

Conf 04 : Les évaluations externes des élèves en mathématiques : apports, enjeux et perspectives*¹

Jean-François Chesné

Directeur scientifique du Conseil national d'évaluation du système scolaire (Cnesco); jean-francois.chesne@education.gouv.fr

Résumé : Les évaluations nationales et internationales révèlent depuis une trentaine d'années une proportion préoccupante pour la France d'élèves en difficulté en fin d'école primaire ou de collège. Les récents résultats de l'enquête internationale TIMSS menée en fin de CM1 en 2015 auprès des élèves de CM1 ont confirmé cette situation.

Après une présentation générale de ces évaluations, de leurs objectifs et de leurs résultats, ce texte se centre plus spécifiquement sur le domaine des nombres et du calcul, en prenant appui sur les travaux du Conseil national d'évaluation du système éducatif (Cnesco) et sur des résultats de recherche en didactiques des mathématiques. À partir du dispositif PACEM (Chesné, 2014), il s'efforce de montrer en quoi des évaluations externes des acquis des élèves, en étant ciblées et adaptées à des projets locaux, peuvent constituer des outils de formation professionnelle des enseignants.

Mots clefs : évaluation, nombres et calcul, développement professionnel, Cnesco.

Introduction

Le débat sur le niveau des élèves en mathématiques se pose régulièrement en France et concerne désormais, à des degrés divers et pour des raisons différentes, un public très large : professionnels de l'enseignement et de la formation des enseignants, chercheurs, décideurs politiques, parents, ... Les évaluations nationales et internationales alimentent de plus en plus ce débat, en révélant une proportion préoccupante d'élèves en difficulté en mathématiques, en fin d'école primaire et en fin de collège. Les récents résultats de l'enquête internationale TIMSS menée en fin de CM1 en 2015 auprès des élèves de CM1 ont confirmé cette situation. Après une présentation rapide du Conseil national d'évaluation du système éducatif (Cnesco), nous dresserons un panorama de ces évaluations, de leurs objectifs et de leurs résultats, puis nous centrerons notre intervention plus spécifiquement sur le domaine des nombres et du calcul, en prenant appui sur les travaux d'évaluation du Cnesco et sur des résultats de recherche en didactique des mathématiques. Nous nous efforcerons enfin, en nous appuyant notamment sur les résultats du dispositif PACEM² (Chesné, 2014), de proposer de nouvelles voies pour l'usage des évaluations externes, et en particulier, de montrer en quoi des évaluations externes des acquis des élèves, en étant ciblées et adaptées à des projets locaux peuvent constituer des outils de formation des enseignants.

Le Conseil national d'évaluation du système scolaire (Cnesco)

Créé par la Loi de la refondation de l'école de 2013, le Conseil national d'évaluation du système scolaire (Cnesco) a été installé en janvier 2014 afin d'assurer une évaluation scientifique et indépendante du système scolaire, complémentaire aux travaux d'évaluation internes au Ministère de l'éducation nationale. Il est composé majoritairement de chercheurs issus de plusieurs champs

¹ Ce texte est celui d'une présentation de juin 2017, antérieur à une reprise d'évaluations exhaustives menées à la rentrée 2017 par le ministère de l'Éducation nationale auprès des élèves de CP et de sixième.

² Projet pour l'acquisition de compétences par les élèves en mathématiques.

scientifiques, spécialistes de l'évaluation, mais aussi de parlementaires et de membres du Conseil économique, social et environnemental. L'article L. 241-12 de la loi du 8 juillet 2013 et la lettre de mission qui le complète, définissent les trois principales missions du Cnesco :

- évaluer le fonctionnement du système scolaire et ses résultats ;
- diffuser les résultats des évaluations des élèves, des dispositifs et des politiques scolaires ;
- évaluer les méthodologies mises en œuvre par les évaluateurs internes au ministère de l'Éducation nationale et celles mises en œuvre par les organismes internationaux.

