

PROJET SPECIAL 2

ÉVALUATIONS INTERNATIONALES: IMPACTS POLITIQUES, CURRICULAIRES ET PLACE DES PAYS FRANCOPHONES

Michèle Artigue, LDAR, Université Paris Diderot - Paris 7, France

Charles Coagri Nassouri, DGIFPE, Burkina Faso

Hikma Smida, Université Tunis El Manar, Tunisie

Carl Winsløw, IND, Université de Copenhague, Danemark

Résumé : Dans cette contribution, nous rendons compte du travail réalisé au colloque EMF2009 dans le cadre du projet spécial 2 dont le thème était : « Evaluations internationales : impacts politiques, curriculaires et place des pays francophones ». Nous présentons d'abord la problématique de ce projet et l'organisation des quatre séances associées lors du colloque. Nous poursuivons par un compte-rendu séance par séance, en reprenant les contributions écrites reçues des intervenants. Nous terminons enfin par un bilan du travail réalisé.

Mots clefs : comparaisons internationales, évaluations internationales, mathématiques, curricula

Introduction

L'enseignement des mathématiques occupe une place centrale dans tous les systèmes d'enseignement du monde, dès les premières années de la scolarité. De plus, en dépit de variations indéniables dans les méthodes et les contenus enseignés d'un pays à l'autre, il se prête particulièrement bien aux comparaisons internationales et, de fait, les études comparatives : évaluations internationales à grande échelle des performances des élèves dans ce domaine, mais aussi recherches secondaires attachées à ces évaluations ou recherches associées à des comparaisons plus locales entre deux pays par exemple, se sont multipliées ces dix dernières années.

Parmi ces études comparatives, nous privilégions dans ce projet spécial les évaluations à grande échelle réalisées dans le cadre d'organismes internationaux comme l'IEA (International Association for the Evaluation of Educational Achievement) ou l'OCDE, et notamment le programme PISA de l'OCDE mis en place en 2000 (voir <http://www.pisa.oecd.org>) dont le but est « d'évaluer dans quelle mesure les jeunes adultes de 15 ans, c'est à dire des élèves en fin d'obligation scolaire, sont préparés à relever les défis de la société de la connaissance¹. »

L'un des objectifs majeurs de ces évaluations est en fait de fournir aux décideurs des pays participants des informations susceptibles de les aider à piloter leurs systèmes éducatifs dans le sens d'une plus grande « qualité » et « efficacité ». Même si les enjeux sont variables selon les pays, nul ne peut nier que ces évaluations et leurs résultats sont de plus en plus utilisés dans les débats concernant l'enseignement des mathématiques, ni qu'elles ont une influence croissante sur les politiques éducatives et les choix curriculaires. Dans le même temps, ces évaluations et/ou leur exploitation font l'objet de critiques sévères au sein des communautés directement concernées par l'enseignement des mathématiques (cf. par exemple (Hopman,

¹ OCDE(2004) : Apprendre aujourd'hui pour réussir demain-premiers résultats de PISA 2003, p20.

Brinek & Retzl, 2007²). Il nous semble donc important de nous interroger sur ces objets au sein de l'espace mathématique francophone, en exploitant la diversité des contextes existant au sein de cette espace pour nourrir la réflexion. C'est l'objet de ce projet spécial.

Nous avons choisi de structurer le travail de ce projet spécial autour de trois thèmes en précisant chacun des thèmes par une liste de questions. Le premier thème concerne le fonctionnement de ces évaluations internationales et il s'agit à travers les réponses apportées de permettre à tous les participants de partager l'information nécessaire au travail envisagé sur ces évaluations et leur impact. Le second thème concerne l'impact politique et curriculaire des évaluations internationales, le troisième leur impact sur l'enseignement et la recherche. Les questions sont formulées dans le cadre général des évaluations internationales mais le programme PISA dans lequel de nombreux pays de l'espace francophone sont engagés sera plus particulièrement visé.

1°. Le fonctionnement des évaluations internationales

- Quelle sont les visées (politiques et officielles) et les modes d'organisation des évaluations internationales ?
- Sur quelles bases théoriques et méthodologiques s'appuient-elles ?
- Quels sont leurs principaux résultats, en particulier sur les compétences mathématiques des élèves ?
- Quelles sont les critiques qui leur sont faites ?
- Y a-t-il des caractéristiques de ces évaluations susceptibles d'avoir un impact particulier sur les pays francophones ?

2°. L'impact politique et curriculaire des évaluations internationales dans les pays francophones.

- Comment s'effectue la diffusion des résultats ? Comment sont-ils traités dans les médias ?
 - Les résultats des évaluations internationales sont-ils source de débats au niveau national ? Si oui, de quelle nature sont ces débats et qui impliquent-ils ?
 - Les résultats des évaluations internationales entraînent-ils des réactions politiques et si oui de quelle nature ? Des décisions politiques récentes concernant l'enseignement des mathématiques ont-elles été au moins en partie motivées par ces évaluations ?
 - Dans les réformes récentes du curriculum ou celles en cours concernant les élèves ou la formation des enseignants, peut-on percevoir une influence plus ou moins directe des évaluations internationales, et si oui, laquelle ?

3°. Impact sur l'enseignement des mathématiques et sur la recherche en didactique dans les pays francophones.

- Observe-t-on des effets des évaluations internationales sur l'enseignement des mathématiques ? Y a-t-il par exemple évolution des formes et contenus de l'évaluation

² S. Hopman, G. Brinek, M. Retzl (Eds) (2007). *PISA according to PISA*. Wien: Lit Verlag.
<http://www.univie.ac.at/pisaaccordingtopisa>

ou incorporation de tâches similaires aux items PISA dans l'enseignement, dans les manuels ? Si oui, quels en sont les effets ?

- Les évaluations internationales ont-elles suscité des travaux de recherche spécifiques ? Si oui, de quelle nature et quels en sont les résultats ?
- Les évaluations internationales ont-elles contribué plus généralement au développement de recherches comparatives en didactique ? Si oui, quelles sont les recherches de ce type menées et quels en sont les résultats ?
- Y aurait-il des travaux spécifiques à mener dans l'espace mathématique francophone concernant les évaluations internationales et si oui lesquels ?

La première séance du projet spécial a été consacrée au thème 1. Après une courte présentation générale faite les organisateurs du projet, Antoine Bodin a fait un exposé centré sur l'organisation et le fonctionnement des comparaisons internationales à grande échelle, en comparant les évaluations TIMSS³ et PISA de ce point de vue. Cet exposé a été suivi d'une phase de discussion où se sont succédés commentaires, questions, et apports d'information. La deuxième séance a été plus particulièrement centrée sur le thème 2. Elle a pris la forme d'une table ronde dont les participants : Lucie DeBlois du Canada, Nadia Mawfik du Maroc, Charles Nassouri du Burkina Faso et Hikma Smida de Tunisie, reflétaient la diversité des contextes présents dans l'espace mathématique francophone. Cette table ronde a, elle aussi, été suivie d'une phase de discussion générale. La troisième séance a été consacrée à l'impact sur l'enseignement et la formation des évaluations à grande échelle. Pour cette séance, le travail collectif a été préparé par un exposé de Richard Cabassut et une réaction de Lucie DeBlois. Enfin, la dernière séance a été consacrée à l'impact sur les recherches en didactique des mathématiques. Elle s'est appuyée sur une méta-analyse d'études comparatives réalisée par Michèle Artigue et Carl Winslow, et considérant plus particulièrement d'une part des recherches secondaires associées ou motivées par les enquêtes internationales comme la TIMSS-Video Study, The Learner Perspective Study ou l'East-West ICMI Study⁴, d'autre part des comparaisons entre deux pays centrées sur un thème précis, réalisées dans le cadre de travaux de thèse. La fin de cette séance a été consacrée à une discussion sur les besoins de recherche dans ce domaine et les prolongements collectifs possibles au niveau de la communauté francophone du travail réalisé dans le cadre de ce projet spécial. Nous présentons ci-après les contributions écrites reçues pour ces actes.

³ TIMSS : Trends in Mathematics and Science Study.

⁴ Voir les références dans la contribution associée.

SEANCE 1 : INTRODUCTION DU PROJET SPECIAL

Comme précisé ci-dessus, l'objet de la première séance était d'introduire le projet spécial en se centrant sur l'organisation et le fonctionnement des évaluations internationales à grande échelle telles que TIMSS et PISA. Nous reproduisons ci-après l'exposé d'Antoine Bodin qui a porté cette introduction.

Organisation et fonctionnement des comparaisons internationales à grande échelle, avec un accent particulier sur PISA⁵

Antoine Bodin⁶, IREM de Marseille

Introduction

Cette communication cherche à préciser les origines, les buts, les démarches et les effets, des études internationales portant sur les mathématiques aux niveaux scolaires.

Compte tenu de l'importance prise par les études conduites par l'OCDE (PISA), nous avons été amenés à leur donner une place importante. Nous avons laissé de côté les questions de nature politique, idéologique, épistémologique, ... qui sont liées à ces études, ainsi que celles qui sont relativement indépendantes du ou des domaines concernés : questions méthodologiques et logistiques, recueil et traitement des données, contrôle de la validité des questions...

Nous avons aussi laissé de côté la question des relations que ces études permettent d'établir entre les acquis des élèves et les variables relatives aux contextes (contexte socio-économique, contexte éducatif), ainsi qu'avec les variables personnelles relatives aux élèves. L'incidence de ces questions sur la qualité et l'intérêt des évaluations du domaine mathématique ne doit pas sous-estimée, mais, dans cette communication, nous chercherons à nous centrer sur les questions mettant directement en jeu les mathématiques.

1. Généralités sur les études internationales

Dans le domaine de l'éducation, les études internationales ont une longue histoire. En cinquante ans, les domaines d'études se sont diversifiés, le nombre de pays y participant est passé d'une petite douzaine à une centaine. Les études portent aussi bien sur les acquis scolaires (TIMSS, PIRLS) que sur la capacité des citoyens à affronter les problèmes de la vie courante (PISA, ALL). Les domaines d'étude traditionnels sont les mathématiques, la lecture et les sciences, mais, depuis quelques années on voit émerger d'autres sujets : langues, citoyenneté, géographie,...

Dans la plupart des pays, en résonance avec ces études et souvent en s'appuyant sur des cadres

⁵ Le diaporama présenté à Dakar ainsi qu'une version plus complète de cette contribution contenant de nombreuses références (50 pages) sont disponibles à l'adresse suivante

http://web.me.com/antoinebodin/pro/ETUDES_INTERNATIONALES_-_PISA.html

⁶ Antoine Bodin a été membre des groupes d'experts mathématiques des études internationales TIMSS et PISA.

de référence proches, on voit se mettre en place des évaluations nationales sur échantillons ou portant sur l'ensemble des élèves d'un ou plusieurs niveaux scolaires.

Même les pays qui ne participent pas directement à ces études sont influencés par elles. En effet, les institutions internationales (EMS, Banque Mondiale,...) leur demandent de mettre en place des évaluations nationales, ce qui les amène inévitablement à emprunter aux cadres de référence des études internationales.

Pour ces études, plusieurs organisations internationales selon les cas coopèrent ou sont en concurrence :

- l'IEA (Association internationale pour l'évaluation des résultats scolaires, basé aux Pays-Bas),
- l'ETS (*Educational Testing Service*, basé aux USA),
- l'OCDE (Organisation de Coopération et de Développement Économiques, dont le siège est à Paris).

Il est à noter que les études menées par ces organismes, en relation étroite avec les ministères de l'éducation des pays concernés, répondent à une demande sociale et politique, qui concerne maintenant l'ensemble des pays et des organisations internationales. Les questions proprement pédagogiques ou didactiques commencent à émerger, mais, dans l'ensemble, il s'agit essentiellement de fournir des indicateurs de pilotage aux responsables des systèmes d'enseignement.

Certaines opérations répondent à un souci de connaissance alors que d'autres émettent des jugements de valeur sur la qualité des systèmes éducatifs. Ce dernier cas signifierait que certaines valeurs seraient reconnues d'une façon universelle, valables de la même façon pour l'ensemble des systèmes éducatifs, ce qui n'est pas sans poser quelques questions.

1.1. Évaluation des connaissances - évaluation des systèmes ?

Plus que l'établissement de constats, ces études cherchent à identifier des points d'entrées susceptibles de faciliter, selon les pays et les systèmes, les améliorations jugées nécessaires.

Reste à savoir si le modèle est consistant, si les études sont valides (i.e. si elles évaluent bien ce qu'elles prétendent évaluer) et si l'usage qui en est fait est bien de nature à aider à l'amélioration des systèmes. Disons de suite que les avis sur cette question sont partagés.

1.2. La place des mathématiques dans les évaluations internationales

Les mathématiques semblent constituer la discipline privilégiée des évaluations nationales et internationales. Le mot « mathématique » est généralement utilisé comme désignant un domaine unifié, bien identifié, et dont chacun s'accorderait à penser que son étude et une certaine maîtrise seraient essentielles à toute formation.

Les mathématiques sont supposées générer une sorte de trait latent qui permettrait de classer de façon univoque aussi bien les individus que les systèmes éducatifs. Nombre d'études s'appuient implicitement sur une conception unidimensionnelle de la compétence mathématique (le niveau !). La recherche et les observations faites dans le cadre des évaluations à grande échelle, ainsi que le rappel de la grande variété des formes de l'activité mathématique devrait faire passer de cette conception naïve à une perspective

multidimensionnelle.

Dans nombre de pays, et pas seulement dans la francophonie, les mathématiciens et les membres de la communauté mathématique, élargie aux enseignants de cette discipline, considèrent les évaluations nationales et internationales avec beaucoup de méfiance. Ils leur reprochent bien souvent un manque de validité (c'est-à-dire de manquer leur objet) et de donner une idée fautive de la discipline.

1.3. L'objet évalué : quelles mathématiques ?

Différents aspects des mathématiques peuvent être concernés par l'évaluation.

A - Les mathématiques pratiques : Il s'agit des mathématiques reconnues comme utiles à tous, pour la vie de tous les jours, ... (littéracie mathématique au sens de PISA⁷).

Compte tenu des choix qui sont souvent faits et des méthodes d'investigation, il est difficile de parler ici de bases, et encore moins de socle. Base ou socle de quoi ? Pourquoi ?

B - La culture mathématique : On pourrait dire mathématiques pour tous, au sens de « l'honnête homme » moderne, qu'il ne saurait être question de réduire au citoyen-consommateur-employable. Il n'est pas question ici de techniques ou de formalisme mathématique, mais d'une dimension culturelle trop souvent absente de nos programmes de formation.

C - Les mathématiques spécialisées : Il s'agit des connaissances, théories et méthodes mathématiques spécifiques sur lesquelles peuvent se greffer des développements ultérieurs.

Le caractère multidimensionnel de l'ensemble des connaissances et compétences relatives à ces trois catégories apparaît clairement. La maîtrise des savoir-faire pratiques n'ouvre pas nécessairement la porte à la culture au sens B ni aux compétences sollicitées en C. Il est toutefois évident, qu'une partie des connaissances et savoir-faire de la première catégorie est indispensable à l'exercice de la troisième, la difficulté est de savoir laquelle et, pour la partie concernée, de savoir jusqu'à quel niveau d'approfondissement.

Par ailleurs, à l'intérieur même de chacune de ces catégories, on peut distinguer des sous-domaines : algèbre, géométrie, analyse, statistiques... De nombreuses observations montrent que, en ce qui concerne les élèves de l'enseignement secondaire ayant en principe suivi le même programme d'enseignement, les réussites dans ces sous-domaines sont assez faiblement corrélées. On peut aussi identifier des types différents de processus cognitifs mis en jeu (reconnaissance et synthèse par exemple) et remarquer que les capacités à faire fonctionner ces différents types de processus sont eux-mêmes faiblement corrélés entre eux.

1.4. Les sujets de l'évaluation

Les niveaux habituellement visés sont les 4^e, 8^e, et 12^e années d'enseignement (en France, CM1, Quatrième et classes terminales), plus, avec PISA, le niveau 15 ans (fixé sur une classe d'âge et non sur un niveau de scolarité).

2. Objet des études, domaines évalués,...

⁷ Le mot anglais « literacy » a été malencontreusement traduit en français par « culture », ce qui est source de nombreux malentendus.

Les cadres de référence relatifs aux contenus ont le plus souvent constitué en une mise à plat des découpages traditionnels du corpus mathématique, obéissant en fait à une logique de syllabus. On trouve classiquement : Nombres et algèbre ; Géométrie ; Fonctions et Analyse ; Statistiques et probabilités. En fait, chacun de ces sous-domaines se dérive de différentes façons suivant les études et des pans entiers des mathématiques scolaires n'y trouvent pas facilement leur place. Par exemple, où placer la question des grandeurs ? les mathématiques discrètes ? Les différentes traditions scolaires s'en arrangent plus ou moins bien, mais les comparaisons inter-curriculaires sont difficiles.

De son côté, PISA s'éloigne des découpages classiques des contenus, pour faire place à une organisation de type qu'il est convenu, en France, d'appeler « par problématiques »⁸. Trois classes de problèmes y sont définies et étudiées : Quantité ; Variations et relations ; Espace et formes ; Incertitude. Cette centration sur les types de problèmes, plutôt que sur les contenus, est favorable à la production de questions dans le traitement desquelles les mathématiques interviennent comme outil.

Le cadre de référence de PISA, le plus récent descripteur et prescripteur de démarche évaluative, est sans doute aussi le mieux élaboré, mais essentiellement dans la perspective « littéracie mathématique » qu'il définit en même temps.

2.1. Les types de questions

Les QCM : Dans les études internationales, en opposition avec les conceptions et les pratiques les plus courantes en France, la prédominance, voire le caractère exclusif des questions à choix multiple a longtemps été la règle. De plus, ces QCM sont en général des questions pauvres, à réponse unique.

Les QROC (Questions à Réponses Ouvertes et Courtes)⁹ : À partir de TIMSS, les études internationales se sont partiellement ouvertes aux QROC (lesquelles représentent 40% des questions dans le cas de PISA). L'ouverture ne porte cependant que sur la réponse et non sur la question. De plus, « court » peut simplement signifier que l'élève est appelé à écrire lui-même une réponse, par exemple numérique, sans justification. Il reste donc peu de place pour la rédaction, l'expression des idées, l'explicitation des démarches, la justification des résultats.

Les QRR (Questions à Réponses Rédigées)¹⁰ : L'utilisation de questions qui pourraient justifier l'appellation QRR est très rare dans les études internationales et dans la plupart des études nationales. C'est évidemment le souci de fidélité, et l'obsession liée à la construction d'échelles aux qualités qui voudraient être indiscutables, qui a conduit à privilégier les QCM dans les études internationales. Les insuffisances des QCM sont cependant largement reconnues. Outre leur tendance à privilégier certains types de connaissances et de compétences, leur validité est souvent discutable et l'est d'autant plus que l'on prétend faire une évaluation authentique (aussi proche que possible des conditions d'utilisation « normales » des connaissances).

⁸ Voir les problématiques de l'APMEP élaborées sous l'égide de Régis Gras.

⁹ *Short answer questions.*

¹⁰ *Long answer ou extended answer questions.*

D'un point de vue éducatif (qui est en fait celui de la psychométrie), les qualités recherchées de fidélité et de validité jouent de façon antagoniste. Tant que l'on privilégie les échelles et que l'on a une conception unidimensionnelle du domaine à évaluer, c'est la fidélité qui l'emporte, et donc les QCM. Lorsque les réflexions épistémologique et didactique parviennent à se faire entendre, le souci de validité prend sa place et les études s'ouvrent à des questions à réponses ouvertes. C'est ce qui s'est largement passé pour les études internationales depuis TIMSS, malgré la difficulté évidente que représente le codage des réponses de ce type dans un contexte multiculturel et multilinguistique.

3. L'étude PISA

3.1. Origine et buts de l'étude PISA (Qui ? Pourquoi ?)

L'acronyme PISA désigne le programme international pour l'évaluation des élèves mis en place par l'OCDE depuis l'année 2000. Les 30 pays de l'OCDE plus un certain nombre de pays dits partenaires (58 pays en tout en 2006, 63 en 2009, représentant 87% de l'économie mondiale !) participent tous les trois ans à une évaluation commune des compétences de base de tous les jeunes de 15 ans (encore scolarisés), à quelque place qu'ils se trouvent dans les systèmes éducatifs concernés. Il s'agit d'évaluer, de façon indépendante des programmes d'enseignement (les curriculums), la façon dont les jeunes sont prêts, vers la fin des scolarités obligatoires, à affronter les défis du monde dans lequel ils sont appelés à vivre.

Avec PISA, l'OCDE s'adresse en premier lieu aux décideurs et aux gestionnaires auxquels elle fournit des indicateurs pour le pilotage des systèmes éducatifs. Cela justifie, dans une certaine mesure, le nombre important des indicateurs produits et l'attention apportée à leur qualité technique. Dans cet article, nous n'évoquerons que quelques-uns de ces indicateurs et nous renverrons à d'autres études pour ce qui concerne les indicateurs de nature économiques, socio-économiques, sociaux, ainsi que les relations de ces indicateurs avec les indicateurs plus directement liés à l'éducation et aux résultats de l'éducation.

Précisons les objectifs de PISA en utilisant les termes utilisés par l'OCDE.

« L'enquête PISA vise à évaluer dans quelle mesure les jeunes adultes de 15 ans, c'est-à-dire des élèves en fin d'obligation scolaire, sont préparés à relever les défis de la société de la connaissance. Cette orientation reflète l'évolution des finalités et des objectifs des programmes scolaires : l'important est d'amener les élèves à utiliser ce qu'ils ont appris à l'école, et pas seulement à le reproduire. » (OCDE 2004 - Apprendre aujourd'hui, réussir demain Premiers résultats de PISA 2003)

Ce qui est évalué est la littéracie, que l'OCDE définit ainsi :

« ...la notion de « littéracie¹¹ », ...renvoie à la capacité des élèves d'exploiter des savoirs et savoir-faire dans des matières clés et d'analyser, de raisonner et de communiquer lorsqu'ils énoncent, résolvent et interprètent des problèmes qui s'inscrivent dans divers contextes. » (idem)

Enfin, ce qui concerne les mathématiques :

¹¹ Littéracie et non "culture" comme cela est souvent écrit en français. Le terme anglais *literacy* ne correspondant en rien à l'idée que l'on se fait de la culture dans la plupart des...cultures !

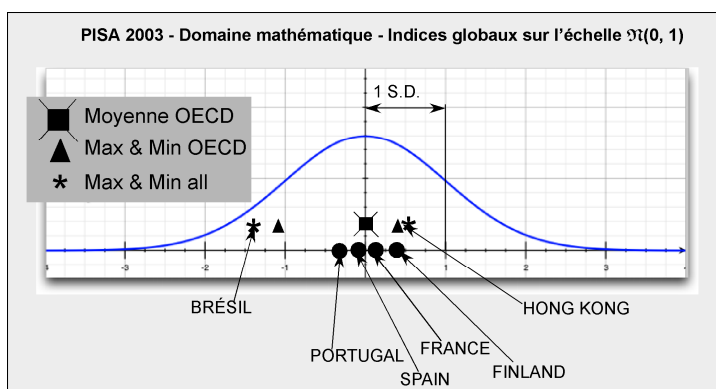
« La littéracie mathématique est l'aptitude d'un individu à identifier et à comprendre le rôle que les mathématiques jouent dans le monde, à produire des jugements fondés sur les mathématiques, et à s'engager dans des activités mathématiques, en fonction des exigences de sa vie en tant que citoyen constructif, impliqué et réfléchi. ¹² » (ibidem)

Il est clair que les objectifs pris en compte par PISA ne recouvrent pas les objectifs du système éducatif français. Dans un article précédent (Bodin, 2005), nous avons pu estimer que, pour les mathématiques, le questionnement de PISA recouvrait 15 à 20% des programmes du collège.

