéresser de plus près à la carrière scolaire antérieure des élèves médiocres lecteurs et à leurs aptitudes ou compétences dans d'autres domaines. Malheureusement, nous ne sommes pas sûr que les scores individuels fournis par PISA aient la fidélité nécessaire pour identifier les lecteurs précaires: le dispositif ne porte pas sur l'évaluation des élèves en tant qu'individus et ne garantit pas une bonne fiabilité à ce niveau, comme le montrent les analyses de Nidegger et Moreau (à paraître) mentionnées ci-dessus. De telles analyses avaient pu se faire dans de bonnes conditions sur les résultats de l'enquête IEA de 1991, chaque élève ayant été soumis aux 89 items du test de compréhension de lecture (Bain, 1993).

Elles nous avaient montré que pour beaucoup d'élèves les difficultés remontaient à la formation initiale et que les tentatives de remédiation (appuis divers, redoublement, passage par une classe spéciale) n'avaient pas apporté les effets attendus.

Il faudrait donc, plus généralement, «élucider pourquoi l'école n'a jamais réussi à résoudre de manière satisfaisante le problème de la lecture pour tous. Les causes de cet échec se trouvent sans doute autant du côté des pratiques culturelles [...] que dans les pratiques scolaires» (Schneuwly, 1993).

D'un point de vue didactique, on ne peut qu'espérer que PISA encouragera des recherches avec une telle visée. On attendra de celles-ci – ainsi que d'investigations plus spécifiquement didactiques – qu'elles affinent le diagnostic au-delà d'une prise de température et permettent de dépasser les remèdes très généraux, voire simplistes, proposés actuellement. Telles ces mesures extraites des «premières recommandations de suivi» de PISA, présentées par un expert: «[...] Instaurer des périodes quotidiennes de lecture silencieuse en classe tout au long de la scolarité obligatoire. Peut-être complété à partir d'un certain âge par la lecture et la discussion du journal. Entraîner à la lecture rapide et efficace. [...] Élaborer à l'usage des maîtres et des maîtresses des outils faciles à utiliser pour le diagnostic différencié des compétences lectrices des élèves¹⁴. Favoriser le plaisir et l'envie de lire chez les enseignants. [...]» (*Le Temps*, 19.09.2002).

PISA comme sismothérapie: quels effets secondaires?

Recourant à une autre image médicale, on a parfois mentionné l'effet d'*électrochoc* provoqué par les résultats de PISA dans certaines régions de la francophonie. Le courant a-t-il passé et secoué tous ceux qui sont concernés par le problème? En fait, comme le note Weiss (2002) dans un article intitulé *Pisa, ce que chacun y a vu*, les réactions à l'enquête et les interprétations des scores obtenus par les divers cantons ont été très diverses. Faute d'explications convaincantes des résultats mis en évidence, faute de diagnostic probant sur les causes des difficultés observées, PISA a fonctionné un peu comme une épreuve projective: beaucoup y ont trouvé une justification à ce qu'ils étaient en train de faire ou avaient déjà prévu de faire.

On peut alors craindre certains effets secondaires de cette sismothérapie. Tout d'abord, un effet de légitimation tous azimuts, comme on vient de le voir, favorisant directement ou indirectement une certaine inertie par rapport aux initiatives à prendre dans le domaine de l'enseignement de la lecture. A ce jeu, les projets des uns seront plus légitimes que ceux des autres, qui auront moins accès aux décisions. Si les enseignants ne s'organisent pas pour faire valoir leur point de vue, leurs besoins ou leurs solutions, on peut craindre qu'ils aient surtout le rôle d'exécuteurs des mesures décidées par autrui.

On peut redouter aussi l'utilisation de ce que Weiss (2002, p. 37) appelle les «produits dérivés» de l'enquête, «en l'occurrence la détermination, au niveau international, des compétences attendues des élèves de tous les pays au terme de leur formation de base, repères vers lesquels chaque système éducatif devrait

tendre, indépendamment de leur curriculum, de leur organisation et de leurs méthodes.» Outre les réserves sur la validité conceptuelle et didactique du test, nous pouvons craindre le processus suivant, très habituel dans le fonctionnement de nos systèmes scolaires (Bain, 1979) et renforcé par la mise en place du projet Harmos¹⁵. Des contrôles dits certificatifs seront probablement instaurés, en particulier fin 9e, s'inspirant de très près du test de compréhension de l'écrit de PISA. De facto, ces examens prendront une fonction sélective: les niveaux de maîtrise définis dans l'enquête risquent de servir à placer les élèves dans les filières du postobligatoire, en particulier à écarter certains d'entre eux des orientations gymnasiales, alors même que la validité prédictive de ces évaluations n'aura pas été établie¹⁶. Soucieux d'aider leurs élèves en prévision de cette échéance importante dans leur carrière scolaire, les enseignants seront poussés à les préparer spécifiquement à de telles épreuves. On aura alors un effet en retour bien connu des moyens d'évaluation sur les pratiques pédagogiques, favorisant des exercices sur la *compréhension de l'écrit* plutôt que des activités de lecture de texte en situation. Ce que Weiss qualifie de «bénéfices scientifiques» du programme PISA pourrait bien correspondre, dans notre perspective, à une régression dans le domaine didactique.