En trois années, en s'appuyant sur un réseau de plus de 200 chercheurs et experts internationaux et en associant largement la communauté éducative, le Cnesco a ainsi investigué un champ très large de thématiques cruciales pour l'école française, depuis les pratiques pédagogiques dans la classe et l'établissement (la lecture, les mathématiques, le redoublement, le traitement de la difficulté scolaire) jusqu'aux politiques scolaires (le handicap, l'enseignement professionnel, l'éducation à la citoyenneté, les inégalités sociales et migratoires, la mixité). Afin de diffuser efficacement les résultats de la recherche et d'enrichir le débat sur l'éducation, le Cnesco a développé six formats d'activités, avec la collaboration de plusieurs partenaires (l'Ifé/ENS de Lyon, le CIEP, Réseau Canopé, l'ESENESR, le réseau des Espé...) :

- **des conférences de consensus** : à partir de ressources d'évaluations scientifiques produites par le Cnesco, un jury d'acteurs de la communauté éducative est chargé d'auditionner des experts de disciplines variées pour produire des recommandations ;
- **des conférences de comparaisons internationales** : à partir de ressources d'évaluations scientifiques sur les politiques scolaires produites par le Cnesco, un public de décideurs français et internationaux sont réunis et échangent autour des politiques publiques afin de proposer des préconisations ;
- **des conférences virtuelles** : des échanges en direct et à distance sont ouverts entre des experts d'une thématique et des acteurs de terrain sur les évaluations du Cnesco ;
- **des rapports scientifiques** : des évaluations quantitatives et qualitatives pluridisciplinaires sont développées dans la durée, pour évaluer l'état de l'école ;
- **des notes d'actualité, et des forums en région.**

Le Cnesco met à disposition de tous des dossiers de ressources complets et multimédias (rapports scientifiques, dossiers de synthèse, préconisations, vidéos...) sur son site Internet : www.cnesco.fr.

De quoi parle-t-on quand on parle d'évaluations externes des acquis des élèves ?

Que sont les évaluations externes des acquis des élèves et quels rôles peuvent-elles jouer dans notre système éducatif ? La question mérite d'être posée. En effet, « aux côtés des évaluations internes des élèves – c'est-à-dire qui relèvent de la responsabilité exclusive de l'enseignant dans sa classe et son établissement – s'est développée depuis les années 1990, dans les pays de l'OCDE une panoplie d'évaluations externes – c'est-à-dire des évaluations qui relèvent d'acteurs extérieurs à l'établissement scolaire – par les administrations centrales des ministères de l'éducation, des agences indépendantes en charge de l'évaluation des élèves ou encore les collectivités territoriales, dans certains pays très décentralisés » (Cnesco, 2014). Hormis les examens certificatifs, comme le baccalauréat, ou les épreuves finales du diplôme national du brevet, les évaluations externes des acquis des élèves font le plus souvent référence à des évaluations qui ne sont pas d'ailleurs

certificatives en France, les évaluations standardisées des acquis des élèves. La conception, l'administration et la correction des tests ainsi que la publication des résultats répondent à un cadre de référence normé (Mons & Pons, 2006). Elles se distinguent en cela d'autres évaluations externes, plus locales, qui peuvent s'en rapprocher, ou s'en inspirer, comme les évaluations pilotées par les académies ou les départements, ou encore les évaluations de l'APMEP³.

Traditionnellement, on oppose deux finalités principales de ces évaluations standardisées : diagnostic et bilan. Plus précisément, Trosseille et Rocher (2015) leur assignent trois objectifs différents :

- *« fournir aux enseignants des outils afin d'enrichir leurs pratiques pédagogiques en évaluant mieux les acquis de leurs élèves ;*
- *disposer d'indicateurs permettant de mesurer, au niveau national, les performances de notre système (évolutions temporelles et comparaisons internationales) ;*
- *doter les « pilotes de proximité » (recteurs, DASEN, IEN) d'indicateurs leur permettant de mieux connaître les résultats des écoles et d'effectuer une vraie régulation. »*

Selon son objectif, une évaluation standardisée peut être exhaustive (elle concerne tous les élèves d'un même niveau scolaire) ou réalisée sur échantillon, ce qui a évidemment une influence sur son coût et son organisation.

Quoi qu'il en soit, les psychométriciens considèrent qu'une confusion des objectifs est dangereux méthodologiquement, et qu'il y a une distinction nécessaire à effectuer et à afficher (rapport IGEN, 2005, rapport du HCE, 2011). Il est couramment admis que cette confusion a existé, au début des années 2000 et entre 2009 et 2012, ce qui a conduit en grande partie à l'abandon actuel des évaluations de type 1 et de type 3.