L'organisation de PISA est lourde et complexe ; nous renvoyons aux documents cités en référence, à notre diaporama et à notre article développé pour plus de détails.

3.2. Comprendre les scores

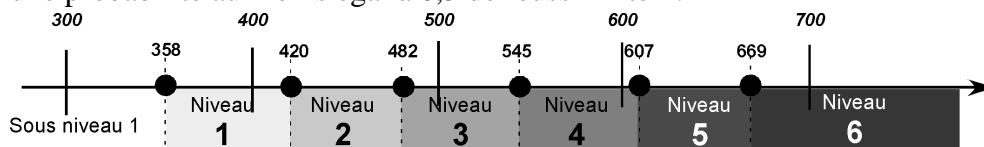
Les résultats des études internationales sont rapportés à une échelle qui reste mystérieuse pour la plupart de ceux mêmes qui l'utilisent. En effet, que signifie la phrase suivante : « *En mathématiques, en 2003 ; le score de la France était 511, tandis que le score de la Finlande était de 544, soit un écart de 32 points....* » ? Sur une échelle d'amplitude apparente de 1000 unités, 32 points, est-ce beaucoup ou est-ce négligeable ?



Pour résumer une méthodologie complexe, disons qu'après un certain nombre d'opérations, un

indice de réussite de l'ensemble des jeunes de 15 ans de l'OCDE est ajusté à la loi normale $N(500; 100)$ et l'on peut alors replacer les résultats d'un pays sur cette échelle (en se restreignant aux résultats de ce pays).

Voir plus haut un exemple de présentation possible des indices de quelques pays. Insistons sur le fait qu'à ce niveau, on a totalement perdu de vue les vrais scores de réussite aux questions. La seule chose que l'on puisse dire de l'écart, par exemple, entre la France et la Finlande est qu'elle est de 33 centièmes d'écart-type sur l'échelle ainsi construite. Conformément aux techniques issues de la psychométrie (théorie des réponses aux items), PISA définit alors l'indice de difficulté d'un item comme étant la valeur de l'indice à partir duquel un individu a une probabilité au moins égale à 0,5 de réussir l'item.



L'échelle de compétence de PISA

Cette démarche a permis à PISA de définir des niveaux de compétence. Par exemple, un élève est au niveau 6 s'il a un indice de compétence égal ou supérieur à 669, tandis qu'un item est

¹² Traduit de l'anglais par l'auteur. La définition habituellement donnée en français est incorrecte (et stupide !).

au niveau 6, si la probabilité d'un élève de niveau 6 de réussir cet item est supérieure ou égale à 0,5.

La façon dont la construction de cette échelle prend en compte l'analyse des tâches permet de donner un sens à ces niveaux et permet de les décrire (cf. Bodin 2005) On peut alors, par exemple, comparer les proportions d'élèves qui, dans chaque pays, se trouvent à tel ou tel niveau de compétence.

3.3. Les mathématiques dans PISA

Nous avons vu que PISA avait une approche utilitaire des mathématiques et se demande dans quelles classes de problèmes les compétences mathématiques pourront s'utiliser. Ce ne sont pas vraiment les connaissances dans les domaines mathématiques classiques (géométrie, algèbre, ...) qu'il s'agit d'évaluer, mais plutôt la façon dont ces connaissances peuvent être mobilisées dans des situations relevant d'une analyse non scolaire des besoins.

Par ailleurs, PISA considère 3 grandes classes de compétences : Reproduction, Connexions et Réflexion. Ces classes servent, simultanément, à l'analyse des compétences (ce qu'il faut être en mesure de mettre en œuvre dans tel type de situation), à la création et à l'analyse des tâches proposées.

PISA, dans son souci de coller au réel, utilise abondamment le langage courant pour présenter les questions d'évaluation. Dans certains cas, la compréhension de textes non mathématiques, devient la principale difficulté que les élèves ont à affronter. Bien sûr cela fait partie de l'activité mathématique normale, mais le vrai travail mathématique ne commence qu'une fois cette étape franchie.

Dans PISA l'idée de démonstration est totalement absente, les processus d'explication et de justification sont très peu pris en compte. De plus, la symbolisation et l'abstraction y sont réduites au minimum. Cela fait une grande différence avec les conceptions françaises habituelles concernant les acquis mathématiques.

4. Principaux enseignements des évaluations internationales et nationales

Que l'on s'intéresse aux curriculums, aux acquis des élèves, ou aux relations entre les variables de contexte et celles, personnelles, relatives aux élèves, les études permettent de mettre en évidence des différences importantes, mais aussi des similarités entre les pays ou les contextes éducatifs.

Dans l'ensemble, les évaluations s'accompagnent de réductions importantes des enjeux curriculaires assumés par les systèmes éducatifs sur ce qu'il apparaît possible d'évaluer¹³. Sauf à considérer que le cadre de référence de telle étude s'imposerait comme standard universel et absolu pour tous les systèmes éducatifs, il n'est pas acceptable de considérer les résultats obtenus comme des indicateurs généraux de l'efficacité globale des systèmes éducatifs concernés.

Précisons ici certaines des réductions opérées :

¹³ C'est l'« écrasement des objectifs sur l'évaluation » dénoncée en France par Guy Brousseau depuis les années 80.

- Pour TIMSS, il s'agit de se limiter à un curriculum central (*core curriculum*) supposé être commun aux pays participants. Des pans entiers des compétences et des connaissances développées dans chaque pays sont ainsi laissés de côté.
- Pour PISA, le point de départ ayant été les besoins supposés du citoyen, il est clair qu'une partie importante des objectifs habituellement poursuivis dans les différents systèmes éducatifs n'a pas été prise en compte. Dans le cas Français, nous avons estimé à seulement 15% la part du curriculum couvert par PISA (A. Bodin, 2005). Il est toutefois exact que ces 15%, s'ils manquaient à la formation d'un individu seraient de nature à l'handicaper plus sérieusement que l'absence des 85% restants.

En ce qui concerne les mathématiques, on peut envisager de s'intéresser à plusieurs types de résultats produits par ces études. Il s'agit des connaissances que ces études apportent :

- sur les acquis des élèves,
- sur l'état des curriculums à travers le monde et sur les différences curriculaires,
- sur les relations entre les acquis des élèves et les curriculums correspondants,
- sur les relations entre les acquis des élèves et les différents facteurs, sociaux, contextuels et personnels, qui les accompagnent.

Pour cela, on peut déjà répartir les pays en trois groupes :

- Groupe A : quatre ou cinq pays du continent asiatique obtiennent régulièrement des résultats nettement plus élevés que tous les autres pays : la Corée, Singapour, Hongkong, Taiwan et le Japon.
- Groupe B : Les pays en voie de développement ou sous-développés, lorsqu'ils participent, obtiennent des résultats bas ou très bas. Cela montre surtout qu'il ne suffit pas d'imposer un curriculum exigeant pour que l'enseignement suive et pour que les apprentissages correspondants se fassent. Nous laisserons de côté ce cas qui, à l'évidence, a davantage à voir avec la question du développement qu'avec notre propos.
- Groupe C : À côté de ces cas extrêmes, il reste la plupart des pays d'Europe, l'Amérique du Nord et la Russie. Ces pays présentent suffisamment de points communs, aussi bien en ce qui concerne les diverses variables contextuelles qu'en ce qui concerne les résultats des élèves, pour que les comparaisons directes présentent quelque intérêt.

D'une façon générale, dans les trois groupes de pays, les différences varient selon les domaines et sous-domaines concernés (numérique, géométrie,..), ce qui peut informer sur les effets des curriculums sur les résultats. Dans l'ensemble, on observe surtout que le niveau des résultats dans un domaine est positivement lié à l'imposition aux contenus correspondants.

Les concepts importants résistent partout. Par exemple, dans l'étude TIMSS 95, un des items les plus faciles du champ « proportionnalité » demandait ce qui se passerait si, dans une classe dont les trois-cinquième des élèves étaient des filles, on ajoutait 5 filles et 5 garçons. Pour l'ensemble des pays, moins des deux-tiers des élèves ont identifié la bonne réponse.

Dans la plupart des cas, les compétences manifestées par les élèves dans l'exécution des tâches proposées sont bien inférieures à celles qui pourraient être inférées de l'analyse des programmes, des commentaires officiels, des manuels, ou... des sujets d'examen.

4.1. Les acquis des élèves en France

Dans PISA 2003, la moyenne des scores des pays de l'OCDE est de 58,2% tandis que la moyenne des scores en France est de 58,8%. Sachant que les pays de l'OCDE incluent des pays qui rencontrent des difficultés importantes qui les classeraient plutôt dans le groupe B ci-dessus, il n'y a pas lieu de pavoiser.

Les études PISA (2000 et 2003) montrent en particulier, et cela est confirmé par les autres études, que, en France :

- Les élèves ont des connaissances, mais que ces connaissances sont peu disponibles. Pour la plupart d'entre eux, si on ne leur dit pas explicitement quelles connaissances mathématiques particulières il convient d'utiliser dans une situation donnée, ils ne le trouveront pas d'eux-mêmes, même s'ils possèdent le ou les éléments de connaissance correspondants.
- Les élèves font preuve de peu d'autonomie. Ils ne s'attaquent qu'aux questions qu'ils pensent pouvoir résoudre, ils ne disposent pas de stratégies pour aborder un problème qui ne leur est pas familier : essayer, expérimenter, bricoler... ne font pas partie de leurs modes d'approche possible.
- Les élèves sont légèrement défavorisés par les QCM. Cela, surtout parce qu'ils ne répondent pas n'importe comment lorsqu'ils ne savent pas. Ce point devrait donc plutôt être considéré comme positif. L'attribution de points négatifs pour réponse erronée serait plus correcte et supprimerait cet inconvénient.

Si l'on s'intéresse aux mathématiques pour tous, PISA ne fait que confirmer ce que l'étude TIMSS 1995 montrait à l'évidence : les élèves qui ne feront pas de mathématiques spécialisées partent mal armés pour la vie, pour comprendre le monde qui les entoure, pour participer aux divers débats de société qui les attendent. Des efforts de tous types devraient être consentis si l'on veut faire face à cette situation inquiétante.

Lorsque l'on évoque les résultats français dans les études internationales, il est souvent évoqué l'idée que ces études seraient culturellement biaisées. Il serait plus exact de dire qu'elles sont curriculairement biaisées, ce qui vaudrait aussi, plus ou moins, pour tous les pays participants. Les biais culturels, lorsqu'ils existent, sont de nature à relativiser la faiblesse des résultats, non à les nier¹⁴.

Malgré l'importance accordée dans notre curriculum à l'expression écrite, à l'argumentation et à la preuve, les résultats que l'on peut observer dans ces domaines sont très loin des attentes.

Les études internationales ne permettent aucune comparaison sur ce qui constitue nos spécificités nationales, telles que la place des aspects formels des mathématiques, de la preuve et de l'argumentation. Il faudrait que des études complémentaires soient menées sur ce point (au niveau européen par exemple). Pour l'instant rien ne prouve que nos faiblesses absolues soient aussi des faiblesses relatives. Tous les systèmes éducatifs éprouvent des difficultés à faire coïncider leurs ambitions avec les résultats observés. Il serait important que des mises à

¹⁴ Le lecteur trouvera sur les sites de l'APMEP et de la SMF un diaporama présentant l'ensemble des questions de PISA qui ont été libérées, avec les résultats obtenus en France et dans quelques autres pays. Il pourra ainsi se faire une idée personnelle du caractère éventuellement culturellement biaisé de ces questions.

plat soient faites qui permettent de baser des actions sur l'état réel de la situation et non sur un imaginaire dépassé.

5. Retombées des études

5.1. En général

Les études et les évaluations dont nous parlons ont toutes pour objectif déclaré l'amélioration de la connaissance des résultats de la formation des élèves pour, finalement, contribuer à l'amélioration de cette formation.

Cette amélioration passe en général par des actions sur le contexte, ou, du moins, sur la prise en compte différenciée de ce contexte (incitation politique à l'amélioration des conditions de vie, discrimination positive,...) qui, malgré leur importance, seront laissées hors du champ de cet article.

Les retombées directes de ces études portent essentiellement sur des modifications des programmes d'enseignement, la formation des enseignants et les instructions qui peuvent leur être données.

5.1.1. Retombées sur les programmes d'enseignement

Les programmes d'enseignement constituent la variable sur laquelle il est le plus facile d'agir ; on ne s'étonnera donc pas que, assez souvent, les retombées les plus visibles aient lieu à ce niveau et que parfois elles s'y arrêtent.

Le pilotage du curriculum par l'évaluation a largement été dénoncé dans les années 80 (Freudenthal, Westbury,..). Il conduit à adapter les programmes aux résultats et donc à diminuer les attentes, ouvrant la voie à un cycle d'ajustements successifs vers le bas.

Il est toutefois des retombées qui peuvent sembler positives, telles que l'introduction généralisée de questions liées à l'incertitude (statistiques et probabilités). Des pays qui avaient un enseignement extrêmement formel et procédural ouvrent leurs curriculums à la résolution de problèmes non stéréotypés, plus ouverts sur l'aspect outil des mathématiques.

La place de la géométrie et en particulier de la géométrie de l'espace est souvent revisitée en relation avec ces études.

5.1.2. Retombées sur la formation des enseignants

Moins évidente et plus importante est la retombée sur la formation initiale et continue des enseignants.

Au pire, il peut s'agir d'un simple conditionnement pour essayer d'obtenir des meilleurs résultats à la prochaine étude internationale. C'est malheureusement ce qui a tendance à se passer, dans quelques pays, avec les résultats de PISA.

Bien entendu, si l'on considère que PISA produit une « mesure » valide de la formation mathématique dans son ensemble, ce que PISA ne fait pas et ne prétend pas faire, alors on peut trouver souhaitable que tout l'effort de formation des enseignants porte sur leur adaptation au cadre de référence de PISA et aux questions de PISA, pour que les élèves soient bien préparés aux futures épreuves (PISA 2006).

Mais chaque fois que l'on préserve un bon équilibre entre la réflexion curriculaire (que faut-il enseigner, pourquoi ? comment ?) et l'utilisation des observations faites dans le cadre des études évaluatives, des actions de formation des enseignants peuvent être organisées avec profit. Cela est d'autant plus vrai que les enseignants sont associés à la réflexion et ne reçoivent pas des instructions préparées dans un cénacle d'initiés.

Dans de nombreux pays, suite aux études internationales, selon les besoins détectés, des formations d'enseignants ont lieu sur la recherche de problème, sur la mathématisation des situations issues de la vie « réelle », sur les statistiques et la pensée probabiliste, sur la géométrie,...

5.1.3. Retombées sur les pratiques pédagogiques

Les retombées sur les pratiques sont plus difficiles à mettre en évidence.

Les études internationales véhiculent toutefois l'idée d'un élève actif, acteur de ses apprentissages et d'un enseignant qui n'est pas seulement un distributeur de connaissances, mais qui est surtout un organisateur d'apprentissages. Cette conception semble faire son chemin dans tous les pays et les études dont nous parlons ne sont pas étrangères à cette évolution.

5.2. En France

Comparé à ce qui se passe dans d'autres pays, peu d'efforts de communication et assez peu de travail d'approfondissement est fait, en France, autour des études internationales. La plupart des travaux sont de nature sociologique et se focalisent sur les résultats et sur les relations entre les variables de contexte (inputs) et les résultats (outputs) ; cela dans la tradition du courant de recherche sur l'efficacité des écoles (orientation *school effectiveness*). Aujourd'hui, les recherches s'orientent plutôt sur l'étude des processus en vue de leur amélioration (orientation *school improvement*) (cf. N. Bottani et Vrignaud, 2005). Ces recherches, qui prennent davantage en compte les pratiques d'enseignement, ne devraient plus faire l'impasse sur la question de la validité épistémologique, ni de la validité didactique, des questions utilisées pour les évaluations.

L'école française de recherche en didactique des mathématiques est particulièrement bien armée pour entreprendre ce genre d'études. Les travaux qui pourraient être menés dans ce cadre seront utiles dans le cas français, mais ils seront aussi de nature à renforcer l'estime que notre culture scolaire rencontre dans le monde.

Malgré cette apparente indifférence le système a cherché à s'adapter :

- Publication d'un décret définissant un socle commun de connaissances et de compétences précisant ce que tous les jeunes doivent maîtriser à la fin de leur scolarité obligatoire. Le décret se réfère explicitement à PISA.
- Modification plus ou moins profonde des questions d'examen (Brevet des collèges et baccalauréat). En particulier les questions du Brevet cherchent à être moins formelles et davantage ancrées sur des situations de la "vie réelle". Toutefois, l'essentiel est oublié, à savoir la dévolution au candidat des démarches à mettre en œuvre. On continue à baliser l'activité par un questionnement du type I. a), I, b) ... en n'oubliant pas de préciser : en

utilisant tel théorème, en appliquant telle procédure,... ce qui s'inscrit en totale contradiction avec les conceptions développées et évaluées par PISA.

- Révision des programmes du collège pour y inclure une dose d'aléatoire.
- Révision des programmes du primaire, pour les élèves de 6 à 11 ans, en se référant aux mauvais résultats enregistrés par PISA entre 15 et 16 ans et sur la base, à mon avis (et pas seulement !), d'une totale incompréhension des objectifs privilégiés par PISA.

6. Impact possible de ces études sur les pays francophones

On observe dans la francophonie des proximités dans les conceptions et les curriculums, mais aussi de grandes différences en ce qui concerne l'enseignement des mathématiques. Il s'en suit une grande variabilité de l'impact que les études internationales ont, ou peuvent avoir, dans ces pays.

Nous avons vu l'impact important que finalement l'étude PISA a en France (socle commun de connaissance et de compétence, modification des programmes d'étude,...). Des effets de même type peuvent être observés en Belgique et en Suisse.

Le cas du Québec mériterait une étude particulière (voir sec. 5 du texte sur la table ronde en session 2 de ce projet spécial). Cette province obtient d'assez bons résultats dans les études internationales et l'enseignement des mathématiques semble y être un genre de compromis de fait entre celui que l'on rencontre habituellement dans les pays de langue anglaise et celui observé dans d'autres pays francophones.

Bien que le français soit l'une des langues officielles de l'OCDE, l'influence de la langue française, comme celle des cultures francophones, dans les études internationales, est très faible. Les recherches en didactique, qui pourraient avoir une influence dans l'élaboration des cadres de référence comme dans l'interprétation des résultats des études, sont méconnues. Cela sans doute dû à l'absence de francophones représentant de ce champ de recherche dans les instances tant techniques que décisionnelles qui encadrent les études. La primauté est donnée aux instances gouvernementales qui n'ont que faire des recherches et aux psychométriciens qui, bien que maîtrisant des techniques intéressantes et utiles, manifestent une certaine naïveté devant la complexité des domaines qu'il s'agit d'évaluer.

La tentation française, de se désintéresser de ces études a fait long feu. En France comme ailleurs dans la francophonie, à juste titre, ou de façon abusive, dans les pays qui y participent comme dans ceux qui n'y participent pas encore, les études internationales sont devenus des déterminants importants des politiques éducatives.

Pour le moment, dans nos pays, l'impact se fait surtout par le biais d'injonctions technocratiques qui donnent peu la parole à la recherche et aux acteurs des systèmes. Nous avons déjà dit l'intérêt que nous portons à ces études et aux enseignements que l'on peut en tirer, mais pour que les effets positifs l'emportent sur les effets négatifs prévisibles, il conviendrait que la communauté mathématique s'investisse davantage à tous les niveaux de leur élaboration et de leur traitement.

Un peu partout dans les pays francophones, l'enseignement des mathématiques est assez formel et prend mal en compte les élèves qui éprouvent des difficultés devant le formalisme et l'abstraction. Cet enseignement réussit plus ou moins bien avec une partie des élèves, mais ne

parvient pas à entraîner les autres. On dénonce alors un manque de motivation puis un manque de bases chez les apprenants et l'on observe, en fin de compte, un rejet de la discipline chaque fois que c'est possible.

Le premier mérite des études internationales est de ramener l'attention sur les mathématiques pour tous, sur le nécessaire accès de tous à un ensemble de compétences où les mathématiques jouent un rôle important. À nous d'agir pour que ce repositionnement de l'enseignement des mathématiques se fasse sur des objectifs à la fois socialement pertinents et épistémologiquement valides.

7. Conclusion

Qu'elles soient nationales ou internationales, les évaluations ne sont pas sans défaut et nous en avons souligné quelques-uns. Toutefois, à condition de ne pas les prendre comme des mesures intrinsèques de la qualité des enseignements, de la qualité des programmes, ou de la qualité du système lui-même, à condition donc de considérer leurs résultats comme des indicateurs qu'il est possible et souhaitable de questionner, ces études présentent une utilité indéniable.

Les évaluations ne sont pas des mesures au sens de la physique et il faudrait en finir avec la mystification qui conduit à publier des scores qui n'ont souvent que peu de signification. Toute évaluation ne produit qu'une image partielle et imparfaite de la réalité dont elle est censée rendre compte. De ce fait, toute évaluation doit rester ouverte à l'analyse et aux interprétations plurielles, pour, ensuite ouvrir la voie à l'action. Cela est bien sûr valable pour les décideurs, mais ça l'est tout autant pour les divers acteurs du système éducatif : professeurs, parents, élèves, associations, ainsi que les autres acteurs du champ social. Évidemment, tout le monde n'a pas la même compétence pour traiter de la question, et les compétences des uns et des autres doivent être respectées, mais tout le monde, est concerné par l'enseignement et tout le monde doit avoir accès à l'intégralité de l'information.

La politique du domaine réservé qui est souvent la règle en matière d'évaluation du système, en France comme dans la plupart des autres pays, correspond, dans le meilleur des cas, à un manque de confiance dans l'intelligence et la bonne foi des différents acteurs. L'inquiétude correspondante n'est pas, il est vrai, dénuée de tout fondement, mais ce n'est qu'en libérant l'information et en favorisant recherches et débats, que ces évaluations pourront contribuer à l'amélioration de l'enseignement des mathématiques et, ce qui va de pair, à une meilleure professionnalisation des enseignants de cette discipline.

Au niveau des études internationales, cette politique d'ouverture est d'ailleurs celle qui est suivie officiellement par l'OCDE comme par l'IEA (TIMSS) mais les relais nécessaires au niveau national ont parfois tendance à se faire discrets.

D'une façon générale, les différentes évaluations sont complémentaires : évaluations nationales officielles ou non, évaluations internationales PISA et TIMSS, etc. Des analyses prenant en compte ces diverses évaluations devraient être conduites par des équipes où les institutions devraient avoir leur place, sans pour autant avoir le monopole de l'interprétation. Des recherches indépendantes devraient être conduites sur les études déjà faites. Des enquêtes complémentaires devraient être suscitées et menées au niveau national et au niveau européen.

Les études internationales sont une bonne occasion de s'interroger sur la solidité et sur la qualité de nos systèmes. Elles fournissent des indicateurs objectifs indépendants dont il serait bien dommage de se priver. Elles sont aussi une occasion d'échanges et de partage à travers pays et cultures qui ne peut qu'enrichir notre propre vision de l'enseignement des mathématiques comme, d'une façon plus générale, de la culture de notre temps.

Adresses et contacts

Note officielle française : <http://www.educ-eval.education.fr/pisa2003.htm>

Cadre de référence et rapports internationaux : <http://www.pisa.oecd.org/>

Sur le site de l'APMEP, article, diaporamas et présentation des questions libérées avec des résultats (en français et en anglais) http://www.apmep.asso.fr/rubrique.php3?id_rubrique=114

Site internet personnel : <http://web.me.com/antoinebodin/pro/>

Références

Bodin, A.. (2005) : *Ce qui est vraiment évalué par PISA en mathématiques. Ce qui ne l'est pas. Un point de vue français*. Communication faite à la conférences Franco Finlandaise sur PISA. www.apmep.asso.fr/IMG/pdf/Com_PISA_FF_Francais-2.pdf

Bottani, N. et Vrignaud, P. (2005) : *La France et les évaluations internationales*. Haut Conseil de l'Évaluation de l'École.