En continuant à filer notre métaphore médicale, nous émettrons encore des doutes quant aux bénéfices à tirer, sur le plan didactique, du suivi du patient au moyen des prises de températures ultérieures, celles-ci visant à analyser et expliquer son évolution à travers les «indicateurs de tendance qui deviendront disponibles en raison du caractère cyclique des données» (OCDE, 1999, p. 12). «Ces études devraient nous permettre de mieux comprendre l'évolution du rendement et les facteurs qui la sous-tendent» (Statistiques Canada, 2001, p. 16). Les passations de 2003 et 2006 comporteront un nombre plus restreint d'items consacrés à la compréhension de l'écrit, une plus grande place étant faite aux mathématiques puis aux sciences. Les mesures de progrès, entachées d'une double erreur, étant toujours d'une fiabilité très relative (De Landsheere, 1992, *Changement (Mesure du)*), nous pensons qu'on ne pourra guère dépasser la simple estimation d'une *tendance* et qu'il sera difficile d'en diagnostiquer avec précision et sûreté les facteurs déterminant l'évolution du rendement. Il est d'ailleurs bien possible qu'une amélioration se manifeste du simple fait qu'à la suite du battage organisé autour de PISA, les autorités scolaires, les enseignants, voire les élèves, prendront plus au sérieux le test qu'en 2000; sans parler du risque d'un certain bachotage du test.

Alors, finalement, le jeu en vaut-il la chandelle? On ne pourra juger de l'efficacité (rapport coût – bénéfice) du programme PISA que vers 2010 au vu des effets globaux ou spécifiques de l'opération dans les domaines de la politique de l'éducation, de l'enseignement et des compétences en lecture de nos élèves. Si l'enquête donne une impulsion aux travaux de didactique relatifs à l'apprentissage – initial et continué – de la lecture, le bénéfice de l'opération sera non né-

gligeable. Mais on a vu que, dans le domaine qui nous intéresse, le test et les questionnaires passés en 2000 n'apportent que peu d'éléments diagnostiques orientant des interventions ciblées et des remèdes spécifiques. En outre, les frais inhérents à une telle enquête sur une dizaine d'années permettront-ils d'investir encore dans des recherches proprement didactiques? Enfin, tôt ou tard, PISA devra probablement affronter et surmonter un autre effet secondaire de la sismothérapie: «La phase excito-motrice, contemporaine du passage du courant, est suivie d'une phase résolutive avec sommeil» doublée d'une amnésie partielle (Desgrez, 2000).