Mais cette prudence statistique n'est pas la seule à souligner à propos des évaluations externes. Il faut également y associer une réserve des enseignants - qui peut aller de l'indifférence à une opposition forte – une réticence des formateurs (sur l'utilisation des évaluations standardisées et de leurs résultats en formation), et une certaine prudence des chercheurs sur la validité de ces évaluations et leurs résultats. Ces réserves, voire ces défiances, sont d'autant plus fortes que :

- le flou/la confusion sur les enjeux, les objectifs et les usages sont importants ;
- la transparence de la part des opérateurs sur les contenus des tests est faible ;
- la montée en puissance des évaluations de pilotage national fait peser une pression médiatique de plus en plus grande sur l'école ;
- des retours négatifs de l'étranger (dérives du concept de « *hard accountability* »; Mons, 2007) font craindre le pire aux établissements et aux enseignants ;
- le travail d'information et de formation des enseignants... et des formateurs est insuffisant.

Quelles évaluations nationales et internationales et que nous apprennent-elles ?

Toutes les évaluations nationales sont menées par le ministère de l'éducation nationale, et le plus souvent par la direction de l'évaluation, de la prospective et la performance (Depp). Elles concernent essentiellement la scolarité obligatoire, et sont de plusieurs natures. Elles peuvent être :

³ Association des professeurs de mathématiques de l'enseignement public.

- diagnostiques, comme les évaluations menées de 1989 à 2008 à l'entrée en CE2 et l'entrée en sixième, puis les évaluations menées en CE1 et en CM2 par la direction générale de l'enseignement scolaire (Dgesco) entre 2009 et 2012 ;
- conçues pour fournir des indicateurs pour le Parlement dans le cadre de la LOLF (loi organique relative aux lois de finances) menées en CM2 et en troisième ;
- bilans, en fin d'école et en fin de collège comme celles du programme Cedre (Cycle des évaluations disciplinaires réalisées sur échantillons) mené pour la première fois en 2008 pour les mathématiques, puis en 2014 ;
- à objectif de comparaisons temporelles comme l'étude menée en 2007 (permettant d'apprécier des variations sur plusieurs années d'intervalles de connaissances analogues (calcul posé et résolution de problèmes arithmétiques)
- destinées à identifier les jeunes en situation d'illettrisme (et en 2013 en situation d'innumérisme), comme les tests de la Journée défense et citoyenneté, adressés aux jeunes de 17 ans. et e, et

Au niveau international, le Programme international de suivi des acquis des élèves (PISA) mis en place pour la première fois en 2000 par l'OCDE s'est aujourd'hui imposé. Les enquêtes PISA concernent l'ensemble des élèves dont l'âge est compris entre 15 ans et trois mois et 16 ans et 2 mois, quels que soient le niveau scolaire et la filière qu'ils occupent dans le système éducatif. L'OCDE part de l'idée que le développement des économies des pays dépend largement de la qualité de leurs systèmes éducatifs et que le citoyen doit savoir mobiliser ses connaissances scolaires pour résoudre les problèmes qu'ils sont susceptibles de rencontrer dans « la vie réelle ». PISA n'évalue donc pas strictement des champs disciplinaires, mais une littératie mathématique, scientifique et en compréhension de l'écrit (Cnesco, 2016).