SEANCE 2 : TABLE RONDE

L'objet de cette table ronde était de présenter pour certains pays francophones comment sont analysés, présentés et médiatisés les résultats des évaluations internationales et de préciser leur impact politique et curriculaire. La séance a été conçue en deux parties. Dans la première partie, Charles Nassouri, Nadia Mawfik, Hikma Smida et Lucie DeBlois ont présenté leurs expériences, mettant en évidence des différences fondamentales quant aux influences des évaluations internationales sur les décisions politiques, les curricula, les manuels et les pratiques enseignantes et la recherche sur l'enseignement de mathématiques. Dans la deuxième partie de la table ronde, consacrée à la discussion, un certain nombre de questions de fond ont été posées et discutées. Dans ce qui suit, nous présentons d'abord une vision globale et synthétique de la séance, puis les contributions de chacun des intervenants.

Impact politique et curriculaire des évaluations internationales dans les pays francophones.

Hikma Smida, Université El Manar, Tunis (coordinatrice)

Lucie DeBlois, Université Laval, Québec

Nadia Mawfik, Ecole Normale Supérieure, Rabat

Charles Nassouri, Inspection de Mathématiques, Ouagadougou

Introduction

L'objet de cette table ronde était de décrire, pour certains pays francophones, les instances et les structures chargées de l'organisation et de la gestion des évaluations internationales et de préciser leur impact politique et curriculaire.

La réflexion a été déclinée en un certain nombre de questions :

- Comment s'effectue la diffusion des résultats ? Comment sont-ils traités dans les médias ?
- Les résultats des évaluations internationales sont-ils source de débats au niveau national ? Si oui, de quelle nature sont ces débats et qui impliquent-ils ?
- Les résultats des évaluations internationales entraînent-ils des réactions politiques et si oui de quelle nature ? Des décisions politiques récentes concernant l'enseignement des mathématiques ont-elles été au moins en partie motivées par ces évaluations ?
- Dans les réformes récentes du curriculum ou celles en cours concernant les élèves ou la formation des enseignants, peut-on percevoir une influence plus ou moins directe des évaluations internationales, et si oui, laquelle.

La séance a été conçue en deux parties : une première partie consacrée à la présentations des expériences de quatre pays : Burkina Faso, Maroc et Tunisie et Canada, la deuxième partie à une discussion sur la mission et les objectifs de l'enseignement des mathématiques définis par les évaluations internationales et l'impact des résultats de ces évaluations sur la communauté des mathématiciens, praticiens et chercheurs.

1. Une présentation globale de la table ronde

1.1. Les communications

Les communications ont concerné quatre pays : le Burkina Faso, le Canada, le Maroc et la Tunisie et ont été présentées respectivement par Charles Nassouri, Lucie DeBlois, Nadia Mawfik et Hikma Smida.

Le Burkina Faso ne participe pas aux évaluations TIMSS et PISA. Toutefois, il apparaît que l'expérience du Burkina dans les évaluations africaines (MLA) et les olympiades africaines a permis à certains inspecteurs d'exploiter le matériel de ces évaluations pour amener les enseignants à réfléchir sur leurs pratiques et à travailler avec leurs pairs. La dynamique impulsée par ces réflexions vise à asseoir un programme de formation des enseignants, à une échelle nationale.

Les interventions ont mis en évidence des différences fondamentales entre l'expérience canadienne et celles du Maroc et de la Tunisie.

Au Canada, les résultats des évaluations sont transmis par le Conseil des Ministres du Canada, le Ministère des Ressources Humaines et Développement des compétences et Statistiques Canada sous forme de conférence de presse et sont ensuite relayés par les médias, les syndicats d'enseignants et les associations. De plus, des études comparatives inter-provinces et inter-pays sont menées. Sur le plan de la recherche, plusieurs projets permettant d'analyser les données primaires ou secondaires, sont menés par des universitaires provenant d'une variété d'universités canadiennes et financés par les ministères.

Au Maroc, les résultats sont médiatisés par la presse, la plupart du temps avec un regard critique sur la politique éducative du pays, les curricula et l'environnement socioculturel. Le ministère de l'éducation marocain a exploité, quant à lui, les résultats pour réformer l'enseignement des mathématiques et mettre en place un modèle standardisé d'évaluation dans certaines disciplines, sur la base d'un référentiel de compétences. En outre, il n'existe quasiment pas d'impact des évaluations internationales sur les enseignants qui leur reprochent d'être établies sur des curricula hypothétiques et d'être sans rapport avec les évaluations certificatives. Il semble, de même, que la communauté des inspecteurs reste peu impliquée dans l'exploitation des résultats.

En Tunisie, les résultats sont rarement relayés par la presse. Le ministère de l'éducation exploite les indicateurs de base et ceux contextuels pour ses stratégies éducatives, et les performances des élèves tunisiens pour négocier les changements dans les curricula et les manuels, ainsi que pour définir le référentiel des programmes de formation des enseignants. Malgré l'absence de débats nationaux à propos des évaluations, il a été possible de constater certaines actions individuelles de la part des inspecteurs de mathématiques en vue d'informer les enseignants, de les sensibiliser aux choix et approches adoptées par les évaluations. De plus, certains inspecteurs organisent des séances de formation des enseignants, dans lesquelles sont discutés les contenus des évaluations internationales, les erreurs des élèves les plus fréquentes et les difficultés que ces élèves rencontrent. Ces séances de formation visent, selon les inspecteurs, à permettre aux enseignants de réfléchir sur leurs pratiques en termes

d'enseignement-apprentissage et d'évaluation et mettent en évidence une grande résistance de la part des enseignants à adhérer à ces évaluations.

En conclusion, comme le prouvent les différentes présentations de la table ronde, l'impact des évaluations internationales sur l'enseignement des mathématiques diffère suivant les pays. Au Canada, il semble que ces manifestations, devenues une partie intégrante de la politique d'enseignement des mathématiques, impulsent une dynamique de changement et de recherche dans les différentes provinces. En Tunisie et au Maroc, malgré une « longue » tradition de participation, il semble que ces évaluations concernent plus les ministères d'éducation et leurs structures que la communauté des praticiens et des chercheurs. Toutefois, les actions menées par certains inspecteurs du Burkina Faso et de la Tunisie pour impulser des actions de formation des enseignants sur la base des données fournies par les évaluations internationales ou les olympiades africaines, suggèrent de réelles opportunités d'une exploitation efficace des données de ces évaluations à des fins d'amélioration de l'enseignement apprentissage des mathématiques et des pratiques enseignantes.

1.2. Discussion

Au cours de la discussion, les avis exprimés par les participants ont été mitigés et parfois opposés.

Certains participants ont apporté un regard critique sur les objectifs, l'organisation et la méthodologie adoptés par les évaluations internationales telles que TIMSS et PISA, certains autres ont fait part de leur doute quant à la contribution de ces évaluations à l'amélioration de l'enseignement des mathématiques, et enfin d'autres participants ont estimé que les évaluations pourraient permettre aux pays participants d'engager une réflexion sur leur rendement, en termes de qualité et d'équité. En résumé, la discussion a suscité, chez participants, un certain nombre de questions de fond que l'on peut articuler autour de trois axes :

1. Les évaluations internationales aident-elles réellement les pays à expliquer et évaluer leur rendement en termes de qualité ou d'équité ou bien, au contraire, contribuent-elles à propager une certaine idéologie ?
2. Les évaluations internationales peuvent-elles contribuer à l'amélioration de l'enseignement des mathématiques ? Constituer un catalyseur pour amener les enseignants de mathématiques à adapter leurs pratiques à un contexte où l'enseignement des mathématiques est interrogé sur sa contribution à aider les élèves à affronter leur vie de citoyen ?
3. Dans quelle mesure la recherche en didactique peut-elle aider à mieux saisir les enjeux de ces évaluations et à rendre leurs résultats plus objectifs et plus explicites ?

Après cette présentation globale, nous reproduisons ci-après les communications présentées par les intervenants.

2. Evaluations internationales : cas de la Tunisie

2.1. Organisation

La Tunisie participe à l'évaluation PISA depuis 2003 et à l'évaluation TIMSS depuis 1999 pour le grade 8 et 2003 pour le grade 4. L'organisation des évaluations TIMSS et PISA est à la charge du département d'évaluation, au sein du centre national de l'innovation pédagogique et de la recherche en éducation (CNIPRE). Le département d'évaluation gère toutes les procédures de passation. Pour ce faire, le département a formé des équipes constituées d'enseignants, de pédagogues et d'inspecteurs qui ont la charge de procéder à l'échantillonnage, la traduction¹⁵, la pré-passation, le codage et la correction. Avant chaque évaluation, des campagnes de sensibilisation et d'information sont menées auprès des directeurs régionaux de l'enseignement, des chefs d'établissement et des enseignants dont les classes ont été retenues pour la passation. Au cours de ces campagnes, des membres du département d'évaluation expliquent les procédures de passation, exposent les consignes et précisent les enjeux de ces évaluations. Les chefs d'établissements sont chargés d'informer et de sensibiliser les élèves retenus par l'échantillon national. Signalons toutefois qu'aucune action n'est entreprise pour sensibiliser les parents des élèves concernés par l'évaluation. L'analyse des résultats est menée par la commission tunisienne pour l'enseignement des mathématiques. Cette analyse se déroule en deux temps et donne lieu à deux rapports :

1. Un rapport national élaboré sur la base des données fournies par TIMSS ou PISA, qui
 - rend compte de la répartition des performances des élèves tunisiens en regard des différents niveaux de performance établis par les évaluations et en regard des moyennes des performances internationales, ainsi que des disparités dans les performances tunisiennes en ce qui concerne le genre et le milieu socio-économique des élèves ;
 - explique les performances des élèves tunisiens, en regard des différences entre les programmes tunisiens et les items de TIMSS ou PISA, des réponses aux questionnaires soumis aux élèves, aux enseignants et aux chefs d'établissement ;
 - propose aux décideurs des recommandations susceptibles d'améliorer les performances.
2. Un rapport interne, élaboré sur la base de l'analyse des items proposés et des réponses des élèves tunisiens. Ce rapport identifie certaines des erreurs et des difficultés rencontrées par les élèves tunisiens, pointe les dysfonctionnements au niveau du curriculum tunisien et des pratiques des enseignants et enfin propose certaines actions à entreprendre.

Le rapport interne est communiqué au ministère qui le transmet aux inspecteurs de mathématiques et à la direction générale des programmes.

2.2. Impact des évaluations internationales

Tous les indicateurs de base et ceux contextuels sont utilisés par les décideurs lors de la mise en place des stratégies éducatives. En général, les résultats relatifs aux performances des élèves et aux pratiques des enseignants sont convoqués par les décideurs lors des négociations pour les changements de programmes, de manuels ou de la mise en place de programmes de

¹⁵ Les épreuves se déroulent en arabe.

formation pour les enseignants. Toutefois, force est de constater l'absence de débat national et une réticence à aborder publiquement les problèmes avec les parties concernées (enseignants, parents, associations, syndicats).

La réforme des programmes de mathématiques tunisiens de 2002 a intégré certains contenus et compétences requis par TIMSS. A titre d'exemples, il a été décidé d'introduire l'enseignement des statistiques dans le primaire et le collège, d'initier les élèves du primaire et du collège à certains phénomènes aléatoires et à un vocabulaire probabiliste, d'inclure des éléments de calcul approché et d'estimation dans la partie numérique, de recommander le recours à divers registres (graphique¹⁶, symbolique, etc.), de développer des compétences de résolution de problèmes et de modélisation dans tous les champs des mathématiques enseignés dans ces cursus.

Une des difficultés majeures rencontrées par les élèves tunisiens concerne le degré de familiarité avec les situations proposées par les évaluations internationales. Ce constat a été pris en compte par les concepteurs de manuels qui ont, tant que faire se peut, tenté de diversifier les familles de problèmes et situations proposés dans les manuels. On a pu en effet constater dans les nouveaux manuels une contextualisation plus importante des problèmes de mathématiques, dans un champ assez riche de situations liées à l'environnement de l'élève (lien avec les autres disciplines, avec la vie quotidienne, avec l'environnement international).

Au niveau des enseignants, seules des actions individuelles menées par certains inspecteurs ont pu être constatées. Après les publications des résultats de TIMSS ou PISA, certains inspecteurs présentent et discutent les performances des élèves tunisiens avec les enseignants dont ils ont la charge. Ils organisent des séances de formation dans lesquelles ils étudient certains items libérés, travaillent avec les enseignants à la production de situations d'apprentissage dans des contextes adoptés par TIMSS et PISA, ainsi qu'à la conception d'évaluations en adéquation avec les objectifs de TIMSS ou PISA. La majorité des inspecteurs impliqués signalent une forte résistance de la part des enseignants à changer leurs pratiques. Selon les inspecteurs, la raison principale de cette résistance au changement est due essentiellement à une divergence de point de vue avec les choix adoptés par les évaluations internationales quant à ce qui doit être enseigné et évalué en mathématiques.

Au plan de la recherche en didactique des mathématiques et de la recherche action, nous n'avons détecté aucun projet provenant des chercheurs en didactique ou des inspecteurs. Toutefois, signalons qu'une recherche sur les pratiques des enseignants à été menée par la commission tunisienne pour l'enseignement des mathématiques, dans le cadre d'un projet mené par TIMSS avec les pays arabes participants.

3. Evaluations internationales : cas du Maroc

3.1. Organisation

Au Maroc, c'est la Direction de l'Evaluation du Système Educatif (DESE) en collaboration avec des instances internationales (IEA, Collège de Boston, UNESCO, UNICEF, etc.) qui organise ces évaluations internationales avec pour objectifs essentiels de :

¹⁶ A ces niveaux de cursus, le registre graphique était réservé uniquement à la géométrie.

- doter le système éducatif d'une base de données rigoureusement et scientifiquement établie ;
- instaurer une culture et une pratique de l'évaluation qui constituent les éléments essentiels sur lesquels se base l'édification d'un système éducatif à la fois performant et évolutif ;
- faire le diagnostic de la qualité des acquis scolaires des élèves dans les matières visées ;
- développer l'expertise nationale en matière d'évaluation ;
- préparer la prise de décision, en mettant à la disposition des responsables et décideurs des résultats fiables relativement au niveau réel des acquis scolaires des élèves considérés.

Les tests de connaissances sont élaborés sur la base d'un socle de compétences défini d'un commun accord par les experts des différents pays de la région. Il s'agit en particulier du test de lecture/écriture, du test de calcul et du test des compétences de la vie courante. En mathématiques, le test de calcul s'articule autour des domaines suivants : arithmétique, mesures, géométrie et résolution de problèmes.

Avant d'être mis sous leur forme définitive, les tests sont expérimentés auprès d'un échantillon d'élèves non retenus pour l'étude proprement dite. Les résultats de cette expérimentation servent à mettre ces instruments sous leur forme définitive. Le choix des établissements est effectué selon des procédures d'échantillonnage aléatoire. Pour chaque établissement choisi, une seule classe de 25 élèves est retenue. Pour le choix des régions et des délégations, les critères de choix sont les performances en matière des services éducatifs offerts, le taux d'urbanisation et la diversité géographique. Des sensibilisations des responsables des délégations provinciales sont réalisées par les cadres de la direction. Ces sensibilisations visent à faire connaître aux responsables régionaux les objectifs de l'étude, l'importance que le Ministère accorde à la dite étude, et ce que la direction attend d'eux pour que l'enquête se réalise dans de bonnes conditions. Par ailleurs, La direction de l'Évaluation du Système Éducatif organise des ateliers régionaux pour présenter l'étude aux différents intervenants, les démarches de passation et les différentes consignes à respecter lors de la passation en vue de garantir un certain niveau de standardisation des conditions de passation. Les dits ateliers sont encadrés par les cadres de la direction (DESE) et par des formateurs expérimentés. La saisie et l'analyse des données sont réalisées à l'aide d'un logiciel d'analyse S.P.S.S, par une équipe technique formée à cet effet au sein de la Direction (DESE).

3.2. Impacts politique, curriculaire et sur les enseignants.

La réforme de l'enseignement lancée au Maroc dès 2000, semble avoir été motivée, au moins en partie, par les résultats des évaluations TIMSS 1999-R. En fait, les résultats n'ont pas été divulgués mais une rencontre nationale a été organisée à Ifran en Août 1999, au cours de laquelle il a été décidé de lancer une réforme de grande envergure de l'enseignement et l'élaboration de la charte nationale d'éducation et de formation. De même, la décennie 2000-2009 a été déclarée décennie de l'éducation et de la formation.

Publiés dans la presse nationale, les résultats de TIMSS 2003, ont montré, une fois encore que les élèves marocains étaient nettement en dessous de la moyenne internationale. Les responsables relativisent cependant ces résultats et avancent des arguments tels que : la réforme est en cours, il faut attendre qu'elle soit achevée pour juger. Néanmoins, un conseil national de l'éducation CNS et une instance nationale d'évaluation du système d'éducation et de formation (l'INSEF) a été créée en 2006, avec pour objectifs de contribuer à la diffusion d'une culture d'évaluation dans tout le système éducatif, de procéder à des évaluations globales, sectorielles ou thématiques du système d'éducation et de formation, et d'apprécier les performances pédagogiques et financières du système éducatif, en regard des objectifs qui lui sont assignés et des normes internationales reconnues en la matière.

Les résultats de TIMSS 2007, publiés par la presse nationale, sont venus confirmer que les élèves marocains sont parmi les moins performants au monde, en ce qui a trait au rendement général en Sciences et en Mathématiques. Ils ont même enregistré un résultat plus faible que celui de 2003 dans la même étude. Avec un score de 341 points en Mathématiques (6 points de moins qu'en 2003), le Maroc reste loin au dessous de la moyenne internationale, pourtant les moyens financiers mis en œuvre sont énormes. Pour relativiser ce constat, les responsables avancent l'argument que les évaluations internationales sont basées sur des curricula hypothétiques et invoquent que les différences culturelles peuvent biaiser les résultats. Ils appellent à organiser des évaluations basées sur les curricula réels officiels au Maroc.

Comme conséquence directe ou indirecte, il y a eu le lancement du programme national d'évaluation des acquis (PNEA), programme piloté par l'INESEF auprès du conseil supérieur de l'enseignement (CSE) en collaboration avec le CNEE (centre national de l'évaluation et des examens). Ce programme vise à instaurer un référentiel national d'évaluation des connaissances et des compétences de base des apprenants marocains, assorti d'indicateurs clairs et pertinents. Son objectif est de contribuer à la prise de décision en matière de politiques éducatives relatives aux programmes, aux curricula et aux apprentissages.

L'évaluation standardisée des acquis a été lancée en juin 2008, sur les apprentissages de l'année scolaire 2007-2008. Les disciplines considérées sont l'arabe, le français, les mathématiques et les sciences (y compris la physique-chimie au collège). Le référentiel de chaque matière concernée par l'étude est basé sur les curricula officiels, les notes ministérielles, les manuels scolaires ainsi que les orientations pédagogiques. L'évaluation a concerné les élèves de la quatrième et de la sixième année du primaire, ainsi que la deuxième et la troisième année du secondaire collégial. Les résultats de l'enquête ont révélé un niveau des élèves «très moyen à faible» en mathématiques. (*L'Economiste du Maroc* 11/03/2009). Par ailleurs, les élèves du primaire semblent mieux faire que ceux du collège. La moyenne des objectifs réalisés par les élèves du primaire dépasse celle du secondaire collégial. Entre la sixième du primaire et la troisième année du secondaire collégial, on a pu constater jusqu'à 16 points d'écart. L'addition, la soustraction, la multiplication et la division sont maîtrisées par le tiers seulement des élèves de la quatrième année du primaire. La publication des résultats de cette enquête a suscité un débat qui a été relayé par les médias nationaux et les associations et certains dysfonctionnements sont pointés du doigt dont nous pouvons citer:

la démarche parcellaire et disjonctive de la réforme des curricula, **la faible harmonie** entre la philosophie des nouveaux programmes et celle d'autres composantes de la pédagogie

scolaire (manuel scolaire, examens, rythmes scolaires, pratiques d'enseignement et d'apprentissage), **la prédominance** dans les contenus des savoirs livresques au détriment des savoirs pratiques, **la faible adéquation** avec l'environnement économique et socioculturel des apprenants, **la modeste implication** des partenaires socio-économiques dans la conception des programmes, **une opérationnalisation** fragile de l'approche par compétences, en raison de la limite des programmes de formation continue des enseignants, **une faible préparation** des acteurs à la nouvelle réforme pédagogique et une absence d'un système responsabilisant et de démarches d'évaluation et de monitoring, **l'absence d'un** processus d'évaluation régulier des programmes et des curricula permettant d'identifier leurs faiblesses et leurs atouts, **le manque** de mécanismes d'incitation et de responsabilité.

Le ministre de l'enseignement a annoncé dernièrement le lancement d'un programme d'urgence de l'éducation nationale pour la période 2009-2012, avec pour objectif d'accélérer le rythme de la réforme. Dire que les résultats des évaluations a influencé directement cette décision est peut être simpliste mais nous pensons que cela y a probablement contribué.

En ce qui concerne le curriculum, on ne peut pas dire qu'il y a « vraiment » un impact visible. Les changements dans les programmes décidés depuis le lancement de la réforme de l'enseignement en 2000 viennent de se généraliser à tous les niveaux de l'enseignement primaire et secondaire et la réforme de l'université n'est pas encore achevée.

Parmi les enseignants marocains, beaucoup relativisent les résultats. Pour ces enseignants, les évaluations sont basées sur des curricula hypothétiques et sont inadéquates avec les évaluations certificatives. Ils incriminent la politique de l'enseignement fondamental obligatoire (pas ou peu de redoublement), ou encore la politique de scolarisation massive et accélérée (manque de moyens, classes surchargées). Selon eux, les décideurs ont misé sur le quantitatif au détriment du qualitatif.

4. Evaluations internationales : cas du Burkina Faso

4.1. Evaluations

D'une manière générale, on rencontre peu d'évaluations internationales dans les pays d'Afrique francophones au sud du Sahara. Pour prendre le cas du Burkina Faso, il y a eu en 2003 une évaluation intitulée *Motoring Learning Accesement II (MLA II)*. Cette évaluation, qui a concerné les sciences et les mathématiques, s'adressait aux élèves de la classe de 5^{ème} et elle concernait un certain nombre de pays dont le Burkina Faso. Elle n'a pratiquement pas eu d'impact sur les politiques curriculaires puisque ses résultats n'ont été évoqués à aucun moment pour relire ou revoir les programmes. Elle est intervenue comme une commande adressée aux autorités ou à un groupe de consultants qui en a remis les résultats aux commanditaires, sans plus.

HPM (Harmonisation des Programmes de Mathématiques) est un projet qui a existé de 1983 à 2003. Son objectif était d'harmoniser les programmes de mathématiques de tous les pays francophones d'Afrique et de l'Océan indien. Dans le cadre de cette grande ambition, il a été envisagé une évaluation au niveau des pays concernés mais cela n'a jamais vu le jour.

Il y a eu cependant au Burkina Faso des évaluations nationales, à partir de 1994. Ces évaluations nationales ont porté sur les acquis des élèves. Elles sont axées sur les contenus. Elles ont concerné les classes de la 6^{ième} à la troisième. On peut parler de l'influence de ces évaluations sur les curricula dans la mesure où elles ont eu pour effet un certain nombre de réajustements au niveau des programmes en cours. En effet :

- au niveau de la géométrie de l'espace, elles ont entraîné une série de formations en direction des enseignants ;
- la question de l'initiation à la modélisation (mise en équation) a été une préoccupation importante des encadreurs pédagogiques, qui ont organisé localement des formations en direction des enseignants ;
- dans le programme lui-même, il a été nécessaire de revoir certaines notions, qui ont été remise en place compte tenu de leur importance dans la résolution de certains problèmes.