Notes

- 1 Nous remercions l'IRDP de nous avoir permis de consulter ce recueil.
- 2 Les divers travaux réalisés par la commission Pédagogie du texte, dans le cadre du Centre de recherches psychopédagogiques du Cycle d'orientation genevois, sont mentionnés dans la brochure bibliographique de C. Pesenti (1998).
- 3 On constate en considérant la modalisation de cette phrase (En général, censés) que sur ce point les responsables ne s'avancent pas trop.
- 4 Les exemples cités correspondent aux textes et aux questions rendus publics; ils peuvent être téléchargés notamment sur le site PISA belge: <http://www.agers.cfwb.be>.
- 5 On peut se demander si la réussite des élèves fréquentant des écoles en milieu plutôt traditionnel comme à Fribourg, en Valais ou en Bavière ne tient pas à une certaine acceptation de ce genre d'exercice et d'évaluation favorisée par le contexte social et éducatif.
- 6 «OBJECTIF DE LA QUESTION: Réfléchir sur la forme d'un texte: évaluer l'usage fait par l'auteur d'une caractéristique particulière du texte.»; cf. site belge <http://www.agers.cfwb.be>.
- 7 Dans les remarques qui suivent, nous nous référerons notamment aux analyses récentes de Moreau et Nidegger sur les données suisses romandes (à paraître).
- 8 N.B. Cette hypothèse de notre part n'est pas purement gratuite; cf. infra, note 15, le projet Harmos.
- 9 Pour une discussion critique sur le rôle des experts dans de telles enquêtes, nous renvoyons le lecteur au chapitre «De la confiance dans les experts», in Blum & Guérin-Pace, 2000, p. 14 ff.
- 10 Cf. la définition du Petit Robert (édition électronique du Nouveau Petit Robert, 1996), sous Température: 3. [...] *On doit se référer à certaines propriétés physiques* (dilatation, changement de résistance électrique, variation de pression des gaz ou des vapeurs, modification dans la viscosité des fluides, etc.), *afin de définir une échelle arbitraire de température* (abrév. °), *grâce à laquelle on peut repérer la température d'un corps. Abusivt Mesure d'une température, son rapport à une échelle arbitraire divisée en degrés** (=> thermomètre).
- 11 Cf. les travaux de la commission Pédagogie du texte, en particulier les *Cahiers de la Section des Sciences de l'Éducation no 40 et 52*, le no 3/1993 de la revue *Éducation et recherche* et Bain, Erard & Séchaud, 1994.
- 12 Pour ne prendre qu'un exemple, les réponses ouvertes sont majoritaires dans les tâches consistant à réfléchir sur un texte, alors que quand il s'agit de rechercher l'information, le test recourt à une majorité de questions fermées.
- 13 Hypothèse qui nous est suggérée notamment par une discussion récente avec des enseignants de sciences ainsi que par nos observations dans nos travaux de didactique des sciences au Cycle d'orientation genevois.

- 14 Nous doutons fortement qu'on puisse actuellement trouver ou élaborer des instruments à la fois «faciles à utiliser» et fournissant un «diagnostic différencié [et fiable] des compétences lectrices».
- 15 Projet d'«Harmonisation de la scolarité obligatoire [...]». But: définir des niveaux de compétences pour les langues, les mathématiques et les sciences qui doivent être atteints à la fin de la 9^e, 6^e et 2^e année de scolarité. Ces niveaux feront l'objet d'un Accord intercantonal sur l'harmonisation de la scolarité obligatoire», EducationCH, 03/06.2002.
- 16 Cf. la mise en garde que Cardinet (1977) a faite à propos de la confusion entre les fonctions sommative (certificative) et prédictive (orientation-sélection) des évaluations du type de celles que nous sommes en train de discuter.

Références bibliographiques et ouvrages consultés

- Adam, R. & Wu, M. (2002). *PISA 2000 Technical Report*. Paris: OECD (diffusé sur le site <http://www.pisa.oecd.org>).
- Bain, D. (1979). *Orientation scolaire et fonctionnement de l'école*. Berne & Francfort: Peter Lang.
- Bain, D. (1993). Les «mauvais» lecteurs au Cycle d'orientation. *Education et Recherche*, 15, (3) 1993, 260-278.
- Bain, D. (1995). Vers une évaluation formative des systèmes de formation: pistes pour une recherche efficiente sur l'efficacité de l'école. In U.P. Trier (ed.): *Wirksamkeit von Bildungssystemen – Analyse de l'efficacité des systèmes de formation – Analysing the Effectiveness of Education Systems*. Aarau: NFP33, SKBF/CSSRE.
- Bain, D., Erard, S. & Séchaud, M. (1994). Comment repérer et aider les lecteurs précaires au premier cycle secondaire. D'un test de lecture à l'intervention en classe. *LIDIL, Revue de linguistique et de didactique des langues*, Université Stendhal, Grenoble, 11, 73-88.
- Bain, D. & Schneuwly, B. (1993). Pour une évaluation formative intégrée dans la pédagogie du français: de la nécessité et de l'utilité des modèles de référence. In Allal, L., Bain, D. & Perrenoud, Ph. (dir.), *Evaluation formative et didactique du français* (pp. 51-79). Neuchâtel & Paris: Delachaux et Niestlé.
- Blum, A. & Guérin-Pace, Fr. (2000). *Des lettres et des chiffres. Des tests d'intelligence à l'évaluation du «savoir lire», un siècle de polémique*. Paris: Fayard.
- Broi, A.-M., Soussi, A. & Wirthner, M. (2001). Les résultats des élèves en compréhension écrite. In Chr. Nidegger, *Compétences des jeunes romands. Résultats de l'enquête PISA 2002 auprès des élèves de 9^e année* (pp. 97-122). Neuchâtel: Institut de recherche et de documentation pédagogique.
- Bronckart, J.-P. (1996). *Activité langagière, textes et discours. Pour un interactionisme socio-discursif*. Lausanne: Delachaux et Niestlé.
- Bronckart, J.-P., Bain, D., Schneuwly, B., Davaud, Cl. & Pasquier, A. (1985). *Le fonctionnement des discours. Un modèle psychologique et une méthode d'analyse*. Neuchâtel-Paris: Delachaux & Niestlé.
- Cardinet, J. (1977). *Objectifs éducatifs et évaluation individualisée*. Neuchâtel: Institut romand de recherche et de documentation pédagogique, 77.05.
- Cardinet, J. (à paraître). Pourquoi faut-il parler d'éduométrie? *Bulletin de l'ADMEE-Europe*.
- Cifali, M., Schneuwly, B. & Schubauer-Leoni, M. L. (1993). *Education et Recherche*, 15, (3), 1993, 279-309.
- De Landsheere, G. (1992). *Dictionnaire de l'évaluation et de la recherche en éducation*. Paris: Presses Universitaires de France.
- De Landsheere, V. (1988). *Faire réussir, faire échouer. La compétence minimale et son évaluation*. Paris: Presses Universitaires de France.