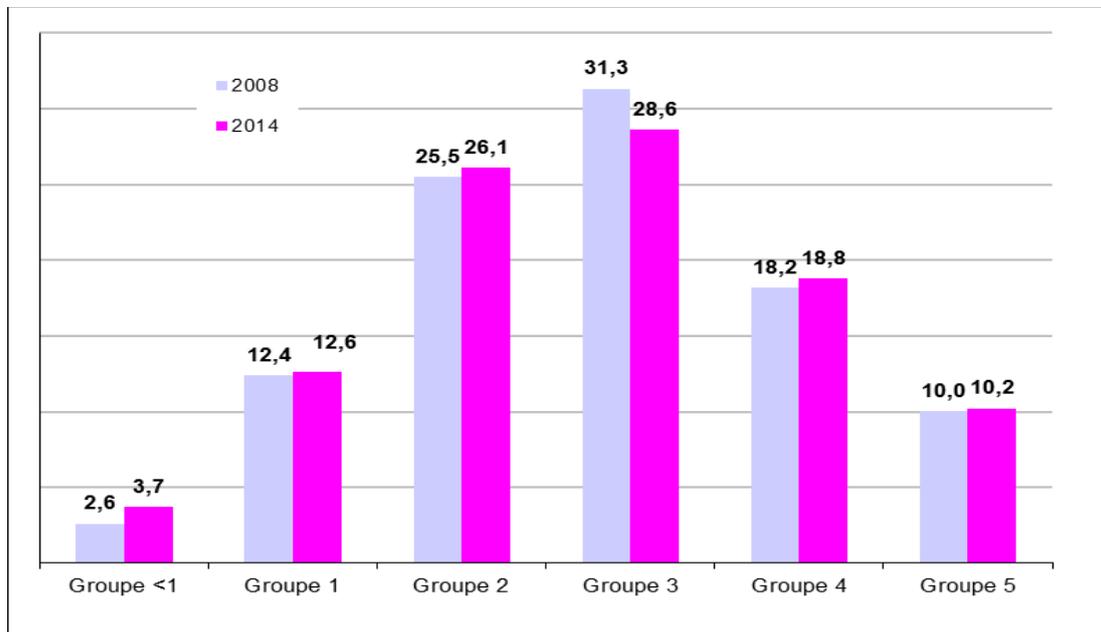
Un autre programme d'enquêtes internationales, TIMSS (*Trends in international mathematics and science study*), est mené depuis 1995. Les enquêtes TIMSS concernent les élèves qui sont dans leur quatrième ou huitième année de scolarité (en France, élèves de CM1 et de quatrième). En 2015, la France a participé pour la première fois au programme TIMSS pour le CM1, aux côtés de 47 autres pays. La France a également participé en 2015 à TIMSS *Advanced* après avoir participé à la première édition en 1995, et après 20 ans de non-participation. Cette enquête a pour population cible les élèves qui sont en classes scientifiques de fin d'études secondaires, la terminale S en France. Les enquêtes TIMSS et TIMSS *Advanced* cherchent à mieux connaître les systèmes éducatifs en ce qui concerne l'enseignement des mathématiques et des sciences et moins directement, à améliorer les politiques scolaires. Elles s'intéressent avant tout aux contenus d'enseignement présents dans les programmes scolaires, et aux acquis des élèves relativement à ces contenus.

Les principaux enseignements de l'ensemble de ces évaluations, convergents et désormais partagés au sein de l'institution, sont les suivants :

- Environ 40 % des élèves français n'ont pas le niveau attendu à la sortie de l'école primaire (Figure 1) ;
- Cette proportion a augmenté entre 2008 et 2014 en fin de collège (Figure 2) ;

- Le nombre d'élèves français « de bas niveau » (obtenant de scores peu élevés) augmente (Figures 2 et 3).

Figure1 : Répartition des élèves par niveaux de compétence aux évaluations Cedre fin d'école en 2008 et en 2014



Source : Depp-MENESR. Note d'information N°18, DEPP, Mai 2015

Les élèves des 3 groupes « de bas niveaux » (groupes inférieurs à 1, 1 et 2) ne maîtrisent pas complètement les nombres entiers, très peu les nombres décimaux, et n'ont une maîtrise assurée que de l'addition et de la soustraction des nombres entiers (avec ou sans retenue). La proportionnalité n'est pas acquise.

TIMSS 2015 (CM1) a confirmé ces résultats avec un résultat supplémentaire très inquiétant : la France obtient le score le plus faible en mathématiques des 26 pays de l'OCDE participant à l'enquête, juste devant le Chili. Plus précisément :

- La proportion d'élèves « de bas niveaux » est supérieure en France (42 %) à celle des pays européens participants (24 % en moyenne) ;
- La proportion « de niveau élevé » est inférieure en France (2 %) à celle des pays européens participants (9 %) ;

Les exemples ci-dessous, issus de TIMSS 2015 (CM1), illustrent les résultats des élèves français dans le domaine des nombres et du calcul⁴ :

Exemple 1 :

⁴ <https://timssandpirls.bc.edu/timss2015/international-database/index.html>

Yassine commence à écrire une suite logique de nombres :

6, 13, 20, 27, ...