Cependant, d'autres problèmes, liés à notre contexte, sont venus freiner cette dynamique.

4.2. Les Olympiades panafricaines de mathématiques

D'autres influences peuvent être notées à travers la participation régulière du Burkina Faso aux olympiades panafricaines de mathématiques (OPAM). En effet, bien qu'on ne puisse parler des olympiades au même titre que des évaluations internationales dont il est question dans ce projet, je vais quand même les aborder dans la mesure où elles sont susceptibles d'avoir une influence sur l'éducation mathématique des élèves.

Les objectifs des OPAM sont:

- de créer une saine et fructueuse émulation au sein de la jeunesse ;
- de détecter de jeunes talents en mathématiques ;
- de tisser des relations d'amitié et de travail entre mathématiciens africains ;
- d'échanger des informations sur les programmes scolaires et sur les problèmes liés à leur mise en œuvre.

Leur impact sur les programmes scolaire est quasiment nul, parce qu'en général un pays ne modifie pas son programme scolaire pour se conformer à celui d'une compétition internationale. Tout au plus, peut-il y découvrir l'importance de telle ou telle notion ou thème qui s'avère très important pour l'éducation mathématique des élèves et l'incorporer dans son programme. Cependant les olympiades peuvent avoir un impact sur les pratiques des enseignants.

En effet, les énoncés de type olympiades présentent un certain nombre de différences avec les énoncés de type traditionnel:

- ils sont souvent très brefs, posés sous forme de problème ouverts : la réponse n'est pas fournie dans la question ;
- ils donnent envie d'être cherchés car ils sont souvent surprenants, se présentant comme un défi à relever ;
- ils permettent de tâtonner, expérimenter, conjecturer. L'induction y est souvent un recours productif. Ils autorisent la mise en œuvre d'une méthode générale de recherche

(examen d'un cas particulier, examen d'une version simplifiée du problème, analyse et synthèse ...)

- la nature de la réponse peut être nouvelle ;
- ils peuvent ne nécessiter aucune connaissance mathématique particulière.

La pratique de la résolution de problème de type olympiade nécessite ou développe un certain nombre de qualités que nous voudrions voir chez nos élèves : le courage, la persévérance, l'esprit d'initiative, l'endurance, la confiance en soi, la patience et d'autres encore qu'il faut conjuguer en même temps. Au-delà des médailles qu'un élève ou un pays peut amasser, le développement de ces qualités devrait être l'objectif majeur.

Pour cela, il est important de ne pas s'en tenir à l'encadrement de quelques élèves, mais de faire en sorte que cela concerne le plus grand nombre d'élèves possible. C'est ce que nous essayons de faire au Burkina Faso. Ainsi, nous commençons (modulo les moyens disponibles) l'encadrement au niveau de la classe de seconde, faute de pouvoir le commencer plus tôt. C'est ce que font également un certains nombres de pays francophones d'Afrique tels que la Côte d'Ivoire qui a obtenu la médaille d'or à la dernière édition des OPAM en Afrique du Sud.

Cette pratique devrait amener beaucoup d'élèves à s'intéresser aux mathématiques comme le montre cette petite histoire : un élève, après une séance en classe sur un problème ouvert de type olympiade, a dit en aparté à son professeur : « maintenant je vois ce que c'est que les mathématiques ; ça permet de libérer son esprit créateur ». Une telle réaction, si elle se produisait au niveau d'un grand nombre d'élèves, vaudrait certainement aussi une médaille.

En résumé, le Burkina Faso ne participe ni aux évaluations TIMSS, ni à PISA. Toutefois, le pays a une certaine expérience dans des évaluations ou des compétitions africaines telles que le MLA et les olympiades africaines. Aucun impact de ces évaluations n'a été signalé au niveau curriculaire. Toutefois, la préparation des élèves burkinabés aux olympiades a permis d'impulser une dynamique auprès de certains enseignants de mathématiques. Avec l'aide de leurs inspecteurs, ces enseignants ont pu s'engager dans une réflexion sur leurs pratiques dans leurs classes, sur les choix de situations d'apprentissage pertinentes et sur des modes d'évaluation non traditionnels.

5. Evaluations internationales : cas du Canada

Dans le cadre de cette table-ronde, je me concentrerai sur les évaluations PISA 2000, 2003, 2006 que je connais mieux. Je présenterai d'abord la façon dont sont diffusés les résultats des élèves canadiens, je préciserai ensuite comment sont développées certaines recherches pour ensuite offrir deux exemples à partir de mes recherches. Je terminerai en discutant l'influence de ces évaluations sur les différentes provinces.

5.1. La Diffusion des résultats

Au plan de la diffusion, entre 22 000 à 28 000 élèves canadiens participent aux évaluations PISA. Il n'existe aucun ministère d'éducation canadien, mais plutôt un Conseil des Ministres de l'Éducation du Canada (CMEC), une structure qui favorise la mise en commun ou la mise en place de projets. C'est dire que sur le plan de l'éducation, le Canada est l'équivalent de

treize pays différents. Au moment de diffuser les résultats d'évaluations internationales, le CMEC, le Ministère des Ressources humaines et Développement des compétences au Canada de même que Statistique Canada organisent une conférence de presse retransmise simultanément sur le Web. Les médias reçoivent le rapport canadien à l'avance sous embargo www.pisa.gc.ca. Habituellement, les médias mettent l'accent sur les résultats comparatifs. Ainsi, les titres suivants apparaissent : « Les élèves canadiens se classent parmi les meilleurs en mathématiques, sciences, lecture et résolution de problèmes, selon les conclusions d'une importante étude de l'OCDE » (décembre 2004). En outre, Statistique Canada présente les résultats par province en les situant par rapport aux autres pays ayant participé à cette évaluation. Par exemple, les résultats des élèves de la province de Québec se situent entre ceux des élèves des Pays-Bas et du Liechtenstein alors que ceux de la province de l'Alberta se situent entre ceux de Hong Kong et de la Finlande.

En outre, une variété d'autres comparaisons est faite. Les provinces et les territoires sont situés les uns par rapport aux autres. Par exemple, pour Pisa 2003, l'écart entre la moyenne des élèves au Québec (537) et la moyenne des élèves en Ontario (530) est faible tout comme l'écart entre la moyenne des élèves de l'Île du Prince-Édouard (500) et du Nouveau-Brunswick (512). Toutefois, l'écart entre ces deux groupes de provinces est grand. Les résultats des élèves anglophones et francophones sont présentés séparément dans cinq provinces (Manitoba, Ontario, Québec, Nouveau-Brunswick et Nouvelle-Écosse), ce qui permet de reconnaître que le Québec est la seule province où les résultats des élèves francophones et anglophones sont comparables. Toutefois, une étude de l'Institut de la statistique du Québec révèle, en février 2010, que les francophones universitaires du Québec demeurent toujours moins scolarisés que les allophones et les anglophones de cette même province, contrairement aux francophones universitaires de l'Ontario qui eux sont plus scolarisés que les anglophones de cette province (D'amours, 2010). Ainsi, malgré des résultats qui ne permettent pas de distinguer les performances des francophones et des anglophones de 16 ans, il en est tout autrement dans la poursuite des études notamment universitaires.

Un phénomène particulier apparaît. Les provinces et territoires anglophones semblent davantage intéressés par les comparaisons avec les résultats des états des USA proches des provinces (Nouveau-Brunswick-Maine) et de certains pays du Commonwealth britannique tels que l'Angleterre, l'Australie, la Nouvelle Zélande, Hong-Kong. Le Québec (francophone) est davantage intéressé par les comparaisons avec la France, la Suisse ou la Belgique.

Enfin, chaque province et territoire étant entièrement responsable de son système d'éducation tel que prévu par la loi constitutionnelle canadienne, les différents Ministères de l'éducation publieront des articles qui sont colorés par les résultats obtenus. Par exemple, l'Ontario se classant relativement bien, présentera un article dont le titre est : « Les élèves de l'Ontario font partie des meilleurs au Canada ». (*Ministère de l'éducation de l'Ontario*, le 30 avril 2008). Le Nouveau-Brunswick ayant les moins bons résultats au Canada offrira plutôt ce titre : « Nouveau-Brunswick - "Les enfants au premier plan" pour de meilleurs résultats aux évaluations internationales » (*Ministère de l'éducation du Nouveau-Brunswick*, 6 décembre 2007). Les médias essaieront de mettre à l'avant-plan certaines caractéristiques qui

différencient les systèmes d'éducation (le financement, les réformes de l'éducation, l'impact des évaluations, les programmes d'étude, etc.).

Le site Internet des syndicats des enseignantes et des enseignants chercheront à conserver un regard plus critique en questionnant les modèles qui peuvent expliquer les résultats (L'école finlandaise : un modèle fondé sur l'égalité (Nouvelles CSQ, hiver 2009, p.21). Les journaux locaux publieront des articles plus louangeurs. Nous retrouverons des titres comme : Les Québécois, des «bollés». (Le soleil 5 décembre 2007). Ces résultats sont encourageants lorsqu'on sait qu'en 1950, les québécois présentaient un taux de décrochage de 66% au cours classique (Baby et DeBlois, 2005). Fortin (Cynthia St-Hilaire, La Voix de l'Est 2010) reconnaît même que le taux de décrochage est actuellement de 12%, lorsqu'on intègre les jeunes jusqu'à 29 ans, contrairement à 40% en 1966.

5.2. La recherche

Au plan de la recherche, le gouvernement du Canada s'impliquait jusqu'à tout récemment de façon plus importante (CESC-CRSH). Ainsi entre 2003 et 2006, 3,2 millions de dollars canadiens ont été consacrés à un programme appelé «Initiative de recherche en éducation CSCE-CRSH (IRECC)». Des universitaires provenant d'une variété d'universités canadiennes ont donc obtenu des subventions sur la base de la proposition de projets permettant d'analyser les données primaires ou secondaires¹⁷.

Un colloque annuel a permis à ces chercheurs et chercheuses de discuter de leurs plans de recherche ou encore de leurs résultats sur des thèmes comme «Les résultats de l'apprentissage et les transitions» (2003) ou encore «Mobilisation du savoir : de la recherche aux politiques et à la pratique» (2006). Ce programme est actuellement en évaluation sur les points suivants : Mobilisation des connaissances, Partenariats internationaux et nationaux, Formation d'étudiantes et d'étudiants, Recherche, analyse et rédaction.

À titre d'exemple, voici deux projets utilisant les bases de données Pisa 2003. Les objectifs du premier projet de recherche consistaient : 1) à décrire, pour l'ensemble des élèves canadiens

-
1. ¹⁷ Les déterminants de la réussite en écriture des élèves de 13 et 16 ans du Canada francophone, Mme Yamina Bouchamma (Université Laval)
 2. Examen des pratiques pédagogiques déterminantes sur le rendement des élèves ontariens et québécois aux évaluations internationales en lecture et en écriture, Mme Renée Forgette-Giroux (Université d'Ottawa)
 3. Développement du capital humain chez des élèves issus de milieux socio-économique défavorisés de 15 ans, Mme Lucie DeBlois, Marc-André Deniger, Xiao Zang, Richard Bertrand (Université Laval)
 4. Pratiques pédagogiques en classe et résultats d'apprentissage, Mme Kadriye Ercikan (Université de Colombie-Britannique)
 5. Rendement scolaire et accès à l'enseignement postsecondaire des jeunes immigrantes et immigrants et des jeunes réfugiés, Mme Lana Stermac (Université de Toronto)
 6. Les facettes multidimensionnelles de la TIC dans les communautés ethniques, Mme Nombuso Dlamini (Université de Windsor)
 7. Élèves risquant d'éprouver des difficultés scolaires : transition au premier cycle du secondaire, Mme Tanya Beran (Université de Calgary)
 8. Facteurs culturels, scolaires et individuels de succès en 1re année, M. Barry Schneider (Université d'Ottawa)
 9. Transitions au postsecondaire. Transition du niveau postsecondaire au marché du travail : rôle de l'éducation coopérative, Mme Maureen Drysdale (Université St. Jerome)
 10. Transitions vers l'université et choix d'un établissement, M. Martin Dooley (Université McMaster)

les forces et les stratégies utilisées pour répondre à certains des problèmes de mathématique, en particulier en ce qui concerne les problèmes ayant 2 codes de correction (Annexe 1), 2) effectuer des comparaisons en fonction des groupes linguistiques (Anglophones et Francophones) et en fonction des territoires (provinces), 3) effectuer des comparaisons selon le sexe. Nous avons pu reconnaître que les différences entre les filles et les garçons sont peu importantes. En ce qui concerne, les variétés de réponses des élèves, nous avons pu reconnaître qu'ils se servent rarement d'un langage relationnel pour justifier leurs explications écrites. Par exemple, pour la situation «Le cambrioleur» (Annexe 1), les élèves s'appuient sur l'apparence du graphique plutôt que sur les relations entre les variables. En outre, nous observons les francophones ont plus de difficultés à justifier par écrit (mathématiquement ou autrement) leurs réponses que les élèves anglophones. Enfin, les élèves ne semblent pas faire de retour métacognitif sur leur démarche (mécanismes de validation de leurs réponses). Par exemple, pour la situation-problème «Les pas» les élèves oublient de réaliser la conversion des unités de mesure.

Un deuxième projet a permis de formuler la question suivante : Les élèves francophones en milieu minoritaire réussissent-ils moins bien en numératie que les élèves anglophones ? La numératie est définie par PISA comme étant la capacité à résoudre une situation écrite de la vie courante en utilisant les mathématiques. À l'instar d'autres recherches, il existe une relation importante entre la littératie (mineure) et la numératie (majeure) (corrélation 0,90). Nous observons peu de différences entre les élèves francophones et anglophones au Québec (Rousseau et al, EMF 2006, thème 7). En comparant les élèves qui ont au haut niveau de compétence en littératie, il y a peu de différences entre les anglophones et les francophones en numératie. À niveau de littératie égal, les élèves francophones en milieu minoritaire présentent une plus grande compétence en numératie que les élèves anglophones (Nouvelle-Écosse, Nouveau-Brunswick, Ontario, Manitoba). Toutefois, en comparant les élèves qui ont au faible niveau de compétence en littératie, il y a une différence entre les résultats des élèves anglophones et les francophones en numératie.

5.3. Influence

Au plan de l'influence des évaluations internationales, elles semblent différentes selon les résultats obtenus par les élèves des différentes provinces et territoires du Canada. Par exemple, des demandes de financement supplémentaires sont formulées lorsque les évaluations sont faibles. Certaines décisions provinciales sont prises. C'est ainsi qu'au Nouveau-Brunswick (au bas du classement parmi les provinces canadiennes), on présentait un «Plan d'apprentissage de qualité» en 2002 et un «Plan d'amélioration en éducation» en 2007. En outre, les examens formatifs sont devenus sommatifs dans le système francophone du Nouveau-Brunswick. Au Québec, peu d'influences sont observés. Le processus curriculaire est cyclique (10 ans) et s'appuie actuellement sur les États généraux de l'éducation (1995). Enfin, nous pourrions penser que certains regroupements ont été initiés à la suite de la publication des résultats aux évaluations internationales, en particulier en ce qui concerne le Conseil Atlantique des ministres de l'éducation et de la formation (CAMEF) mis en place en avril 2004. Cette entente entre l'Île-du-Prince-Édouard, le Nouveau-Brunswick, la Nouvelle-Écosse et Terre-Neuve-et-Labrador vise à élaborer des projets qui répondent aux besoins

identifiés dans les secteurs d'éducation publique et d'éducation postsecondaire. Dans l'Ouest canadien, un Protocole de l'Ouest et le Nord canadien (Alberta, Colombie-Britannique, Yukon, Saskatchewan, Manitoba, Nunavut 2000) existe depuis 1993. Il semble toutefois que, lors de l'élaboration d'un cadre commun des programmes d'études, un souci de tenir compte des résultats des évaluations internationales se manifeste de la part de l'Alberta. Par exemple, les évaluations comportent des questions sur les notions de probabilité, ce qui a influencé le choix de commencer l'apprentissage de ces concepts de probabilité plus tôt. En Ontario, l'Office de la qualité et de la responsabilité en éducation existe depuis 1996. Cet organisme indépendant créé par le gouvernement de l'Ontario a le mandat de rehausser le niveau de la responsabilité et d'améliorer la qualité du système éducatif de la province en fournissant des renseignements sur le rendement des élèves à partir des résultats des tests écrits administrés annuellement.

En conclusion, au Canada, il est possible d'identifier des effets que j'appellerai «Effets directs» souvent immédiats. Il s'agit des comparaisons diffusés dans les médias ou de la reformulation des curriculums et les évaluations pour certaines provinces. D'autres effets, que j'appellerai «Effets indirects» semblent plus structurants et mobilisant comme le montrent les différents regroupements de proximité réalisés.

Références

BABY, A., DEBLOIS, L. (2005). La réussite éducative, d'hier à aujourd'hui. Dans L. DeBlois (Eds.), *La réussite scolaire: comprendre et mieux intervenir* (D. Lamothe coll.), pp. 1-14. Québec: Presses de l'Université Laval.

DEBLOIS, L., FRIEMAN, V. ROUSSEAU, M. (2008). Les résultats des élèves aux tests internationaux et leur possible influence sur les thèmes de recherche. *La didactique des mathématiques au Québec: genèse et perspectives*. Actes du colloque du Groupe des didacticiens des mathématiques du Québec 2007. 135-147.

LES ÉTATS GÉNÉRAUX DE L'ÉDUCATION (1995-1996). *Exposé de la situation*. Gouvernement du Québec. Ministère de l'Éducation, 1996 95-0963.

<http://www.quebec.ca/menu/tabmat.htm>

ST-HILAIRE, CINTHIA (2010) Une baisse de 30% en 40 ans. *La Voix de l'Est*. 18 février 2010. <http://www.cyberpresse.ca/la-voix-de-lest/actualites/201002/18/01-952839-une-baisse-de-30-en-40-ans.php> (Consulté en mars 2010).

D'AMOURS YVAN (2010) *La scolarité des francophones et des anglophones à travers les groupes d'âges au Québec et en Ontario*. Institut de la statistique du Québec Février Vol. 14 (2). http://www.stat.gouv.qc.ca/publications/conditions/pdf2010/sociodemobref_fev10.pdf

ROUSSEAU, M., FRIEMAN, V., SAVARD, D., & DEBLOIS, L. (2008). Les défis de l'enseignement et de l'évaluation des mathématiques chez les élèves francophones vivant en milieu minoritaire : Pistes de réflexion. *CD-Rom des Actes du colloque Espace Mathématique Francophone 2006, thème 7*, Sherbrooke, Québec.

Annexe 1

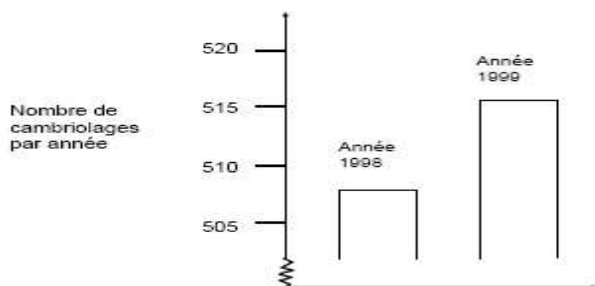
CAMBRIOLAGES

Question 1 : CAMBRIOLAGES

M179Q01 - 01 02 03 04 11 12 21 22 23 99

Lors d'une émission télévisée, un journaliste montre ce graphique et dit :

« Ce graphique montre qu'il y a eu une très forte augmentation du nombre de cambriolages entre 1998 et 1999. »



Considérez-vous que l'affirmation du journaliste est une interprétation correcte de ce graphique ? Justifiez votre réponse par une explication.

Crédit complet

Code 21 : Non, ce n'est pas correct. La réponse met l'accent sur le fait que seule une partie limitée du graphique est présentée.

- Ce n'est pas correct. Il aurait dû montrer la totalité du graphique.
- Je ne pense pas que ce soit une interprétation correcte du graphique, car s'ils avaient montré tout le graphique, on aurait vu qu'il y a eu seulement une légère augmentation des cambriolages.
- Non, parce qu'il a utilisé la partie supérieure du graphique, et si on avait regardé le graphique complet de 0 à 520, cela n'aurait pas augmenté tant que cela.
- Non, car le graphique donne l'impression qu'il y a eu un accroissement important, mais si on regarde les chiffres on voit qu'il n'y a pas eu une grosse augmentation.

Code 23 : Indique qu'il faut avoir des indications sur les tendances au cours du temps pour pouvoir former un jugement.

- On ne peut pas dire si l'accroissement est très fort ou non. Si le nombre de cambriolages en 1997 a été le même qu'en 1998, alors on pourrait dire qu'il y a eu un très fort accroissement en 1999.
- On ne peut pas savoir ce que veut dire « très fort », parce qu'il faut au moins deux changements pour dire que l'un est grand, l'autre petit.

SEANCE 3

La séance 3 a été consacrée à un travail sur l'impact des évaluations internationales sur l'enseignement et la formation. Nous reproduisons ci-après les textes associés à l'exposé de Richard Cabassut, chargé d'introduire ce thème et la réaction de Lucie DeBlois.

Impact sur l'enseignement et la formation des évaluations internationales à grande échelle à partir de l'exemple de PISA

Richard Cabassut
IUFM, Université de Strasbourg & LDAR, Université Paris 7

Introduction

Nous proposons de réfléchir sur l'impact des évaluations internationales. Nous nous limiterons à l'évaluation de PISA en observant son impact dans trois pays européens : le Bade-Wurtemberg, la France, et l'Andalousie, choisis pour la différence de leurs systèmes éducatifs et la différence de nature de l'impact. Nous questionnerons alors cette observation : Comment mesurer cet impact ? Quels changements observer ? Quelles questions pour l'enseignement et la formation mathématiques ?

Il est bien entendu difficile d'observer cet impact : les systèmes éducatifs sont des systèmes complexes soumis à des variables multiples qu'il est difficile d'isoler. Nous repèrerons, quand c'est possible, les déclarations et les textes officiels affirmant la prise en compte des évaluations internationales. Il faut ensuite préciser quels impacts sont observables.

Le programme PISA évalue « dans quelle mesure les élèves en fin d'obligation scolaire ont acquis certaines des connaissances et compétences qui sont essentielles pour pouvoir participer pleinement à la vie de la société » (OCDE 2007 p.18). En mathématiques, PISA évalue la « littératie » ou la « culture » mathématique, c'est-à-dire « l'aptitude d'un individu à identifier et à comprendre le rôle joué par les mathématiques dans le monde, à porter des jugements fondés à leur propos et à s'engager dans des activités mathématiques, en fonction des exigences de sa vie en tant que citoyen constructif, impliqué et réfléchi. Le terme de *culture mathématique* renvoie ici à l'utilisation fonctionnelle au sens large des mathématiques. La notion d'engagement fait référence à la capacité de reconnaître des problèmes mathématiques et de les formuler dans diverses situations » (OCDE 2007 p.23).

« L'enquête PISA étudie les capacités des élèves à analyser des idées mathématiques, à raisonner à leur propos et à les communiquer à autrui, au moment où ils posent, formulent, résolvent et interprètent des problèmes mathématiques relevant de diverses situations. Pour résoudre ces problèmes, les élèves doivent exploiter les savoir-faire et les compétences qu'ils ont acquis tout au long de leur scolarité et de leurs expériences de vie. Dans

l'enquête PISA, le terme de mathématisation désigne le processus fondamental appliqué par les élèves pour résoudre des problèmes de la vie courante » (OCDE (2006, p. 107).