- Demeuse, M. (2002). Edumétrie et psychométrie. *Bulletin de l'ADMEE-Europe*, no (2), 3-4.
- Desgrez, H. (2000). Thérapeutique – Physiothérapie: Electrothérapie, *Encyclopedia Universalis*, version électronique no 6.
- Dolz, J., Noverraz, M. & Schneuwly, B. (2001). *S'exprimer en français. Séquences didactiques pour l'oral et pour l'écrit. Notes méthodologiques, vol. IV*. Bruxelles: De Boeck, Corome.
- Gabriel, F. & Kaiser, Cl. (2002). ...Qu'en est-il des résultats aux tests de PISA. *Cap C'O*. Département de l'instruction publique genevois, octobre 2002, 2-4. (diffusé sur le site www.geneve.ch/co/capco).
- Lafontaine, D. (2002). Au-delà des performances des jeunes de 15 ans, un système éducatif se profile... *Le Point sur la Recherche en Education*, n° 24, juin 2002 (diffusé sur le site <http://www.agers.cfwb.be>).
- Lafontaine, D., Baye, A. & Matoul, A. (2001). *PISA 2000: Programme international de l'OCDE pour le suivi des acquis des élèves. Présentation de l'enquête*. Université de Liège: Faculté de Psychologie et des Sciences de l'Education, Département Education et Formation, Pédagogie théorique et expérimentale.
- Moreau, J. & Nidegger, Chr. (à paraître). PISA: Vers une définition de compétences universelles? Symposium ADMEE S3.1 – Evaluation des compétences pour piloter le système: l'exemple de PISA. *Actes du congrès ADMEE/SSRE 2002*. Lausanne: ISPFP & URSP.
- Moser, U. (2001) *Pilotage du système de formation en Suisse. Préparés pour la vie? Les compétences de base des jeunes – Synthèse du rapport national PISA 2000*. Neuchâtel et Berne. OFS/CDIP (diffusé sur le site <http://www.statistik.admin.ch>).
- Nidegger, Chr. (Coord.). (2001a). *Compétences des jeunes romands. Résultats de l'enquête PISA 2002 auprès des élèves de 9e année*. Neuchâtel: Institut de recherche et de documentation pédagogique.
- Nidegger, Chr. (2001b) Résultats internationaux et suisses. In Chr. Nidegger (Coord.), *Compétences des jeunes romands. Résultats de l'enquête PISA 2002 auprès des élèves de 9e année* (pp. 29-38). Neuchâtel: Institut de recherche et de documentation pédagogique.
- Nidegger, Chr. (2002). Les acquis des jeunes Romands. Exploitation locale de l'enquête internationale PISA 2000. *Psychoscope*, 23, (8).
- OCDE (1999). *Mesurer les connaissances et compétences des élèves. Un nouveau cadre d'évaluation*. Paris: OCDE.
- OFS-OECD (2000). *Introduction to DeSeCo*. Neuchâtel: Office fédéral de la statistique (diffusé sur le site http://www.statistik.admin.ch/stat_ch/ber15/desecco/intro.htm).
- Perrenoud, Ph. (1984). *La fabrication de l'excellence scolaire: du curriculum aux pratiques d'évaluation: vers une analyse de la réussite, de l'échec et des inégalités comme réalités construites par le système scolaire*. Genève: Droz.
- Pesenti, C. (1998). *CRPP, centre de recherches psychopédagogiques: bibliographie 1963-1998*. Genève, DIPCO, Direction générale du Cycle d'orientation.
- Pittard, J. (1990). *Et pourtant ils lisent*. Genève: Centre de recherches psychopédagogiques du Cycle d'orientation genevois, Commission Pédagogie du texte.
- Schneuwly, B. (1993). Mauvais lecteurs – Que faire? ou Le regard du didacticien de la langue maternelle. *Education et Recherche*, 15, (3), 280-286.
- Schneuwly, B. & Bain, D. (1993). Mécanismes de régulation des activités textuelles: stratégies d'intervention dans les séquences didactiques. In Allal, L., Bain, D. & Perrenoud, Ph. (dir.), *Evaluation formative et du didactique du français*. (pp. 219-238) Neuchâtel & Paris: Delachaux et Niestlé.
- Statistiques Canada (2001). *À la hauteur: La performance des jeunes du Canada en lecture, en mathématiques et en sciences Étude PISA de l'OCDE – Premiers résultats pour les Canadiens de 15 ans*. Ottawa: Statistiques Canada / Conseil des ministres de l'éducation (Canada) (diffusé sur le site www.pisa.gc.ca).
- Wehrlich, E. (1976). *A Text Grammar of English*. Heidelberg: Quelle and Meyer.
- Weiss, J. (2002). Pisa, ce que chacun y a vu. *Educateur*, 9/2002, 36-37.