Il ajoute à chaque fois le même nombre pour obtenir le nombre suivant.

Quel est le nombre suivant qu'il devrait écrire dans sa suite logique ?

Le taux de réponses correctes pour la France est 59 % alors qu'il s'élève à 70 % pour l'ensemble des pays participants à l'enquête. A cet exercice, comme à bien d'autres, de niveaux de difficulté divers, le score de la France est le plus bas des pays de l'OCDE (hors le Chili).

Exemple 2 :

Célia a 12 longueurs de fil, 40 perles rondes, et 48 perles plates.

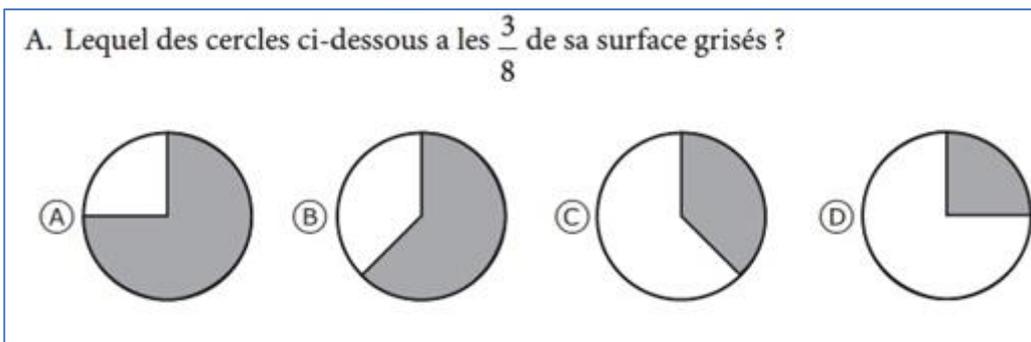
Elle utilise 1 longueur de fil, 10 perles rondes, et 8 perles plates pour fabriquer 1 bracelet.

Si Célia fabrique des bracelets tous identiques, combien peut-elle en fabriquer ?

(A) 40
(B) 12
(C) 5
(D) 4

Le taux de réponses correctes pour la France est 28 % alors qu'il s'élève à 37 % pour l'ensemble des pays participants à l'enquête.

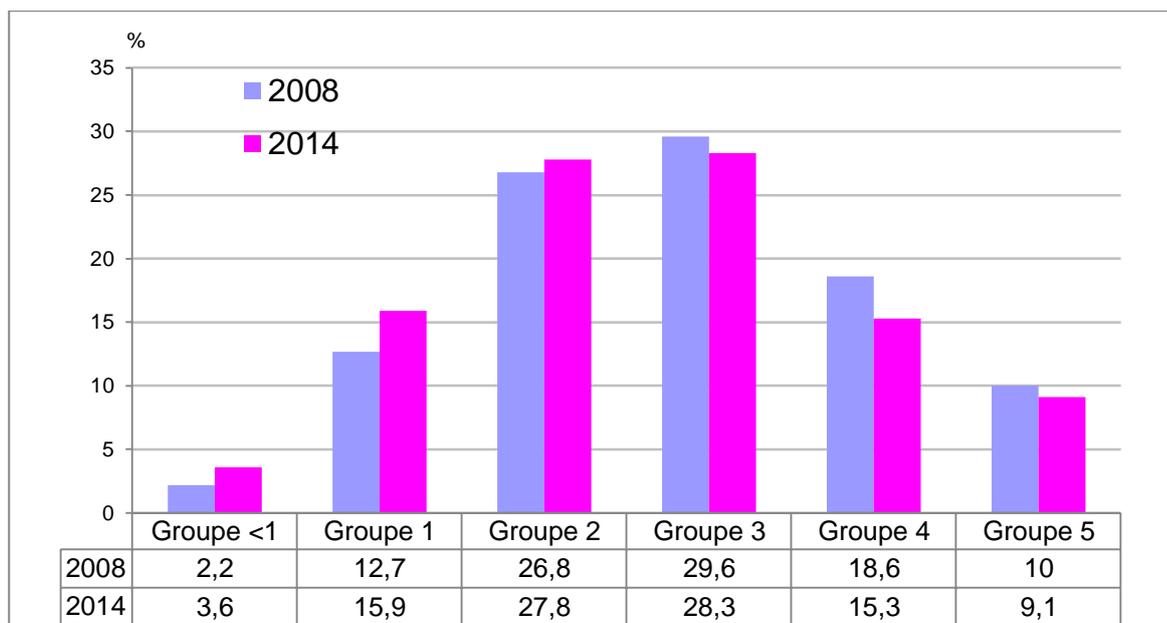
Exemple 3 :



Le taux de réponses correctes pour la France est 15 % alors qu'il s'élève à 24 % pour l'ensemble des pays participants à l'enquête.