« Les épreuves PISA sont constituées de questions demandant aux élèves d'élaborer leurs propres réponses ainsi que de questions à choix multiple. Les questions sont regroupées par unité. Ces unités s'articulent autour de textes ou de graphiques que les élèves sont susceptibles de rencontrer dans la vie courante ». (OCDE 2007, p.21)

Les notions suivantes sont donc importantes dans PISA : les compétences à acquérir, le processus de mathématisation, l'évaluation des compétences acquises par les élèves. Nous allons donc observer la place que ces trois notions, les compétences, la mathématisation et l'évaluation, peuvent avoir dans l'enseignement et la formation en mathématiques. Il est très difficile de savoir quelle est la part de l'impact de PISA sur la place accordée ces notions. Notre approche sera essentiellement qualitative.

1. Un nouveau curriculum en Allemagne : le cas du Bade-Wurtemberg

L'Allemagne est un pays constitué de seize états fédérés. Chaque état (Land) est responsable de l'organisation de l'enseignement primaire et secondaire. Une conférence des Ministres de l'Education et de la Culture (KMK) coordonne la coopération entre les différents états en matière d'éducation et de formation. La KMK relie explicitement la réforme de la mise en place de standards de formations (Bildungsstandards) aux résultats des évaluations internationales TIMSS et PISA : « En octobre 1997, la conférence des Ministres de la Culture et de l'Education (KMK) a décidé de laisser comparer le système scolaire allemand dans le cadre de recherches scientifiques internationales (décision de Constance). L'objectif est d'obtenir des diagnostics assurés sur les forces et les faiblesses des élèves dans les principaux domaines de compétences à obtenir. A travers les résultats de TIMSS, PISA et IGLU, il s'est avéré en Allemagne que ce n'est pas seulement le choix prioritaire des stimulations dans le système scolaire, qui conduit aux résultats souhaités. On doit ajouter le choix et la vérification des performances attendues [...] C'est pourquoi la conférence des Ministres de la Culture et de l'Education (KMK) a placé comme axe essentiel de son travail le développement et la mise en place de normes de formation (Bildungsstandards) valables dans l'ensemble du territoire fédéral¹⁸ » (KMK 2005 , p.5, trad. R.C.).

Nous allons illustrer cet impact sur un land, le Bade-Wurtemberg, qui a réorganisé à partir de 2004, son enseignement primaire et secondaire en reconnaissant l'influence de PISA :

« Beaucoup d'évolutions, de changements structurels et d'innovations, lancés expérimentalement depuis la dernière réforme du plan de formation en 1994, ont été intégrés dans la planification de la réforme du plan de formation de 2004. D'autres

18. Oktober 1997 hat die Kultusministerkonferenz beschlossen, das deutsche Schulsystem im Rahmen wissenschaftlicher Untersuchungen international vergleichen zu lassen (Konstanzer Beschluss). Ziel ist es, gesicherte Befunde über Stärken und Schwächen der Schülerinnen und Schüler in den zentralen Kompetenzbereichen zu erhalten. Durch die Ergebnisse von TIMSS, PISA und IGLU ist deutlich geworden, dass die in Deutschland vorrangige Inputsteuerung allein nicht zu den erwünschten Ergebnissen im Bildungssystem führt [...] Deshalb hat die Kultusministerkonferenz einen besonderen Schwerpunkt ihrer Arbeit auf die Entwicklung und Einführung von bundesweit geltende Bildungsstandards gelegt. (KMK 2005 , p.5, traduction Richard Cabassut)

développements ont été des réactions à des études internationales comme TIMSS et PISA, comme l'augmentation du poids des médias ou de l'enseignement des sciences¹⁹» (Ministerium 2009), trad. R.C.). L'étude PISA de 2003 portait effectivement sur les compétences en mathématiques et en sciences et a donné lieu à une étude élargie en Allemagne (Prenzel, Zimmer, 2006).

Alors que jusqu'aux programmes de 1994, le programme de mathématiques était constitué de contenus mathématiques répartis annuellement, les programmes de mathématiques de 2004 répartissent, par cycle de deux années, de l'école primaire à la fin de la scolarité secondaire, des compétences à atteindre en fin de cycle.

Examinons tout d'abord un exemple concernant la filière Gymnasium²⁰ qui prépare du grade 5 (10-11 ans) au grade 12 (17-18 ans) à l'entrée à l'Université.

1.1. L'entrée par les compétences

Les nouveaux programmes de mathématiques (Ministerium 2009b) distinguent quatre domaines de compétences : apprendre, justifier, résoudre un problème, communiquer. Huit idées directrices sont précisées : nombre, algorithme, variable, mesurer, espace et forme, relation fonctionnelle, données et hasard, mise en réseau, modéliser. On peut rapprocher ces domaines de compétences et ces idées directrices des capacités et des domaines de contenus proposés par PISA pour la scolarité obligatoire :

- quatre idées majeures sur les contenus : quantité, espace et forme, variations et relations, incertitude (OCDE 2006, p.108),
- huit catégories de compétences : pensée et raisonnement ; argumentation ; communication ; modélisation ; formulation et résolution de problèmes ; représentation ; utilisation d'un langage et d'opérations de nature symbolique, formelle et technique ; utilisation d'instruments et d'outils (OCDE 2006, pp. 111, 112).

On observe par exemple que la notion de modélisation est une idée directrice en Bade-Wurtemberg et une catégorie de compétence pour PISA.

Illustrons les compétences relatives à l'idée directrice de « nombre », pour les classes 5 et 6 (âge 10-12 ans).

« Les élèves disposent, en relation avec les idées directrices citées, des compétences suivantes :

Idée directrice « nombre »

- connaître différentes formes de représentation des nombres, les choisir adaptées à la situation et changer de l'une à l'autre,
- comparer et ordonner les nombres,
- conduire des calculs approximatifs et contrôler les résultats d'un calcul.

19. Viele Entwicklungen, strukturelle Änderungen und Innovationen, die seit der vergangenen Bildungsplanreform im Jahr 1994 in Schulversuchen auf den Weg gebracht wurden, flossen in die Planungen für den Bildungsplan 2004 ein. Andere Entwicklungen ergaben sich aus Reaktionen auf internationale Studien wie TIMSS und Pisa; wie die stärkere Gewichtung der Medien oder des naturwissenschaftlichen Unterrichts (Ministerium 2009, trad. Richard Cabassut)

20. En Bade-Wurtemberg (ce qui n'est pas le cas de tous les Länder), l'école primaire se termine au grade 4 (âge 9-10 ans) et se poursuit essentiellement en trois types d'écoles secondaires : la Hauptschule, la Realschule et le Gymnasium.

Contenu :

- Nombres entiers, nombres rationnels.
- Puissances de dix simples, fractions, fractions décimales, problèmes de pourcentages ». (Ministerium 2009b, p. 95, trad. R.C.)

Le Ministère met à la disposition des enseignants, sur des sites internet du Ministère, des ressources tirées d'expérimentations de ce programme qui, bien entendu, n'ont pas un caractère prescriptif. Voici un exemple décrivant une programmation annuelle possible d'une mise en œuvre possible du curriculum mathématique pour les classes 5 et classe 6. L'extrait choisi (Ministerium 2009c, p.1, trad. R.C.) concerne le thème des entiers naturels.

La colonne I décline des compétences pouvant être mises en œuvre. Ces compétences relèvent d'idées directrices différentes :

- 1 nombre : connaître différentes formes de représentation des nombres, comparer et ordonner des nombres ...
- 4 mesurer : comprendre la structure et l'utilisation d'un système d'unités de mesure ...
- 7 données et hasard : collecter, ranger et représenter clairement des données ...

La colonne II précise le thème traité : les entiers naturels.

La colonne III précise le contenu : les nombres et leurs représentations (les élèves représentent en diagramme des résultats de sondages effectués dans la classe ou autre), évaluer les grandeurs et mesurer (sur des exemples concrets, développer un « sens » des grandeurs) ...

La colonne IV précise la place dans l'année de l'expérimentation (septembre et octobre) et le nombre d'heures de cours consacrées (20 heures).

Les colonnes II à IV concernent le noyau du curriculum mathématique et représentent deux tiers du temps disponible.

La colonne V contient des informations sur le tiers de temps restant à consacrer au curriculum scolaire : découverte possible et approfondissement du curriculum scolaire / travail commun avec d'autres disciplines et réunion des disciplines. Le thème de début de classe 5 est « notre nouvelle école et notre nouvelle classe »²¹. Un objectif d'apprentissage prioritaire est : autonomie et capacité à travailler en équipe.

21 Rappelons qu'en Bade-Wurtemberg, l'école primaire se termine en classe 4 à l'issue de laquelle l'élève choisit entre trois écoles différentes : la Hauptschule, la Realschule et le Gymnasium. Donc le début de classe 5 de Gymnasium marque bien un changement d'école. Rappelons qu'au sein de la même école il est assez fréquent qu'un groupe d'élèves se suit de classe en classe au sein de la même école, le changement de classe signifiant un changement de groupe d'élèves principalement lors du changement d'école, ce qui est le cas du début de la classe 5.

I	Kerncurriculum 2/3 der Zeit			Hinweise auf das Schulcurriculum 1/3 der Zeit	
	II	III	IV		V
Kompetenzen	Thema	Inhalt	Monat/ Jahr	Zeit/ Unter- richts- stun- den	Mögliche Ergänzung und Vertiefung im Schulcurriculum/ Zusammenarbeit mit anderen Fächern und Fächerverbänden (nur Hinweise/Vorschläge)
Die Schülerinnen und Schüler können 1.1 verschiedene Darstellungsformen von Zahlen kennen, situationsgerecht auswählen und ineinander umwandeln; 1.2 Zahlen vergleichen und anordnen; 1.3 Überschlagsrechnungen durchführen und zur Kontrolle von Rechenergebnissen einsetzen. 4.1 die Struktur und den Gebrauch von Maßsystemen verstehen; 4.2 geeignete Maßgrößen und Einheiten nutzen, um Situationen zu beschreiben und zu untersuchen; 4.3 Maße schätzen und bestimmen; 4.4 Messergebnisse sachangemessen darstellen. 7.1 Daten systematisch sammeln, anordnen und übersichtlich darstellen; 7.2 Daten bewerten und aus ihnen Schlüsse ziehen.	Natürliche Zahlen	<ul style="list-style-type: none"> Zahlen und ihre Darstellung (Schüler stellen Umfrageergebnisse der Klasse u.a. als Diagramme dar) Größen schätzen und messen (An konkreten Beispielen ein „Gespür“ für Größen entwickeln) → Lernzirkel Mit Größen rechnen (einfache Umrechnungsaufgaben) Größen mit Komma Zusammenfassung und Übungen zum Thema Natürliche Zahlen 	09/10 2004	20	<p>→ Einstiegsthema in Klasse 5 : Unsere neue Schule und unsere Klasse</p> <p>→ Übergeordnetes Lernziel: Selbständigkeit und Teamfähigkeit</p> <p>Heftführung und Gestaltung</p> <p><i>Wiederholung (auch älterer Inhalte) zur Sicherung von Basiswissen</i></p>

Le Ministère propose un autre exemple de programmation annuelle du curriculum, correspondant à une mise en œuvre dans une autre école. L'idée centrale est que le curriculum n'est pas un produit fini mais un processus de développement prenant en compte les contextes des situations et de la classe. De manière générale, un site de ressources du Ministère (Ministerium 2009d) met à disposition les comptes rendus des expérimentations qui ont été conduites dès 2002 ainsi que différentes ressources (fiches de séquences d'activités pour la classe comprenant notamment des épreuves d'évaluation).

Les éditeurs scolaires ont proposé des manuels de classe pour élèves, des livres de professeurs les accompagnant et des ressources pour les professeurs conformes à ce nouveau curriculum (Lambacher Schweizer 2009). Des équipes de didacticiens et d'enseignants ont créé également des ressources pour l'enseignement des "Bildungsstandards" (Blum 2007).

Les universités proposent des formations initiales pour les étudiants se destinant à l'enseignement (Borromeo, Blum 2009) et des formations continues (Schmidt 2009) prenant en compte ce nouveau curriculum, par exemple des formations à la mise en œuvre de la compétence « modéliser ».

2. Le socle commun en France

2.1. Le socle commun de connaissances et de compétences à l'école obligatoire

En septembre 2007, le socle commun de connaissances et de compétences est mis en œuvre à partir de l'école primaire. Il est explicitement fait référence aux évaluations internationales, notamment PISA : « La scolarité obligatoire doit au moins garantir à chaque élève les moyens nécessaires à l'acquisition d'un socle commun constitué d'un ensemble de connaissances et de compétences qu'il est indispensable de maîtriser pour accomplir avec

succès sa scolarité, poursuivre sa formation, construire son avenir personnel et professionnel et réussir sa vie en société. [...] La définition du socle commun [...] se réfère enfin aux évaluations internationales, notamment au Programme international pour le suivi des acquis des élèves (PISA) qui propose une mesure comparée des connaissances et des compétences nécessaires tout au long de la vie » (BOEN 2006 p.III)). Concernant les capacités mathématiques, la modélisation apparaît explicitement : « À la sortie de l'école obligatoire, l'élève doit être en mesure d'appliquer les principes et processus mathématiques de base dans la vie quotidienne, dans sa vie privée comme dans son travail. Pour cela, il doit être capable :[...] - de saisir quand une situation de la vie courante se prête à un traitement mathématique, l'analyser en posant les données puis en émettant des hypothèses, s'engager dans un raisonnement ou un calcul en vue de sa résolution, et, pour cela : savoir quand et comment utiliser les opérations élémentaires ; contrôler la vraisemblance d'un résultat ; reconnaître les situations relevant de la proportionnalité et les traiter en choisissant un moyen adapté ; utiliser les représentations graphiques ; utiliser les théorèmes de géométrie plane [...] L'étude des sciences expérimentales développe les capacités inductives et déductives de l'intelligence sous ses différentes formes. L'élève doit être capable de pratiquer une démarche scientifique : savoir observer, questionner, formuler une hypothèse et la valider, argumenter, modéliser de façon élémentaire ; comprendre le lien entre les phénomènes de la nature et le langage mathématique qui s'y applique et aide à les décrire » (Ibid. p.VIII-IX).

A la fin de chaque cycle²² de la scolarité obligatoire, une grille de références est mise à la disposition des enseignants pour les aider à concevoir les évaluations nécessaires qui leur permettront de mettre en place les aides individualisées nécessaires et de renseigner le livret de compétences de chaque élève. Nous proposons, à titre d'illustration, l'extrait suivant de la grille de référence de fin de cycle 3 (fin de grade 5, autour de 11 ans).

22. La scolarité obligatoire se déroule essentiellement dans deux écoles successives à filière unique : l'école primaire (grade 1 à 5, de 6 à 11ans) et puis le collège (grade 6 à 9, de 11 à 15 ans). Elle se poursuit dans différents types de lycées (professionnels, techniques et généraux). Cette scolarité obligatoire est décomposée en cycles :

- fin du cycle 2 ou premier palier du socle commun : fin de grade 2 (CE1),
- fin du cycle 3 ou deuxième palier du socle commun : fin du grade 5 (CM2),
- fin du cycle d'adaptation : fin du grade 6 (6ème),
- fin du cycle central : fin du grade 8 (4ème),
- fin du cycle d'orientation : fin du grade 9 (3ème).

Connaissances et capacités attendues en fin de scolarité obligatoire	Éléments du socle attendus en fin de cycle 3	Indications pour l'évaluation dans des situations simples
Reconnaître des situations de proportionnalité, utiliser des pourcentages, des tableaux, des graphiques. Exploiter des données statistiques et aborder des situations simples de probabilité.	Se repérer sur une droite graduée et dans un plan quadrillé.	Sur une droite graduée de 0,1 en 0,1, de 1 en 1, de 10 en 10 ou de 100 en 100 : l'élève sait lire la position (abscisse) d'un point, et placer un point dont il connaît l'abscisse. Dans un plan quadrillé : lire la position d'un point (ou d'un nœud) ou d'une case, et placer un point ou une case à partir d'une description de sa position (coordonnées, trajet codé, ...).
	Lire, interpréter et construire quelques représentations : diagrammes, graphiques.	- identifier les variables représentées (temps, durée, température, hauteur, longueur, couleur, etc.), et lire leur valeur en respectant l'échelle. Exemples de supports : lignes brisées, diagrammes en bâtons ou histogrammes à pas réguliers. - compléter une ligne brisée ou un diagramme en bâtons, toutes les indications lui ayant été fournies.
	Se repérer dans l'espace : utiliser une carte, un plan, un schéma.	Sur un plan d'un lieu familier (classe, école, quartier) : - se situer - représenter un déplacement simple - évaluer la distance entre deux points ;
Connaître et utiliser les nombres entiers, décimaux et fractionnaires. Mener à bien un calcul mental, à la main, à la calculatrice, avec un ordinateur.	Connaître les désignations orales et écrites des nombres entiers et décimaux.	- lire, et écrire sous la dictée, en chiffres et en lettres, les nombres entiers jusqu'à la classe des millions et les nombres décimaux (3 chiffres après la virgule au plus) exprimés sous la forme ... unités et ... dixièmes, centièmes ou millièmes. - connaître la valeur des chiffres d'un tel nombre, et le décomposer en utilisant 10, 100, ... ou encore 0,1 ; 0,01 ; 0,001. - reconnaître les multiples de 2, de 5 et de 10.
	Ordonner ou comparer des nombres entiers, des nombres décimaux.	- comparer deux nombres entiers, ou deux nombres décimaux dont les parties décimales sont de même longueur. - encadrer un nombre décimal par deux nombres entiers consécutifs (les symboles < et > doivent être connus et utilisés). - produire des suites écrites ou orales de nombres de 0,1 en 0,1, à partir d'un nombre donnée, dans les deux sens.
	Calculer mentalement.	- les tables de multiplication "dans tous les sens" : récitation d'une table donnée, mais aussi réponse à des questions du type : $6 \times 4 = ?$ 24, c'est combien de fois 6 ? puis 24, c'est ? - les multiples courants de 25, 50 et 250 - les doubles, moitié et triple des nombres d'usage courant. Il sait aussi effectuer des calculs mentaux portant sur les quatre opérations, dans des cas où un calcul réfléchi simple peut être mis en œuvre.

Concernant les nouveaux programmes de mathématiques de 2008 de l'école primaire et du collège, nous n'avons pas observé de différences qui pourraient être expliquées par l'impact d'évaluations internationales, notamment PISA, si ce n'est la référence claire au socle commun.

Les programmes de l'école primaire de 2008 (BOEN 2008) marquent clairement la différence entre les compétences du programmes et celles du socle commun. Les programmes du collège de 2008 précisent également les compétences du socle commun : « les points du programme (connaissances, capacités et exemples) qui ne sont pas exigibles pour le socle sont écrits en italiques. Si la phrase en italiques est précédée d'un astérisque, l'item sera exigible pour le socle dans une année ultérieure. Dire que l'exigibilité pour le socle est différée ne veut pas dire que la capacité ne doit pas être travaillée – bien au contraire ! mais que les élèves pourront bénéficier de plus de temps pour la maîtriser. (BOEN 2008b, p.13).

Le livret de connaissances et de compétences de chaque élève enregistre la validation progressive de chacune des sept grandes compétences du socle commun de connaissances et de compétences à chacun des paliers (fin CE1, fin CM2, fin 6ème, fin scolarité obligatoire).

Voici un extrait de livret de compétences expérimentée au cours de l'année 2007-2008.

<p>ORGANISATION ET GESTION DES DONNÉES</p> <p>Savoir lire, utiliser et construire tableaux, diagrammes et graphiques</p> <p>Reconnaître des situations de proportionnalité, utiliser des pourcentages</p> <p>Exploiter des données statistiques et aborder des situations simples de probabilité</p> <p>NOMBRE ET CALCUL</p> <p>Connaître et utiliser les nombres entiers, décimaux et fractionnaires</p> <p>Mener à bien un calcul mental, à la main, à la calculatrice, avec un ordinateur</p> <p>Savoir quand et comment utiliser les quatre opérations</p> <p>GÉOMÉTRIE</p> <p>Connaître et représenter des figures géométriques et des objets de l'espace ; utiliser leurs propriétés</p> <p>Effectuer des tracés avec des instruments usuels (règles, équerres, compas)</p> <p>Connaître les notions géométriques de parallèle, perpendiculaire, médiatrice, bissectrice, tangente</p> <p>Utiliser les théorèmes de géométrie plane</p> <p>Se repérer dans l'espace</p>	<p>GRANDEURS ET MESURES</p> <p>Connaître les unités de mesures usuelles pour les longueurs (m, cm), masses (kg, g), contenances (litre), durées (jours, heures, minutes)</p> <p>Réaliser des mesures (longueurs, durées), calculer des valeurs (volumes, vitesses...) en utilisant différentes unités</p> <p>RÉSOLUTION DE PROBLÈMES</p> <p>Rechercher et organiser les informations</p> <p>Calculer, mesurer, appliquer des consignes</p> <p>Engager une démarche, raisonner, argumenter, démontrer</p> <p>Communiquer à l'aide d'un langage mathématique adapté</p> <p>PRATIQUER UNE DÉMARCHE SCIENTIFIQUE OU TECHNOLOGIQUE</p> <p>Observer, rechercher, et organiser les informations pour les utiliser</p> <p>Réaliser, manipuler, mesurer, calculer, appliquer des consignes</p> <p>Raisonner, argumenter, démontrer</p> <p>Communiquer à l'aide de langages ou d'outils scientifiques</p>
--	--

L'évaluation de l'acquisition des compétences du socle commun devrait faciliter la mise en place de programmes personnalisés de réussite éducative à l'école et au collège : lorsqu'il apparaît qu'un élève risque de ne pas maîtriser les connaissances et les compétences indispensables à la fin d'un cycle, le directeur d'école ou le chef d'établissement propose aux parents ou au responsable légal de l'élève de mettre en place un programme personnalisé de réussite éducative. Il est encore trop tôt pour évaluer ce dispositif. Avec la prise en compte du socle commun, une difficulté de l'enseignement au collège est de traiter la totalité du programme et d'assurer à tous les élèves la maîtrise des éléments du socle. Une pédagogie différenciée basée sur la résolution de problèmes et la mise en activité de la totalité des élèves devrait permettre de vaincre cette difficulté.

On observe donc peu de changements dans l'écriture des programmes. Quelques nouveaux domaines apparaissent. Au collège, "grandeurs et mesures" assureront une meilleure continuité avec l'école primaire où ce domaine était déjà présent ; ce domaine est favorable à l'émergence de problèmes de modélisation. Dans les projets de programmes de seconde, l'introduction des domaines "algorithmique" et "probabilités" (joint au domaine "statistiques" déjà présent), favorisent également les problèmes de modélisation. Ceci est conforté par l'apparition en 2006 des thèmes de convergence en collèges, et dans les projets de 2009 de thèmes d'études dans le programme de seconde, qui devraient favoriser le traitement de problèmes de modélisation, tout comme les travaux personnels encadrés en première et terminale.

Dans l'académie de Strasbourg ont été proposées, ces dernières années, des formations continues sur la modélisation pour professeurs de l'école primaire et pour les professeurs de mathématiques du secondaire, et une formation continue sur PISA pour professeurs scientifiques de l'école secondaire.

3. Evaluation diagnostique en Andalousie

En Espagne, l'éducation primaire et secondaire est de la responsabilité des communautés autonomes. Le décret royal du 29 décembre 2006 fixe un cadre commun à toutes les Communautés pour les enseignements minimaux de l'éducation secondaire obligatoire dans les termes suivants.

Les enseignements minimaux sont les aspects de base du curriculum référés aux objectifs, aux compétences de base, aux contenus et aux critères d'évaluation. L'objectif du décret royal est d'établir les enseignements minimaux de l'éducation secondaire obligatoire (qui dure 4 ans de 12 ans à 16 ans). La finalité des enseignements minimaux est d'assurer une formation commune à tous les élèves du système éducatif dans le système scolaire espagnol et de garantir la validation des diplômes correspondants. A la fin de la deuxième année, une évaluation diagnostique à valeur formative et d'orientation permet à chaque élève de déterminer les moyens adéquats d'amélioration pour achever son éducation secondaire obligatoire.