PISA und die Lesekompetenz: der Standpunkt eines Didaktikers. Kritische Analyse der didaktischen Gültigkeit einer internationalen Studie

Zusammenfassung

Ungeachtet der Absichten der für PISA verantwortlichen Instanzen sollte man die Frage nach der Gültigkeit eines solchen Unterfangens – und den damit verbundenen Interessen – aus didaktischer Sicht stellen. Wir überlassen es anderen, die positiven Aspekte dieses Programms zu verteidigen, und beschränken uns darauf, seine Grenzen für eine didaktische Nutzung, beziehungsweise seine möglichen negativen Folgen für die gegenwärtigen Veränderungsbestrebungen im Lesunterricht aufzuzeigen. Wir stellen abschliessend die Frage, welche sekundären Auswirkungen eine solche Untersuchung haben kann, und ob die Sache der Mühe überhaupt wert ist. Unsere Analysen, Überlegungen und Thesen stützen sich auf unsere Erfahrungen im Bereich der Didaktik Französisch als Muttersprache, in Zusammenarbeit mit Lehrern des Cycle d'orientation in Genf (Sekundarstufe I), sowie auf unsere Beteiligung an zwei internationalen Studien (IEA, *International Association for the Evaluation of Educational Achievement, Reading Literacy*, 1991; TIMSS, *Third International Mathematics and Sciences Study*, 1995).

PISA e la didattica. Analisi critica della validità didattica di un'inchiesta internazionale sulla comprensione dello scritto

Riassunto

Indipendentemente dalle intenzioni dei responsabili di PISA, è importante chiedersi quale sia la validità – e dunque l'interesse – di una tale operazione dal punto di vista didattico. Lasciando ad altri il compito di difendere gli aspetti positivi di questa inchiesta, mostreremo i suoi limiti sul piano didattico, in particolare taluni ostacoli che comporta in relazione agli sforzi innovativi attualmente messi in atto nel campo dell'insegnamento della lettura. Ci chiederemo infine quali siano gli effetti secondari dell'inchiesta e se il santo vale la candela.

Analisi, riflessioni e interrogativi proposti si basano essenzialmente sulle nostre esperienze nella didattica del francese lingua materna, esperienze fatte in collaborazione con gli insegnanti del Cycle d'orientation genevois (secondario I), così come sulla nostra partecipazione a due inchieste internazionali (IEA, *International Association for the Evaluation of Educational Achievement, Reading literacy*, 1991; TIMSS, *Third International Mathematics and Sciences Study*, 1995).

Pisa and Reading Literacy: a didactical point of view. A critical analysis of the didactical validity of an international survey

Summary

Whatever the intentions of the initiators of PISA, it is important to examine the validity – and thus the interest – of such an operation from the didactical point of view. We leave it to others to defend the positive aspects of this programme, seeking here only to highlight the limits of its usefulness for didactics – perhaps even the obstacles which this type of investigation might erect with respect to current efforts to improve the teaching of reading. We will conclude by asking about the secondary effects of the study and whether the effort is worth the investment.

Our analysis, reflections and questions are based essentially on our experience in didactics of French as a mother tongue, in collaboration with teachers of the *Cycle d'orientation* in Geneva (lower secondary school) and on our participation in two other international surveys *IEA, International Association for the Evaluation of Educational Achievement, Reading Literacy, 1991*; *TIMSS, Third International Mathematics and Science Study, 1995*).