En mathématiques, il existe deux évaluations avec le même sujet pour toute l'Andalousie : une au milieu de l'enseignement primaire (de 6 à 12 ans), et une au milieu de l'enseignement secondaire obligatoire (de 12 à 16 ans). L'évaluation régionale dans le secondaire se situe en général au début de la troisième année. Pour l'élève, elle est diagnostique et formative (mise en place de soutien, aide à l'apprentissage ...) ; elle aide à l'orientation de l'élève; elle permet enfin d'avoir une vue systémique de l'enseignement des mathématiques à ce niveau.

Voici deux extraits du sujet d'évaluation diagnostique d'octobre 2006.

Les nombres.

Situation-problème : chez le marchand de fruits

Question 1 :

La mère des frères Jean et Antoine leur demande d'effectuer les achats suivants chez le marchand de fruits et légumes:

- $\frac{1}{2}$ kg de carottes à 0,70 €/kg,
- $\frac{1}{4}$ kg de piments à 2,20 €/kg,
- 1kg et $\frac{1}{2}$ d'oranges à 0,80
- 1 kg et $\frac{3}{4}$ de pommes à 1,40 kg.

Quel est le poids total des produits achetés? Explique comment tu obtiens le résultat.

Question 1	
Compétence	Organiser, comprendre et interpréter l'information
Elément de compétence	Identifier le sens de l'information numérique et symbolique
Contenu	Nombre et mesure
Barème	2 La réponse correcte est 4 kg. Elle s'obtient au moyen de la somme de toutes les quantités données.

	<p>On peut obtenir la somme des fractions $1/2+1/4+1/2+3/4=8/4=2$ et ajouter à ce résultat les 2 kg restant.</p> <p>On peut exprimer les fractions en nombres décimaux et les additionner : $0,5 + 0,25 + 1,5 + 1,75 = 4$</p>
1	<p>On donne la réponse correcte sans argumentation.</p> <p>On n'indique pas les unités.</p>
0	<p>Toute autre réponse.</p>

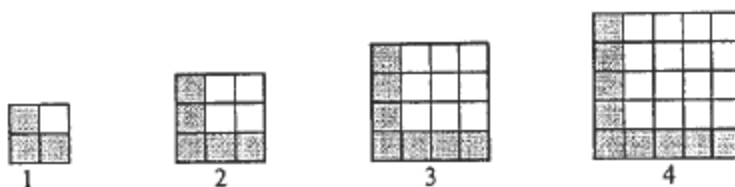
Question 2 :

La mère pense qu'on a dépensé davantage pour les fruits que pour les légumes. A-t-elle raison ? Expliquer.

Question 1		
Compétence	Exprimer	
Elément de compétence	Justifier les résultats exprimés en exprimant des arguments sur une base mathématique.	
Contenu	Nombre et mesure	
Barème	2	<p>La réponse correcte est qu'elle a raison parce que le coût des fruits est $1,5 \text{ kg} \times 0,8 \text{ €/kg} + 1,75 \times 1,4 \text{ €/kg} = 3,65 \text{ €}$ tandis que le coût des légumes est : $0,5 \times 0,7 \text{ €/kg} + 0,25 \times 2,2 \text{ €/kg} = 0,9 \text{ €}$.</p> <p>Il n'est pas nécessaire que les opérations figurent sous la forme ci-dessus, mais il est suffisant qu'apparaissent les calculs précédents.</p>
	1	<p>On répond qu'elle a raison mais on évoque une argumentation pauvre qui ne repose pas sur les opérations qui doivent être effectuées.</p>
	0	<p>Toute autre réponse.</p>

La géométrie

Situation-problème : L en construction



Question 7 :

Dans les dessins suivants on montre la construction des quatre premières figures d'une suite utilisant des carreaux noirs.

Considérons la question suivante : Combien de carreaux noirs faut-il pour dessiner la pièce qui occupe la place 100 ?

Pour répondre, on pourra commencer de différentes manières.

Une première manière serait :

Chaque branche du L contient 100 carreaux noirs et on ajoute celui du coin.

Une deuxième manière serait :

Une branche contient 100 carreaux noirs et l'autre 101, soit un total de ...

Une troisième manière serait :

La figure s'obtient en enlevant à un carré de 101 carreaux de côté un carré de côté 100 ...

Lis attentivement chacune des réponses précédentes et signale celle qui te paraît la plus simple, en expliquant pourquoi et en complétant pour terminer de répondre à la question sur le nombre de carreaux noirs nécessaire pour dessiner la pièce qui occupe la place 100.

Question 7	
Compétence	Poser et résoudre des problèmes.
Elément de compétence	Apprécier la pertinence de différentes façons de résoudre des problèmes avec une base mathématique.
Contenu	Géométrie
Barème	2 La réponse est 201 parce que : - si on choisit la première manière : $2 \times 100 + 1 = 201$ - si on choisit la 2ème manière : une branche a 100 carreaux et l'autre 101 soit un total de $100 + 101 = 201$ - si on choisit la 3ème manière : la figure s'obtient en enlevant à un carré de côté 101 un carré de côté 100 : $101^2 - 100^2 = 10201 - 10000 = 201$ On peut indiquer qu'aucune des manières n'est meilleure qu'une autre. Il faut argumenter si l'une paraît plus simple ; si on utilise le sens commun on considèrera la réponse comme valable
	1 Donne la réponse 201 en justifiant exclusivement la manière choisie pour calculer. Donne une réponse incorrecte, mais l'argumentation est valide.
	0 Toute autre réponse.

On observe que l'évaluation ne se fait pas sur la seule base de la correction des traitements ou la seule validité des raisonnements, mais qu'elle prend en compte des compétences plus larges. L'un des exercices fait appel à une situation de la vie quotidienne et effectivement, sur les 17 questions de cette évaluation, 14 concernent des situations en lien avec la réalité, ce qui n'est pas sans rappeler les situations de l'enquête PISA. On observe également l'évaluation en terme de compétences.

4. Conclusion

Il semblerait que PISA ait eu un impact important dans le Bade-Wurtemberg, avec une réforme importante du curriculum, introduisant avec force l'enseignement défini en termes de compétences et le thème de la modélisation. En France, le curriculum a peu changé ces dernières années mais l'introduction du socle commun correspond bien au même intérêt que PISA, de mesurer l'acquisition des compétences de base chez tous les élèves au cours de la scolarité obligatoire, et de transformer l'obligation de moyens en obligation de résultats. En Andalousie on retrouve, avec l'évaluation diagnostique, la volonté d'aider les élèves à acquérir les compétences minimales. Dans les trois pays on observe une reconnaissance du thème de la modélisation dans les programmes et dans les évaluations, comme le propose PISA.

Mais redisons la complexité des systèmes éducatifs, la difficulté à attribuer un changement observé à l'impact de PISA, la difficulté à observer la distance entre le curriculum officiel et le curriculum réel. Cependant il existe des cas où cet impact est réel. Le programme européen LEMA (Cabassut 2008, Lema 2009) en est une illustration. Ce programme vise à construire une formation à l'enseignement de la modélisation et des applications à destination des enseignants de mathématiques de l'école primaire. Il adopte clairement le cadre théorique de PISA.

Mais s'il est difficile d'assurer que les changements observés sont un impact de PISA, rien ne nous interdit de questionner ces changements. L'importance donnée à l'évaluation ne risque-t-elle pas de développer une culture de l'évaluation (teaching to the test) contre une culture de la réflexion et de l'apprentissage ? Modéliser, n'est-ce pas une compétence trop complexe, dont la trop grande importance pourrait être un obstacle à l'apprentissage ? PISA est orienté vers les productions du système éducatif (output orientation) (évaluation des élèves en sortie du système obligatoire) alors que les systèmes sont réformés par les entrées dans le système éducatif (nouveau curriculum, nouvelles ressources, nouvelles formations des maîtres). Comment concilier ces orientations ?

PISA développe des dispositifs pour repérer les élèves en difficultés et les trois exemples précédents en font de même. Par contre les dispositifs efficaces pour aider les élèves en difficultés ne sont pas évoqués par PISA et n'apparaissent pas clairement dans les exemples observés. Et c'est sans doute le défi à relever pour l'espace mathématique francophone : à la suite des évaluations internationales et nationales qui repèrent les élèves en difficulté, comment aider ces élèves ?

Références :

BLUM W. (2007). *Bildungsstandards Mathematik: konkret. Sekundarstufe 1: Aufgabenbeispiele, Unterrichts Anregungen, Fortbildungsideen*. Editeur Cornelsen-Scriptor.

BORROMEO FERRI R., BLUM W. (2009): Mathematical modelling in teacher education – experiences from a modelling seminar. In *Proceedings of 6th Cerme (congress of European society for research in mathematics education)*. Lyon. France.(à paraître).

BOEN Bulletin Officiel de l'Education Nationale (2006) Socle commun de connaissances et de compétences, *Bulletin officiel de l'éducation nationale* n° 29 du 20 juillet 2006.

BOEN (2008) Horaires et programmes d'enseignement de l'école primaire. *Bulletin officiel de l'Education Nationale* Numéro hors-série n°3 du 19 juin 2008.

BOEN (2008b) Programmes du collège. *Bulletin officiel de l'Education Nationale* spécial n°6 du 28 août 2008.

- CABASSUT Richard (2008) Problèmes dans un exemple de formation continue à la modélisation. *Actes du XXXV^{ème} Colloque Copirelem*. Bombannes.
- KMK Kultusministerkonferenz (2005) *Bildungsstandards der Kultusministerkonferenz Erläuterungen zur Konzeption und Entwicklung*. Wolters Kluwer Deutschland GmbH, München, Neuwied.
- Lambacher Schweizer (2009) <http://www.klett.de> . Accès à ce site le 23/06/09.
- LEMA (2009) Site du projet européen Comenius <http://www.lemma-project.org/>
- Ministerium Baden-Württemberg (2006) *Schulversuch am Bildungszentrum Nord, Gymnasium, Reutlingen. Erprobung der Bildungsstandards im Fach Mathematik Klasse 5 Schuljahr 2002 / 03 Juni 2006*. Accès à ce site du Ministère de l'Education du Bade-Wurtemberg le 23/06/09.
- Ministerium Baden-Württemberg (2009) *Schule in Baden-Württemberg Allgemein bildende Schulen* <http://www.kultusportal-bw.de/servlet/PB/-s/1w9knnw1oijd6v165r5sbjz4zec4chcna/menu/1176661/> Accès à ce site du Ministère de l'Education du Bade-Wurtemberg le 23/06/09.
- Ministerium Baden-Württemberg (2009b) *Bildungsstandards für Mathematik Gymnasium – Klassen 6, 8, 10, Kursstufe* pp. 92-102 http://www.bildung-staerkt-menschen.de/service/downloads/Bildungsstandards/Gym/Gym_M_bs.pdf Accès à ce site du Ministère de l'Education du Bade-Wurtemberg le 25/06/09.
- Ministerium Baden-Württemberg (2009c) *Bildungsplan 2004. Allgemein bildendes Gymnasium. Kerncurriculum für Mathematik Standard 6 Beispiel 1*. http://www.bildung-staerkt-menschen.de/unterstuetzung/schularten/Gym/curricula/kerncurricula/GY_M_5_6_Bsp1_kc.doc Accès à ce site du Ministère de l'Education du Bade-Wurtemberg le 26/06/09.
- Ministerium Baden-Württemberg (2009d) *Umsetzungsbeispiel-Mathematik* <http://www.bildung-staerkt-menschen.de/service/downloads/Umsetzungsbeispiele/Gym/M> Accès à ce site du Ministère de l'Education du Bade-Wurtemberg le 26/06/09.
- Ministère français (2007) *Socle commun de connaissances et de compétences. Livret de connaissances et de compétences. Grille de référence « Les principaux éléments de mathématiques et la culture scientifique et technologique »*. Ministère de l'Education Nationale. Eduscol. DGESCO. Octobre 2007.
- OCDE (2006). *Compétences en sciences, lecture et mathématiques : le cadre de l'évaluation de PISA 2006*. Editions OCDE. Paris.
- OCDE (2007) *PISA 2006 Les compétences en sciences, un atout pour réussir Volume 1 – Analyse des résultats*. Editions OCDE. Paris.
- PRENZEL M., ZIMMER K. (2006). Les prolongements de PISA 2003 en Allemagne : principaux résultats et enseignements, *Revue française de pédagogie*, n°157.
- SCHMIDT B. (2009). Modelling in the classroom – motives and obstacles from the teacher's perspective. In *Proceedings of 6th Cerme (congress of European society for research in mathematics education)*. Lyon. France (à paraître).
- Junta de Andalucía (2006). *Prueba de la evaluación de diagnóstico . Competencias básicas en matemáticas*. Consejería de Educación. Dirección General de Ordenación y Evaluación Educativa. <http://www.adelat.org/media/docum/prueba.pdf> Accès à ce site le 25/06/09.

En réaction au texte de Richard Cabassut.

Lucie DeBlois, Université Laval

L'auteur tente de mesurer l'impact des évaluations internationales, notamment celle du PISA, en étudiant les changements observés dans les programmes de trois unités administratives dans trois pays. Il étudie surtout des effets que j'appellerai directs. Ainsi, dans deux des trois pays, les apprentissages à réaliser se définissent selon la notion de compétences. Dans le troisième pays, un accent est développé pour l'évaluation diagnostique. Le barème critérié qui est développé serait sous l'influence de PISA. Enfin, la modélisation semble prendre de l'importance pour les trois unités administratives étudiées. Ainsi, dans un premier temps, des comparaisons sont faites avec les résultats des états situés à proximité culturelle ou physique. Pour étudier l'impact sur l'enseignement et la formation des évaluations internationales à grande échelle, il sera nécessaire de développer des recherches à partir des données secondaires de ces évaluations (Rousseau et al, 2008) et de se doter de cadres théoriques qui permettront une approche systémique. Certains cadres théoriques sont développés autour d'effets directs. Ainsi, Alexander (2009) identifie des conditions qui permettent une comparaison entre «les pédagogies» de cinq pays. En considérant un modèle générique d'enseignement (structure, forme, pratiques), les théories et les valeurs sous-entendues dans les pratiques, le système éducatif et la culture sociétale, il devient possible d'étudier ce qu'il appelle les spécifications, les translations, les transpositions et les transformations du curriculum. Par exemple, nous étudierons les curriculums comme l'a fait notre collègue pour considérer les spécifications. Nous analyserons l'interprétation, les modifications et les ajouts de sens que les enseignants et les élèves donnent à ce curriculum, ce que l'auteur appelle la translation. L'étude des ajustements faits par les enseignants pour planifier une situation d'enseignement-apprentissage permettra de cerner la transposition faite, rejoignant ainsi les travaux de Chevallard (1996). Enfin, l'étude du passage entre les prescriptions et les actions dans les tâches et les activités manifesterait des transformations réalisées. Des études de cas dans des contextes institutionnels différents (DeBlois, Roditi, Freiman, en cours) pourraient mettre en évidence les contrastes et les éléments communs pour étudier le poids de l'institution (Chevallard, 1996), les ressources interprétatives et les ressources d'actions des enseignants (Bednarz et al, 2001), la sensibilité des enseignants (DeBlois, 2006) situant ainsi les activités des élèves dans les ressources structurant les apprentissages (Lave, 1988; DeBlois, 2003) et précisant les interactions porteuses d'apprentissage dans leurs contextes. L'auteur identifie aussi certains effets indirects. Ainsi, le programme européen LEMA évoqué pourrait être considéré non seulement comme une illustration de l'influence de Pisa sur la formation à l'enseignement, mais comme un exemple d'effet structurant. De même, certains effets pourraient être considérés mobilisant. Ce sont les réactions consistant à regrouper des acteurs de l'éducation autour de questions évoquées par l'auteur comme celle des besoins régionaux ou encore de la responsabilité de l'éducation offerte. Il nous faudra mettre au point de nouveaux cadres théoriques qui permettent l'étude d'effets indirects, notamment en articulant et en intégrant les cadres théoriques et les résultats des recherches réalisées dans le champ de la recherche comparative.

Références

ALEXANDER, R. (2009) Towards a Comparative Pedagogy. *International Handbook of Comparative Education*. R. Cowen and A. Kazamias (eds.), Springer Science + Business Media B.V. 923-941.

BEDNARZ, N., DESGAGNÉ, S., DIALLO, P., & POIRIER, L. (2001). Approche collaborative de recherche : une illustration en didactique des mathématiques. Dans P. Jonnaert et S. Laurin (Dir.), *Les didactiques des disciplines, un débat contemporain*, (pp. 177-207). Québec: Presses de l'Université du Québec.

CHEVALLARD, Y. (1996). *La Transposition didactique*. Grenoble : La Pensée sauvage.

DEBLOIS, L. (2003). Interpréter explicitement les productions des élèves : une piste... *Éducation et Francophonie*, XXXI(2), 176-198.

DEBLOIS, L. (2006). Influence des interprétations des productions des élèves sur les stratégies d'intervention en classe de mathématiques. *Educational Studies in Mathematics*, 62(3), 307-329.

LAVE, J. (1988). *Cognition in Practice*. Cambridge University Press, 206 pages.

ROUSSEAU, M., FRIEMAN, V., SAVARD, D., & DEBLOIS, L. (2008). Les défis de l'enseignement et de l'évaluation des mathématiques chez les élèves francophones vivant en milieu minoritaire : Pistes de réflexion. *CD-Rom des Actes du colloque Espace Mathématique Francophone 2006, thème 7*, Sherbrooke, Québec.

SEANCE 4

La quatrième et dernière séance du projet spécial a été principalement consacrée à l'impact des évaluations internationales sur la recherche didactique et nous reproduisons ci-après le texte associé à l'exposé d'introduction présenté par Michèle Artigue et Carl Winslow.

L'enseignement des mathématiques dans une perspective internationale: d'évaluations vers la recherche

Michèle Artigue, LDAR, Université Paris Diderot – Paris 7
Carl Winslow, IND, Université de Copenhague

Résumé : L'objet de cette contribution est de présenter un cadre inspiré par la théorie anthropologique du didactique pour l'analyse des études comparatives en didactique, d'illustrer son fonctionnement et montrer sa pertinence, en considérant non seulement les évaluations internationales à grande échelle comme PISA qui sont au cœur de la thématique de ce projet spécial mais aussi des exemples illustrant la diversité et la richesse des recherches comparatives qui se sont développées au cours de la dernière décennie, et enfin d'inférer de cette analyse des besoins et perspectives de recherche, en ayant à l'esprit plus particulièrement la communauté francophone de didactique des mathématiques.

Introduction

Dans cette contribution, nous élargissons, comme précisé dans l'introduction générale, la perspective du projet spécial à celle, plus générale, des études comparatives internationales qui se sont multipliées pendant la dernière décennie. Nous présentons d'abord un cadre d'analyse dérivé de l'approche anthropologique du didactique (notée TAD dans la suite) qui nous semble pertinent pour approcher la diversité des travaux menés dans ce domaine et en fonder une méta-analyse. Nous présentons ensuite de façon synthétique les résultats de la mise en œuvre que nous en avons faite, dans le cadre de la préparation de ce projet spécial, en considérant un échantillon de travaux comparatifs nous semblant raisonnablement représentatif de la diversité existante et sans nous limiter à la communauté francophone. Cet échantillon contient en effet d'une part des évaluations à grande échelle comme le sont les évaluations TIMSS et PISA amplement discutées dans ce projet spécial et des études internationales impliquant plusieurs pays comme la *TIMSS Video Study*, la *Learners' Perspective Study* ou l'Etude ICMI 13 comparant les cultures éducatives de pays d'Asie de l'Est et de pays de l'Ouest, d'autre part des comparaisons binaires effectuées dans le cadre de thèses doctorales récentes préparées en France. Il s'agit plus particulièrement dans ce cas de sept thèses soutenues à l'université de Grenoble 1 dans le cadre d'une collaboration avec l'université pédagogique d'Ho-Chi-Minh au Vietnam et de quatre thèses comparatives préparées au sein de l'équipe DIDIREM de l'université Paris Diderot – Paris 7. A l'aide du cadre d'analyse élaboré, nous essayons de mettre en évidence les caractéristiques de ces différents travaux comparatifs et d'en pointer les potentialités et limites. Nous concluons en inférant de ce travail de méta-analyse des besoins et perspectives de recherche, en particulier en ce qui concerne la communauté francophone de didactique des mathématiques.

1. Un cadre d'analyse inspiré de la TAD

La TAD, à travers sa modélisation des pratiques humaines en termes de praxéologies, l'attention qu'elle porte aux différentes conditions et contraintes qui façonnent l'enseignement des mathématiques et tout particulièrement l'échelle des niveaux de co-détermination, plus récemment inscrite dans cette théorie, nous semble un cadre particulièrement utile pour analyser les études comparatives qui se sont multipliées dans la dernière décennie, les situer les unes par rapport aux autres, penser leurs connexions et complémentarités possibles. Dans cette première partie, nous présentons de façon très synthétique les éléments de ce cadre théorique que nous utilisons dans cette étude.

1.1. La notion de praxéologie

Dans le cadre de la TAD, l'activité mathématique comme tout type d'activité humaine est organisée autour de la réalisation de tâches au moyen de techniques qui dépendent non seulement de la tâche mais de l'environnement institutionnel et culturel dans lequel se situe la réalisation. Plus précisément, les activités humaines sont modélisées en termes de praxéologies. Dans sa forme élémentaire, celui des praxéologies ponctuelles, le modèle praxéologique est constitué d'un bloc pratique : un type de tâche et une technique pour accomplir cette tâche et d'un bloc théorique formé lui-même d'un discours technologique qui sert à décrire, expliquer, justifier la technique, et d'une théorie qui joue le rôle de technologie pour le discours technologique lui-même, le rapportant à un ensemble structuré d'objets et d'énoncés. Ainsi, la résolution d'une équation du second degré par la technique classique du déterminant peut être expliquée par un discours technologique s'appuyant sur la mise sous forme canonique, qui s'inscrit lui-même dans le cadre de la théorie des équations algébriques. Pour les besoins de l'analyse comparative, il est important de ne pas se limiter à ce niveau de modélisation mais de considérer comment s'organisent et se structurent progressivement les praxéologies. Dans la TAD, ceci s'effectue par l'identification de regroupements de praxéologies ponctuelles en organisations locales, puis régionales et globales, selon les éléments partagés. Un ensemble de praxéologies partageant la même technologie constitue une organisation locale de praxéologies, et des organisations locales partageant une même théorie ou, plus souvent, une même partie de théorie constituent une organisation régionale. Selon les institutions et les cultures, les praxéologies ponctuelles mais aussi la façon dont elles se structurent progressivement se différencient. Nous avons ici considéré une modélisation des activités mathématiques, les praxéologies en jeu étant des praxéologies mathématiques, mais l'analyse comparative ne peut se limiter à celle des organisations mathématiques, elle doit nécessairement prendre en charge les formes de l'enseignement au-delà de son contenu. C'est pourquoi interviennent dans la TAD, en relation de co-détermination avec les praxéologies mathématiques des praxéologies didactiques, dont les types de tâches sont celles de la professionnalité enseignante. Un défi important est ici que les praxéologies didactiques sont souvent lacunaires, le discours technologique étant peu développé et explicite, et difficile à relier, même lorsqu'il existe, à des cadres théoriques précis. Ceci contribue aux difficultés rencontrées, dans de nombreuses études comparatives, dans le codage et l'analyse de données

vidéo issues de différents contextes, et l'attention méthodologique de plus en plus portée à ces questions.

Nous avons évoqué dans le paragraphe précédent la co-détermination existant entre organisations mathématiques et didactiques. Dans les dernières années, cette idée de co-détermination a pris une signification à la fois plus large et plus précise dans la TAD, aboutissant à l'échelle des niveaux de co-détermination didactique que nous présentons ci-après.

1.2. L'échelle des niveaux de co-détermination didactique

Comme nous l'avons souligné dès le début de ce texte, un point fondamental dans la vision portée par la TAD est que les activités humaines et les organisations praxéologiques associées sont marquées par les contextes institutionnels dans lesquels elles se situent. Ces contextes déterminent des conditions et des contraintes qui conditionnent les formes praxéologiques susceptibles de se développer et de vivre, elles en déterminent l'écologie. De plus, comme les organisations praxéologiques, les institutions présentent souvent une structure emboîtée, les différents niveaux s'influencent mutuellement. L'échelle des niveaux de co-détermination représentée dans la figure 1 vise à modéliser cette situation (Chevallard, 2002). Il est important de préciser que lorsqu'il est question de discipline, domaine , ce dont il est question c'est la façon dont ces objets se réalisent dans le contexte d'enseignement concerné. Il est aussi important de souligner que de nombreuses conditions, en particulier celles se situant aux niveaux les plus élevés de la hiérarchie, ne peuvent pas être modifiées par un enseignant individuel. Certaines peuvent l'être cependant par des groupes constitués d'enseignants au sein d'associations ou par l'action de la noosphère du système éducatif (Chevallard, 1985).

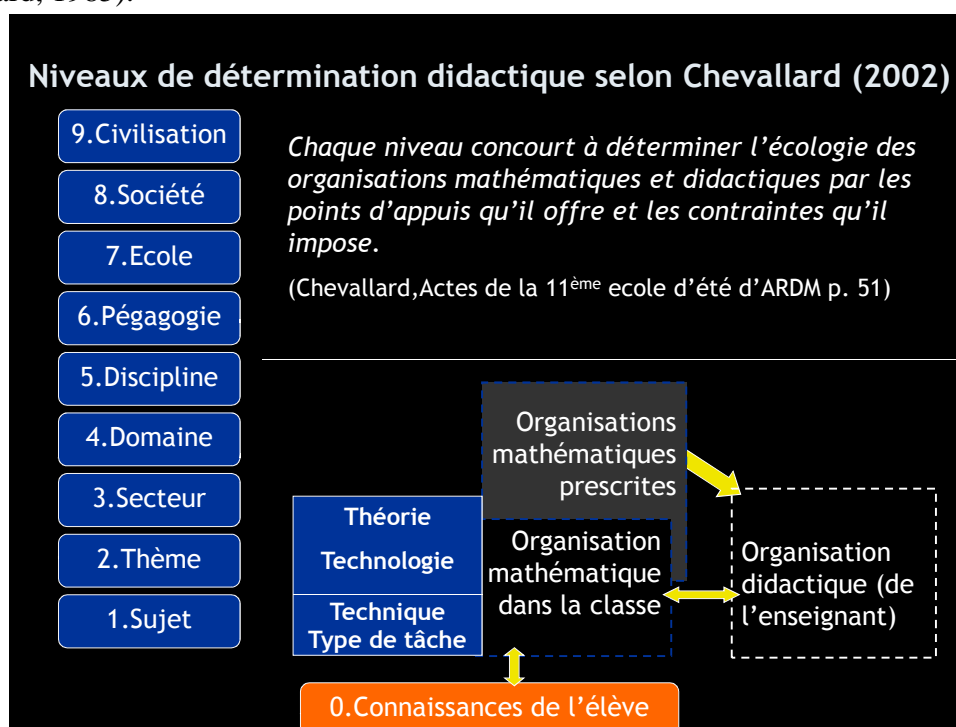


Figure 1 : L'outil d'analyse construit

Dans cette échelle, l'échelon médian et celui de la discipline, ici les mathématiques, mais le modèle se veut d'une portée plus générale. Par rapport à ce niveau, sont définis quatre sous-niveaux, ceux des domaines, secteurs, thèmes et sujets. Un *domaine* est une collection d'organisations régionales engageant plusieurs théories et formant une part bien identifiée dans le contexte institutionnel d'enseignement de la discipline concerné (par exemple l'algèbre, la géométrie, l'analyse). Un *secteur* se caractérise par une théorie unifiant une ou plusieurs organisations régionales. En algèbre, on peut penser par exemple au secteur des polynômes. Les différents types de tâches concernant ces objets dans l'enseignement peuvent être regroupés en *thèmes* unifiés par un même ensemble technologique, comme celui des équations polynomiales. Enfin, au premier niveau, se trouvent les *sujets*. La résolution des équations du second degré par la formule usuelle mentionnée plus haut en est un exemple. Comme on le perçoit bien dans cette description, même s'ils ne sont pas univoques, des liens évidents existent entre la hiérarchie de ces sous-niveaux et celle des organisations mathématiques présentée dans la partie précédente. De plus, il nous semble important de souligner, une fois de plus, que l'exploitation de cette hiérarchie, notamment via l'identification de structures verticales allant des sujets aux domaines, va bien sûr dépendre des contextes institutionnels analysés.

Les niveaux supérieurs à celui de la discipline : *pédagogie, école, société, civilisation*, expriment, quant à eux, les influences qui s'exercent plus globalement sur l'enseignement des mathématiques, en allant de celles exercées par les modèles pédagogiques prônés à un moment et dans un contexte donné, à des influences plus pérennes et partagées par de nombreux systèmes éducatifs comme celles se situant au niveau d'une civilisation.

Tous ces niveaux déterminent et conditionnent l'activité d'enseignement mais ils sont extérieurs à la situation d'enseignement elle-même. Les connaissances personnelles (à la fois dans leurs composantes pratiques et théoriques) que les élèves construisent à travers leur participation à ces situations, varient suivant les élèves et peuvent être distantes des attentes des enseignants, de l'école et de la société. C'est d'ailleurs un des points mis en évidence par les évaluations internationales qui font l'objet de ce projet spécial. C'est pourquoi nous rajoutons dans le modèle que nous construisons pour l'analyse des études comparatives, ce « produit final » de l'entreprise d'enseignement, que nous situons au niveau 0.

1.3. Niveaux de comparaison suggérés par cet outil d'analyse

L'outil d'analyse que nous avons élaboré nous conduit à envisager, pour les études comparatives, les dix niveaux suivants de comparaison entre deux contextes ou plus.

0. Les connaissances des étudiants sur un ou plusieurs sujets, situés au sein de thèmes et secteurs
1. Les praxéologies existantes sur des sujets spécifiques, telles que prescrites dans les documents curriculaires, les évaluations officielles, ou observées dans les pratiques réelles des contextes concernés
2. Les organisations locales associées à des thèmes spécifiques, telles que prescrites dans les documents curriculaires, décrites par les enseignants ou inférées de l'observation de l'enseignement de sujets relevant du thème

3. Les organisations régionales de secteurs spécifiques, telles que prescrites dans les documents curriculaires, décrites par les enseignants ou inférées de l'observation de l'enseignement de plusieurs thèmes relevant du secteur (études de niveau 2)
4. Les organisations globales de domaines spécifiques, telles que prescrites dans les documents curriculaires, décrites par les enseignants ou inférées de l'observation de l'enseignement de plusieurs secteurs relevant du domaine (études de niveau 3)
5. Les organisations d'une discipline en domaines ou plus globalement, sur la base de documents curriculaires ou d'autres données (incluant par exemple des déclarations d'enseignants)
6. Les formes pédagogiques, telles que prescrites par les écoles ou les programmes, observées ou décrites par les enseignants (ceci incluant des principes pour l'enseignement transcendant ou spécifique aux disciplines enseignées, les relations prescrites ou observées entre disciplines...)
7. Les conditions et caractéristiques d'institutions d'enseignement spécifiques, par exemple les obligations et l'autonomie des enseignants
8. Les conditions et caractéristiques des sociétés, incluant en particulier la façon dont l'École est administrée, financée et structurée
9. Le contexte plus large des cultures ou civilisations, leurs principes pour la société humaine, en particulier en ce qui concerne le rôle et le sens de l'éducation

Les études comparatives concrètes peuvent inclure plusieurs de ces niveaux, et un point crucial dans leur analyse est de déterminer jusqu'à quel point et comment les différents niveaux de comparaison envisagés sont reliés, et en particulier quelles relations de causalité sont inférées ou affirmées entre ces niveaux de comparaison. De plus, il nous semble essentiel de considérer si la question de la comparabilité elle-même est posée, et en particulier de déterminer si les discours technologiques des contextes considérés, aux différents niveaux des pratiques institutionnelles, font l'objet d'un examen critique, et si c'est le cas avec quels moyens.

En résumé, pour identifier et caractériser une étude comparative, nous proposons de considérer les quatre questions suivantes :

- a) Auquel ou auxquels des niveaux 0-9 définis ci-dessus la comparaison est-elle menée ?
- b) En supposant que la comparaison engage des organisations mathématiques et didactiques, quels moyens sont utilisés pour la fonder ? (ex., observations directes, entretiens avec des enseignants, analyse de manuels ...)
- c) Quels outils méthodologiques (en particulier, technologiques au sens de la TAD) sont-ils utilisés pour l'interprétation et pour s'assurer que la comparaison fait sens à un niveau donné ? (ex., au niveau 2, comment des thèmes comparables sont-ils identifiés et comment la comparabilité est-elle justifiée? En particulier comment les différences de langue sont-elles prises en compte ?)
- d) Comment l'étude relie-t-elle *comparaison horizontale* (entre deux contextes au même niveau) et *relations verticales* entre les niveaux trouvés dans chaque contexte ?) En particulier qu'en est-il des relations causales formulées ?

Toutes ces questions ne sont pas d'égale importance pour toutes les études mais, dans ce qui suit, nous essayons de montrer leur pertinence, en considérant un échantillon assez large de

travaux comparatifs, sans nous limiter au monde francophone. De fait, notre corpus est double. Il comprend d'une part un ensemble de travaux impliquant un nombre important de pays et en particulier les évaluations internationales comme TIMSS et PISA qui ont motivé ce projet spécial, d'autre part un ensemble de thèses de doctorat comparatives préparées en France et où la comparaison, plus limitée vu les contraintes d'un travail de thèse, concerne deux pays.

2. Quelques grandes études comparatives internationales

Dans cette partie, nous évoquons successivement trois types de travaux : les séries d'enquêtes quantitatives de TIMSS et PISA ; deux études qualitatives à grande échelle : la *TIMSS Video Study* et la *Learners' Perspective Study*, et enfin la 13^{ème} étude internationale de la CIEM, dont le thème était la comparaison de l'enseignement mathématique dans deux contextes culturels : l'Occident et l'Extrême-Orient. Dans chaque cas, nous nous appuyons sur le cadre théorique développé dans la partie précédente pour situer le travail comparatif mené. Le lecteur intéressé pourra se référer à (Artigue & Winslow 2010) pour un exposé plus détaillé.

2.1. Les enquêtes TIMSS et PISA

TIMSS (*Trends in mathematics and science study*) et PISA (*Programme for international student assessment*) sont très certainement les enquêtes internationales les plus connues. Dans chaque cas, l'objectif est de mesurer les compétences mathématiques des élèves dans un grand nombre de pays. La comparaison principale est donc horizontale et se situe au niveau 0. Mais elle comporte aussi un recueil de données se situant à différents niveaux dans l'échelle de co-détermination, niveaux 2 à 8 pour TIMSS, niveaux 6 à 8 pour PISA. Ces données dites « de base » (*background data*) sont censées fournir, de par les corrélations repérées avec les données principales au niveau 0, des hypothèses d'explication pour les différences majeures observées entre les prestations des élèves, suivant les pays.

Les deux enquêtes ont en commun d'être des opérations organisées par des consortiums internationaux, sous les auspices de l'IEA (Association internationale pour l'évaluation du rendement scolaire) pour TIMSS, et de l'OCDE pour PISA. Les enquêtes se répètent à des intervalles plus ou moins réguliers. La participation d'un pays est normalement décidée et financée au niveau gouvernemental, et le nombre de pays participants varie donc d'une enquête à l'autre, avec une tendance pourtant croissante. Il y a aussi des différences majeures entre TIMSS et PISA, que nous allons décrire rapidement.

TIMSS a pour origine deux études internationales sur l'état de connaissances des élèves en science et en mathématiques, menées dès les années 60 (l'IEA ayant été fondée en 1958). La première enquête de la série actuelle date de 1995, et a été suivie des enquêtes de 1999, 2003 et 2007. La particularité majeure de TIMSS, par rapport à PISA, est de considérer plusieurs niveaux d'âge (grade 4, 8 et 12) et d'adapter la mesure aux programmes nationaux dans le sens où les élèves ont à travailler des items correspondant à des objectifs réellement au programme de leur pays. La comparaison n'est donc, en principe, que partielle. Afin de rendre possible cette différenciation des tests distribués, un cadrage pour identifier les éléments des programmes a été construit. Il permet de regrouper les items selon des « objectifs », des

« secteurs de thèmes » et des « domaines de contenus », qui correspondent respectivement très bien aux niveaux 2, 3 et 4 de l'échelle de co-détermination, les items individuels correspondant plutôt à un sujet (niveau 1). Le cadrage est bien sûr en développement continu, dans une collaboration entre experts nationaux qui sont aussi chargés de repérer les parties pertinentes pour leur pays. Mullis, Martin, Ruddock, O'Sullivan, Arora & Erberber (2005) fournit un aperçu de l'état actuel de ce cadrage. En plus du test élève, des données sont recueillies dans chaque pays sur les conditions de l'enseignement, situées dans notre modèle aux niveaux 6, 7 et 8.

L'enquête PISA est conduite tous les trois ans depuis l'an 2000, 66 pays y ayant participé en 2009. Elle est fondée sur une vision unique des compétences mathématiques (entre autres) requises de tous les citoyens des pays de l'OCDE, à l'issue de la scolarité obligatoire (OCDE, 2006). Une seule tranche d'élèves est ainsi soumise au test (âge 15-16 ans) et les mêmes items sont distribués dans tous les pays, sans prendre en compte les différences de programmes. Les items étant secrets et le cadrage théorique plutôt transversal au niveau 5, il est difficile de situer les données par rapport aux niveaux 2 à 4 de notre modèle. Comme dans TIMSS par ailleurs, des données sont recueillies concernant des conditions situées aux niveaux 6, 7 et 8. Les deux enquêtes ont donc, comme résultats principaux, des données au niveau 0, permettant une comparaison horizontale plus ou moins précise des performances d'élèves dans les pays participants, filtrées d'ailleurs par un traitement statistique avancé. La discussion des résultats prend ensuite en compte les données de base, souvent sous forme d'hypothèses fortes de causalité, les différences observées au niveau 0 étant expliquées par des différences corrélées situées aux niveaux supérieurs (2 à 8 pour TIMSS ; 6 à 8 pour PISA). Voici un exemple typique où l'on peut affirmer que des corrélations sont ainsi interprétées comme directement causales :

« L'organisation par l'établissement d'activités visant à renforcer l'engagement des élèves vis-à-vis des mathématiques (des concours de mathématiques, des clubs de mathématiques ou d'informatique mettant un accent spécifique sur les mathématiques) a également un impact positif, au-delà de l'effet de tous les autres facteurs. Chaque activité supplémentaire de ce type donne lieu à un gain de performance moyen d'un point de score. » (OCDE 2004, p.276).

Il est évident que ce type de raisonnement peut être critiqué, par exemple parce que la présence d'activités périscolaires est susceptible d'être corrélée à d'autres éléments de co-détermination, plus difficiles à modifier ou évaluer que le nombre d'activités en question. Il y a bien d'autres critiques faites et faisables par rapport aux cadrages, aux items, aux méthodes de recueil et de traitement de données, à l'usage peu problématisé de l'anglais comme langue de référence, etc. Ceci dit, la consistance de certains résultats des deux enquêtes, notamment les hautes performances de pays comme la Corée et le Japon, ont attiré l'attention sur des phénomènes qui étaient restés invisibles aux chercheurs de façon générale et au grand public.

2.2. Deux études qualitatives à grande échelle

Que l'on adopte un point de vue didactique ou un regard plus large, on ne peut s'empêcher de se sentir insatisfait de ce qui est offert par les données dites « de base », lorsque l'on cherche des explications plus opératoires et précises des différences manifestes et consistantes

repérées au niveau 0. On aimerait, en particulier, en savoir plus sur ce qui se passe réellement dans les classes de mathématiques, dans les pays concernés. Les méthodes de tests et de questionnaires ne donnent sur ce sujet qu'une information indirecte et partielle. Aussi, dès 1995 (au moment de la première enquête TIMSS), a-t-on mis en place dans trois pays (Etats-Unis, Allemagne, Japon) l'étude TIMSS vidéo, une étude complémentaire pour laquelle, dans chaque pays, 50 à 100 leçons de mathématiques ont été filmées, transcrites et codées par un groupe international de chercheurs (Kawanaka, Stigler & Hiebert, 1999). Le codage a permis, en particulier, de produire des données quantitatives, par exemple par rapport au temps consacré à des types d'activités définis par une catégorisation élaborée au cours de l'analyse initiale des vidéos. Des analyses qualitatives plus fines ont bien sûr été également menées mais ce sont surtout les résultats au niveau des structures de leçons dans un pays donné (niveau 7, co-détermination de niveaux 6 et 5) qui ont retenu l'attention des chercheurs et du grand public :

« ...il y a des propriétés typiques pour un grand nombre de leçons dans un même pays et qui séparent les leçons de pays différents. Ces propriétés ou modèles récurrents définissent les parties différentes d'une leçon et la manière dont elles sont ordonnées » (Stigler and Hiebert, 1999, p. 77f)

Ces structures ou, comme disent les auteurs cités, « scripts » de leçon, montrent en particulier la prédominance, au Japon, du travail autonome des élèves, à la différence des deux pays occidentaux où le travail des élèves est la plupart du temps restreint à l'entraînement à l'usage de techniques fournies par l'enseignant.

L'identification plus précise de « scripts » réguliers pour les leçons de mathématiques est certes controversée (par ex. par Keitel and Kilpatrick, 1999; Clarke, Mesiti, O'Keefe, Xu, Jablonka, Chee Mok & Shimizu, 2007). Cependant, on ne peut nier le rôle pionnier de l'étude vidéo TIMSS pour avancer vers une comparaison plus fine de ce que sont à l'évidence les lieux de production des connaissances des élèves, à savoir les séances de classe, et ainsi produire des hypothèses plus sérieuses et opératoires sur les causes possibles des différences identifiées concernant ces connaissances. Cette étude a d'ailleurs été répétée en 1999 dans 7 pays cette fois.

Il y a pourtant des raisons de douter que les différences majeures de pratiques d'enseignement correspondent aux frontières entre pays (voir par ex. (Clark et al., 2007)). Et l'on peut penser que, dans une perspective constructive, au sein d'une communauté internationale de chercheurs, il soit plus intéressant d'étudier en profondeur des exemples d'enseignement estimés « performants » dans des pays divers que de chercher à identifier des caractéristiques globales des pratiques suivant les pays, en faisant l'hypothèse que, dans un pays donné, les pratiques sont comparables et que l'on peut moyenniser les données recueillies. L'étude LPS sur les *learners' perspectives* (perspective des élèves) s'est en fait fixé ce but, en réunissant des équipes de chercheurs de 9 pays à l'équipe dirigeante située à Melbourne. Dans chaque pays, trois séquences de 10 leçons ont été filmées sous trois angles distincts (la classe entière, la perspective de l'enseignant et celle des élèves). Les trois séquences sont aussi choisies pour assurer une certaine « représentativité » du curriculum. Elles offrent de ce fait une diversité par rapport au niveau 4 même si chaque séquence est plutôt limitée à un secteur (niveau 3). Les vidéos sont réalisées dans des contextes urbains, avec des enseignants considérés « de

qualité » par rapport aux priorités des programmes nationaux. Les données vidéos sont complétées par des entretiens avec les enseignants et les élèves, une diversité de matériaux liés aux séquences et des productions d'élèves. L'objet principal de cette étude est une analyse approfondie des pratiques des enseignants et des élèves dans les séquences filmées. La co-détermination didactique relève donc, pour chaque séquence, de l'articulation verticale des niveaux 0 à 7 plutôt que de la comparaison directe entre pays. Celle-ci peut être entreprise seulement aux niveaux supérieurs (5 à 7), du fait de l'arbitraire relatif des domaines présents dans les données pour chaque pays. Toutefois, on constate que l'absence de projet de comparaison horizontale entre pays (et donc de moyenniser sur ceux-ci) semble faire disparaître l'impression d'homogénéité des structures de leçons pour un pays donné, tout en laissant ouverte la possibilité que les enseignants choisis pour leur qualité aient aussi des pratiques plus autonomes que l'enseignant moyen :

« Une conséquence inévitable de n'importe quel échantillon représentatif de leçons individuelles pour une nation est de moyenniser sur les éléments distincts de leçons dont la position dans la leçon reflète de façon directe et informative la position de la leçon dans une séquence de sujets. L'ensemble de données détaillées de LPS permet une analyse fine non seulement de la forme mais aussi de la fonction des événements d'une leçon » (Clarke et al. 2007, 292 ; notre traduction).

Cette analyse permet donc un aperçu assez fin de pratiques d'enseignement dans une variété de pays, y compris des pays d'Extrême-Orient. Il faut cependant noter qu'aucun pays francophone n'a participé aux deux études considérées dans cette partie. Pour ce qui est des enquêtes TIMSS et PISA, il nous semble que les chercheurs francophones en didactique ont été moins présents dans les équipes nationales, ce qui explique en partie l'absence de leur pays dans les études de suivi considérées ici, qui ont été entreprises à l'initiative de chercheurs autonomes plutôt que de gouvernements.

2.3. Comparaison « Est-Ouest », la 13^{ème} étude de la CIEM

Par contre, les chercheurs francophones ont bien entrepris des études comparatives, notamment à plus petite échelle (au sens par exemple du nombre de pays considérés) et dans le contexte de thèses de doctorat (voir la section 3). Nous nous arrêterons ici sur une étude regroupant des chercheurs venant de tous les continents du monde (dont la francophonie), organisée par la CIEM en 2002 dans le but d'examiner, sous une diversité d'angles, les pratiques d'enseignement mathématique dans deux contextes de civilisation (niveau 9) : l'Occident et l'Extrême-Orient. La motivation de cette comparaison vient en partie des enquêtes et études à grande échelle qui ont fait surgir, entre autres, l'hypothèse de différences manifestes à ce niveau. En particulier, on a constaté dans les grandes enquêtes une variation interne des résultats de pays multiculturels comme les Etats-Unis qui suggèrent que les élèves issus de familles immigrées de pays asiatiques sont aussi performants que ceux vivant en Asie (Clarke in (Leung et al. 2006), p.354f). Ceci contribue à suggérer que le contexte culturel (niveau 9) joue peut-être autant voire plus que le contexte national (niveau 8) :

« Notre hypothèse est que les divisions culturelles sont beaucoup plus significatives que les divisions politiques ou géographiques pour expliquer les différences de pratiques

enseignantes en mathématiques. L'Extrême-Orient et l'Occident sont donc, dans cette étude, des démarcations culturelles (...) en gros identifiées comme la tradition chinoise/confucéenne d'un côté, et la tradition gréco-latine-chrétienne de l'autre » (Leung et al. 2006, p.4; notre traduction).

La treizième étude de la CIEM a donc rassemblé une centaine de chercheurs représentant de façon équilibrée les deux traditions étudiées. Le volume issu de l'étude présente environ 40 travaux, allant de présentations de pratiques de pays individuels à des analyses des grandes enquêtes internationales (considérées en sec. III.1-2) menées en rapport au thème de l'étude. Il y a donc une diversité de niveaux considérés, mais avec un point commun : la référence au niveau 9 des civilisations.

On constate, en parcourant les chapitres du livre, que non seulement l'objet d'étude – l'enseignement mathématique – mais aussi les perspectives des chercheurs, sont fortement marqués par le contexte culturel, même si les chercheurs de l'Asie se servent plus souvent d'outils théoriques et méthodologiques d'origine occidentale que l'inverse. Il devient alors évident que nous ne disposons pas d'un point de vue supérieur ou neutre pour mener des comparaisons qui se situent au niveau 9, et que mesurer les pratiques issues d'une tradition civilisatrice dans l'optique d'une autre risque de mener involontairement à des positions colonisatrices peu susceptibles de faire avancer la recherche. Il nous semble donc que le lecteur occidental a beaucoup à gagner en se penchant sur les parties de l'étude qui ont adopté des modèles théoriques d'origine locale pour étudier le contexte culturel de l'Extrême-Orient. Un exemple important d'étude de ce type est celle de Hirabayashi (in Leung et al. 2006, pp.51-65). Ce chercheur est un des fondateurs de la recherche japonaise en didactique des mathématiques, plus précisément de l'école importante centrée à l'Université de Hiroshima. Dans son chapitre, il développe l'importance pour l'étude de l'enseignement mathématique au Japon, de deux notions appariées : *dô* (道, en gros 'chemin') et *jutsu* (術, 'techniques'), que l'on retrouve dans les écoles d'arts martiaux connues comme *ju-dô* et *ju-jutsu*. La seconde, d'origine militaire, a pour but de parfaire des compétences techniques, tandis que la première, plus philosophique, inclut une façon partagée de penser. De même, l'apprentissage des mathématiques peut se concevoir, étroitement, comme l'appropriation de techniques utiles, ou de façon plus holistique, comme l'initiation à une façon cohérente et partagée de penser et d'agir – un *iemoto* (家元, littéralement « une maison ou famille spirituelle », cf. aussi Emori & Winsløw in Leung et al., 2006, pp.553-566). Hirabayashi déplore, en citant des exemples, la tendance actuelle de l'enseignement japonais à se limiter au *jutsu*, tout en laissant entrevoir un fond traditionnel de le concevoir comme *dô*. L'analyse pourra être rapprochée de la différence entre un programme focalisé sur la maîtrise de sujets et de thèmes précis (niveau 1, 2) et un programme qui met en valeur la cohérence transversale de toute la discipline (niveau 5), conçu comme *iemoto*.

Il faut cependant noter que, tandis que plusieurs contributions à cette étude semblent confirmer l'existence d'affinités de pensée et d'approche par rapport à l'éducation mathématique en Chine et au Japon, il paraît très difficile d'identifier des traits stables et partagés pour les pays dits « occidentaux ». Par exemple, Kaiser, Hino and Knipping (in Leung et al., 2006, pp.318-350) ont étudié l'enseignement mathématique en Angleterre,

France, Allemagne et Japon sous l'optique de huit variables empiriques plutôt globales (niveau 5 et 6), par exemple la priorité accordée dans les programmes de mathématiques aux applications dans des contextes extra-mathématiques. Ils trouvent que, pour ces huit variables, l'Angleterre et la France correspondent aux extrémités du spectre, « l'Allemagne et le Japon étant situés quelque part entre ces deux extrêmes ».

Le cas d'autres pays d'Asie, comme le Vietnam décrit par Bessot et Comiti (in (Leung et al., 2006), pp. 159-179 – cf. aussi Sec. 3), montre comment l'histoire des interactions avec les pays occidentaux, notamment avant, durant et après l'époque coloniale, fait apparaître une succession d'idéologies et de philosophies de l'éducation parfois contradictoires (niveaux 6 à 8). L'analyse de la co-détermination des pratiques didactiques et mathématiques actuelles, à un niveau plus local (niveaux 1 à 4) est donc loin de s'articuler simplement à des différences de civilisations orientale et occidentale.

3. Un corpus de thèses comparatives

Après avoir considéré des études comparatives impliquant de nombreux pays, mais souvent peu de pays francophones, nous nous intéressons dans ce qui suit à un corpus très différent. Il s'agit en effet de comparaisons binaires menées dans le cadre de doctorats préparés en France. Ce sont donc l'œuvre d'un chercheur et non d'un collectif, et ces œuvres ont été développées au sein d'une culture didactique particulière où la TAD que nous utilisons ici comme outil de méta-analyse est une théorie influente. Pour préparer cette séance du projet spécial, nous n'avons pas recherché l'exhaustivité mais nous avons sélectionné un corpus de onze thèses. Il s'agit d'une part des quatre thèses comparatives qui ont été préparées au sein de l'équipe DIDIREM (devenue depuis Laboratoire de didactique André Revuz : LDAR) à l'université Paris 7, d'autre part de sept thèses préparées à l'université de Grenoble 1 dans le cadre d'une coopération de longue durée avec l'université pédagogique d'Ho-Chi-Minh au Vietnam. Le tableau ci-après présente ce corpus.

Auteur	Année	Pays	Thème
Nadia Amra	2004	France-Palestine	L'enseignement des fonctions au lycée
Valentina Celi	2002	France-Italie	L'enseignement de la géométrie dans le secondaire (11-16 ans) et ses effets
Richard Cabassut	2005	Allemagne-France	Raisonnement et preuve dans l'enseignement des mathématiques
Pascal Stoelting *	2008	Allemagne-France	Pensée fonctionnelle chez des élèves de 10 à 16 ans. Son enseignement en Allemagne et en France
Lê Thi Hoai Châu*	1997	France-Vietnam	L'enseignement des vecteurs en classe de seconde
Doan Huu Hai*	2001	France-Vietnam	L'enseignement de la géométrie dans l'espace au lycée
Lê Van Tien *	2001	France-Vietnam	Les relations entre fonctions et équations au lycée
Nguyen Chi Thanh*	2005	France-Vietnam	L'introduction de l'algorithmique et de la programmation dans l'enseignement secondaire

			avec des calculatrices
Nguyen Hai Quoc*	2006	France-Vietnam	La résolution des équations du second degré dans l'enseignement secondaire
Tran Luong Cong Khanh*	2006	France-Vietnam	La notion d'intégrale au lycée
Lê Thai Bao Thien Trung*	2007	France-Vietnam	Les relations entre l'enseignement de la notion de limite et la "décimalisation" des nombres réels dans l'environnement de calculatrices

Tableau 1 : Corpus des thèses étudiées (les thèses marquées d'un astérisque sont des thèses en co-tutelle)

En quoi l'inscription dans un travail de thèse oriente-t-elle le travail comparatif et influe-t-elle sur la nature des résultats obtenus ? L'outil de comparaison qui a été développé est-il pertinent pour rendre compte de ces travaux et permet-il de mettre en lumière des régularités intéressantes ? C'est à ces questions que nous essayons de répondre dans cette partie.

3.1. Les thèses Franco-Vietnamiennes

Pour situer ce premier corpus, nous nous appuyons au-delà des seules thèses sur l'article de Comiti et Bessot (2009) qui retrace l'histoire de ce projet de coopération entre les deux universités. Cet article met bien évidence en effet la cohérence globale du projet dans lequel s'inscrivent ces différents travaux. Ses objectifs sont triples :

1. réinvestir et mettre à l'épreuve des connaissances didactiques élaborées dans le contexte culturel français
2. répondre aux besoins du système éducatif vietnamien
3. favoriser l'intégration des nouveaux docteurs dans la communauté scientifique de leur pays.

Ce qui est aussi souligné, c'est que ces études sont perçues comme un moyen privilégié de questionner la tendance naturelle à la naturalisation des praxéologies mathématiques et didactiques, qui conduit chacun à considérer son contexte particulier comme le contexte normal. L'influence de la TAD sur les travaux menés est par ailleurs évidente, comme le montre la citation suivante :

« Loin d'une approche 'cognitiviste' centrée sur les particularités des élèves et des professeurs, ce type de recherche situe donc le problème des difficultés des élèves et des professeurs dans l'ensemble des conditions et des contraintes qui pèsent sur l'enseignement des mathématiques ».

Il s'ensuit que, pour ces recherches, dont l'objet déclaré se situe aux niveaux 3 ou 4 de l'échelle de co-détermination (cf. les thèmes précisés dans le tableau 1), l'analyse institutionnelle basée sur l'étude de documents curriculaires et considérée à la fois dans ses dimensions diachronique et synchronique, joue un rôle fondamental. Et elle met en lumière les contraintes de différents niveaux pesant sur l'enseignement au Vietnam, utilisant pour cela de façon privilégiée les moments de réforme, et pointant par exemple l'influence conjuguée d'une tradition confucéenne millénaire et de la colonisation française (niveau 8). En utilisant la terminologie introduite dans la partie 2, l'accent est donc mis d'abord sur le *curriculum visé*, mais des connections sont établies avec les *curricula effectif et atteint*. Des

questionnaires sont régulièrement utilisés pour évaluer l'impact des différences praxéologiques identifiées sur les pratiques des étudiants. La comparaison institutionnelle peut également mener à la conception et à l'expérimentation d'ingénieries didactiques, en particulier dans le contexte vietnamien. C'est le cas par exemple des deux thèses de Chi Thanh N'guyen et Lê Thai Bao Thien Trung moins centrées sur la comparaison que les autres thèses. Dans ce cas, comme le soulignent Comiti et Bessot :

« La réalisation d'une ingénierie didactique [adossée à une telle comparaison institutionnelle] dans l'un des systèmes et son analyse permettent non seulement de poser des questions sur ce système mais aussi sur l'autre système, et favorisent ainsi l'émergence de questions génériques dont la pertinence est attestée dans l'autre système. »

Mais ces parties des thèses qui se situent au-delà de l'analyse fine des processus de transposition didactique restent des études qualitatives et locales. De plus, on ne trouve dans ces travaux aucune étude directe du curriculum implémenté à travers l'observation et l'analyse de pratiques de classes analogue à celle observée dans certaines études comparatives de la partie III.

3.2. Les thèses comparatives de l'équipe DIDIREM

Les thèses préparées au sein de l'équipe DIDIREM constituent un corpus moins homogène mais on observe cependant certaines similarités intéressantes avec le corpus précédent. La problématique de la comparaison se situe également aux niveaux 3 et 4 et elle est soutenue par des analyses de type épistémologico-mathématiques approfondies ; la TAD, même si elle est complétée par d'autres constructions théoriques relatives au domaine mathématique considéré ou par des cadres comme celui des registres sémiotiques de Duval, est le cadre théorique dominant dans trois des quatre thèses et l'analyse institutionnelle faisant intervenir différents niveaux de co-détermination, même si la notion elle-même n'est pas utilisée, joue un rôle fondamental. Par exemple Amra pointe les diverses influences coloniales anglo-saxonnes qui ont façonné les programmes mathématiques en Cisjordanie et plus généralement les visions pédagogiques pour expliquer les différences praxéologiques importantes avec la France dans le domaine des fonctions. De même Celi fait intervenir les caractéristiques de l'organisation curriculaire en Italie comme la plus grande autonomie accordée aux enseignants, le rôle différent attribué aux manuels, pour expliquer les différences qu'elle observe dans l'enseignement de la géométrie. On note des études d'impact au niveau 0 mais, comme dans le corpus précédent, l'impasse est faite sur l'étude du curriculum implémenté. Enfin, comme dans le premier corpus, les questions linguistiques sont faiblement problématisées, si l'on excepte la thèse de Cabassut qui étudie de façon systématique le langage didactique associé aux processus de raisonnement et de preuve dans les deux pays concernés par son étude : l'Allemagne et la France.

Au-delà de ces similarités, des différences existent, au niveau des motivations de l'étude qui apparaissent plus variées que dans le premier corpus, au niveau des importances relatives accordées à l'étude institutionnelle et à l'étude du curriculum atteint (niveau 0), ainsi qu'au niveau des moyens mis en œuvre pour cette étude au niveau 0 : questionnaires, entretiens, utilisation de données provenant d'autres sources comme PISA, d'études nationales ou régionales, de compétitions mathématiques entre régions. Et il nous semble intéressant de

souligner que c'est à ce niveau 0 que les problèmes de comparabilité entre systèmes émergent, quand il s'agit de construire des instruments adaptés aux deux contextes, comme le montrent particulièrement bien les thèses d'Amra et de Céli.

Enfin, on observe une plus grande symétrie des contextes comparatifs que dans le premier cas, ce qui en retour pose des problèmes de mise en compatibilité de cadres théoriques. Ceci est en particulier visible dans la thèse de Stölting où sont coordonnées les notions de champ conceptuel due à Vergnaud (1990), de registre sémiotique déjà citée et de Grundvorstellungen (idées fondamentales) due à vom Hofe (1995).

4. Conclusions et perspectives

La méta-analyse qui a été menée nous semble bien montrer la diversité et la richesse des recherches comparatives concernées, l'importance aussi de continuer à développer ce type de recherches dans le contexte actuel de médiatisation et d'influence politique croissante des évaluations internationales à grande échelle comme TIMSS et PISA. Faute de pouvoir s'appuyer sur les résultats de telles recherches, le risque est grand de comparer l'incomparable, de s'attacher à des traits de surface, de faire des inférences abusives. Mais elle montre aussi les difficiles problèmes méthodologiques que posent ces recherches, et la vigilance dont les chercheurs doivent faire preuve pour résister à la tentation de faire parler les données au-delà de ce qu'elles permettent, que ce soit en opérant des regroupements et des moyennisations abusives, en extrapolant des résultats obtenus dans des contextes très spécifiques, ou en interprétant abusivement en termes de causalité des relations observées entre niveaux de co-détermination. Elle montre aussi la vigilance à exercer par rapport au pouvoir exercé par les cultures dominantes dont les modèles culturels ne sont en rien des références neutres et, non indépendamment, la nécessité de porter une attention plus grande aux questions linguistiques.

S'il importe de développer ces recherches, il est aussi important de mieux les coordonner, et que la communauté sache mieux tirer parti des énormes bases de données qui en résultent et ne sont que très partiellement exploitées. Mais pour soutenir cette coordination, faciliter l'exploitation des données et la capitalisation des résultats, il est sans aucun doute nécessaire de continuer à développer des outils de comparaison et de méta-analyse comme celui que nous avons élaboré pour cette contribution. En nous lançant dans ce travail nous avons fait l'hypothèse que la TAD offrait, pour répondre à ces besoins, des caractéristiques très intéressantes, et cette conviction s'est renforcée au cours de l'avancée de ce travail.

Les pays francophones ont été jusqu'ici assez peu partie prenante dans les grandes études comparatives internationales. La partie III de cette contribution le montre nettement. La question linguistique n'y est cependant pas étrangère. Pourtant, l'analyse du corpus des thèses le montre, les recherches que les chercheurs francophones mènent dans ce domaine présentent des caractéristiques intéressantes et montrent que les chercheurs concernés ont réussi à élaborer, notamment en s'appuyant sur la TAD, des cadres théoriques qui permettent de problématiser de façon tout à fait pertinente les comparaisons et fournissent des notions opérationnelles pour les mener à bien. Il nous semble important que ces travaux ne demeurent

pas isolés de ce courant international de recherche, et nous espérons que notre travail pourra y contribuer.

Références :

- ARTIGUE, M., WINSLOW, C. (2010) International comparative studies on mathematics education : a viewpoint from the anthropological theory of didactics. *Recherche en Didactique des Mathématiques*, 30/1, 47-82..
- BESSOT, A., COMITI, C. (2009) Apport des études comparatives aux recherches en didactique des mathématiques : le cas Vietnam / France. In Hache C., Coulange L. (Eds.) *Actes du Séminaire National de Didactique des Mathématiques (ARDM)*.
- CHEVALLARD, Y. (1985) *La transposition didactique. Du savoir savant au savoir enseigné*. Grenoble : La Pensée Sauvage.
- CHEVALLARD, Y. (1999) L'analyse des pratiques enseignantes en théorie anthropologique du didactique. *Recherches en Didactique des Mathématiques* 19(2), 221-265.
- CHEVALLARD, Y. (2002). Organiser l'étude 3. Écologie & régulation. In Dorier, J. L. et al. (Eds.), *Actes de la 11e école de didactique des mathématiques*. Grenoble : La Pensée Sauvage.
- CLARKE, D. (2003). International comparative research in mathematics education. In Bishop, A. et al., *Second international handbook of mathematics education*, pp. 143-184. Dordrecht: Kluwer.
- CLARKE, D., MESITI, C., O'KEEFE, C., XU, L., JABLONKA, E., CHEE MOK, I., AND SHIMIZU, Y. (2007). Addressing the challenge of legitimate international comparisons of classroom practice. *International Journal of Educational Research* 46, 280–293.
- CLARKE, D. (2006). The LPS research design. In D. Clarke, C. Keitel & Y. Shimizu (Eds.) *Mathematics in twelve countries: the insider's perspective*, pp. 15-36. Rotterdam: Sense Publishers.
- DUVAL R. (1995). *Semiosis et Pensée Humaine*. Paris : Peter Lang.
- HOFE R. VOM (1995). *Grundvorstellungen mathematischer Inhalte*. Heidelberg: Spektrum Akademischer Verlag GmbH.
- HOPMANN, S., BRINEK, G., RETZL, M. (Eds.) (2007). *PISA according to PISA – does PISA keep what it promises?* Wien: LIT Verlag.
- Houdement C., Kuzniak A. (2000). Formation des maîtres et paradigmes géométriques. *Recherches en Didactique des Mathématiques* 89-116.
- KAWANAKA, T., STIGLER, J., HIEBERT, J. (1999). Studying mathematics classrooms in Germany, Japan and the United States: lessons from the TIMSS videotape study. In G. Kaiser, E. Luna and I. Huntley (Eds.), *International comparison in mathematics education*, pp. 86-103. London: Falmer Press.
- KEITEL, C., KILPATRICK, J. (1999). The rationality and irrationality of international comparative studies. In G. Kaiser, E. Luna and I. Huntley (Eds), *International comparison in mathematics education*, pp. 241-256. London: Falmer Press.
- LANGE, J. DE (1996). Using and Applying Mathematics in Education. in: A.J. Bishop, A. et al. (Eds). 1996. *International handbook of mathematics education, Part one*, pp. 49-97. Dordrecht: Kluwer.
- F. LEUNG, K.-D. GRAF, F. LOPEZ-REAL (2006) (Eds). *Mathematics education in different cultural traditions – a comparative study of East Asia and the West*. New York: Springer Science.

- MULLIS, I., MARTIN, M., RUDDOCK, G., O'SULLIVAN, C., ARORA, A., ERBERBER, E. (Eds.) (2005). *TIMSS 2007 assessment frameworks*. Boston: TIMSS & PIRLS international study center.
- MULLIS, I., MARTIN, M., GONZALES, E., CHROSTOWSKI, S. (2008). *TIMSS 2007 International Mathematics Report: Findings from IEA's Trends in International Mathematics and Science Study at the Fourth and Eighth Grades*. Boston: TIMSS & PIRLS international study center.
- NISS, M. (2003). Quantitative literacy and mathematical competencies. In B. L. Madison and L. A. Steen (Eds.), *Quantitative literacy : why numeracy matters for schools and colleges*, pp. 215-220. Princeton: National Council on Education and the Disciplines.
- OCDE (2004), Apprendre aujourd'hui, réussir demain - Premiers résultats de PISA 2003. Paris: OCDE.
- OCDE (2006), Compétences en sciences, lecture et mathématiques : Le cadre d'évaluation de PISA 2006. Paris: OCDE.
- TRAVERS, K. (1988). Introduction to the second international mathematics study. In Robitaille, D. and Garden, F. (Eds.), *The IEA Study of Mathematics II: Contexts and Outcomes of School Mathematics*, pp. 1-16. London: Pergamon Press.
- VERGNAUD, G. (1990). La théorie des champs conceptuels. *Recherches en Didactique des Mathématiques*, 10(2-3), 133-170.
- WINSLØW, C. (2005). Définir les objectifs de l'enseignement mathématique : La dialectique matières-compétences. *Annales de didactique et des sciences cognitives* 10, 131-156.

Bilan du projet Spécial 2

Le projet spécial a rassemblé une trentaine de participants. C'est un effectif relativement modeste qui traduit sans doute l'intérêt jusqu'ici limité de la communauté didactique francophone sur ces questions. Mais ce projet spécial, nous semble-t-il, a joué le rôle qui était attendu de lui. Il a permis aux participants de mieux comprendre l'organisation et le fonctionnement de ces évaluations internationales qui sont des objets complexes, les valeurs et les conceptions des mathématiques et de leur apprentissage qui les sous-tendent, la nature réelle des résultats qu'elles permettent d'obtenir au-delà des présentations simplistes qui en sont généralement faites dans les médias. Il a permis de rassembler des informations très intéressantes sur la façon dont ces évaluations internationales influencent les politiques éducatives et les choix curriculaires dans les pays francophones notamment, de prendre la mesure de la diversité de ces influences suivant les contextes, de leurs caractères producteurs mais aussi de leur possible dangerosité. Il a montré que ces évaluations internationales sont source de multiples questions de recherches qui nécessitent des conceptualisations spécifiques ou au moins des adaptations des conceptualisations existantes, ainsi que des méthodologies adaptées. Il a permis de mieux situer les travaux de la communauté francophone dans le champ plus global des études comparatives internationales, qu'elles soient ou non motivées par des évaluations internationales, et de questionner le potentiel de certaines des approches théoriques développées au sein de cette communauté, comme la théorie anthropologique du didactique, pour ce champ d'étude.

Sur un thème dont la sensibilité politique est évidente, les échanges ont été vifs, parfois passionnés, mais ils ont été guidés par la volonté de soumettre les questions à l'étude à une réelle réflexion scientifique. Il est cependant clair que, malgré la richesse des contributions et des interventions, le thème est loin d'être épuisé par le travail qui a pu être fait à Dakar. Le nombre très réduit des travaux secondaires existants sur les évaluations internationales au sein de la communauté francophone rend l'identification de phénomènes et effets plus spécifiques qui pourraient affecter l'enseignement des pays de cette communauté encore difficile, même si l'on voit s'exprimer le ressenti d'une domination de perceptions d'origine anglo-saxonne de ce que devraient être les résultats de l'enseignement mathématique.

Des recherches relevant du champ des comparaisons internationales existent bien dans les pays francophones mais elles ont surtout visé jusqu'ici des comparaisons à petite échelle, impliquant souvent deux pays, un contexte institutionnel bien précis et une partie spécifique du curriculum. Les projets internationaux plus vastes ont été le plus souvent initiés par des chercheurs d'origine anglo-saxonne et les chercheurs francophones y sont très peu présents.

C'est dans ce champ plus vaste des études comparatives internationales qu'il nous semble qu'il serait intéressant de poursuivre le travail, en se centrant sur *le sens et l'usage de comparaisons des pratiques de l'enseignement mathématique* à travers les pays francophones et au-delà, les évaluations internationales n'étant plus sur le devant de la scène, en faisant appel à une variété d'expériences, qu'elles émanent de travaux de recherche ou de pratiques plus larges.

Auteurs :

Michèle Artigue, LDAR, Université Paris Diderot - Paris 7, France, artigue@math.jussieu.fr

Antoine Bodin, IREM de Marseille, France, antoinebodin@mac.com

Richard Cabassut, IUFM Université de Strasbourg & LDAR, Université Paris Diderot - Paris7, France, richard.cabassut@unistra.fr

Lucie DeBlois, Université Laval, Québec, Lucie.DeBlois@fse.ulaval.ca

Nadia Mawfik, Ecole Normale Supérieure de Rabat, Maroc, nmawfik@yahoo.fr

Charles Coagri NASSOURI, DGIFPE Ouagadougou, Burkina Faso, Charles_nasouri@univ-ouaga.bf

Hikma Smida, Université Tunis El Manar, Tunisie, hikma.smida@minedu.edunet.tn

Carl Winsløw, IND, Université de Copenhague, Danemark, winslow@ind.ku.dk