Au collège, les difficultés des élèves s'accroissent, comme le montre les résultats de l'évaluation Cedre en fin de troisième.

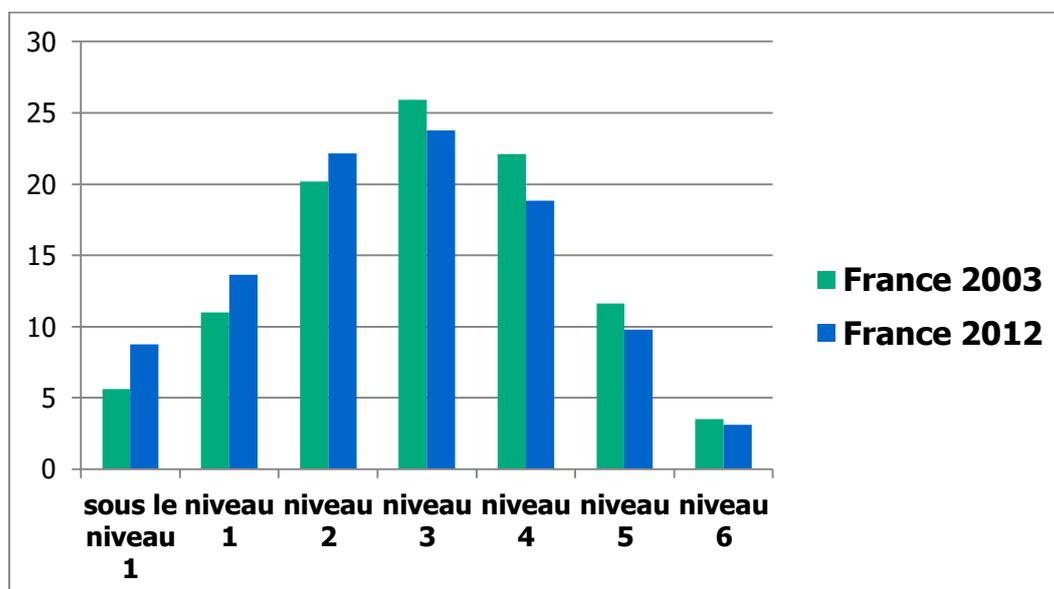
Figure 2 : Répartition des élèves par niveaux de compétence aux évaluations Cedre fin de collège en 2008 et en 2014



Source : Depp-MENESR. Note d’information N°19, DEPP, Mai 2015

L’augmentation de la proportion des élèves des groupes 1 et inférieurs à 1 est nette, alors que celle des groupes 4 et 5 diminue, ce qui traduit une nette détérioration du niveau des élèves entre 2008 et 2014. Les élèves des groupes 1 et 2 maîtrisent les opérations sur les nombres entiers, relatifs compris, et savent résoudre des situations de proportionnalité dans des cas simples. Mais l’utilisation du calcul littéral est une difficulté pour eux, tout comme celle des nombres rationnels. Là aussi, cette évolution se retrouve dans les résultats des évaluations internationales

Figure 3 : Répartition des élèves par niveaux de compétence aux évaluations PISA en 2003 et en 2012



Source : Depp-MENESR. Note d’information 13.31

De plus, et c’est un résultat bien identifié désormais mais néanmoins très inquiétant, il existe en France une forte corrélation entre l’environnement socio-économique des élèves et leur réussite scolaire. Ainsi, selon Cedre école 2014, les scores des écoles socialement défavorisées baissent, et notamment dans les écoles en éducation prioritaire, le nombre d’élèves en grande difficulté

augmente. Cette corrélation devient de plus en plus marquée, ce qui se traduit dans PISA par une détérioration de l'équité entre 2003 et 2012, telle qu'elle est mesurée par l'enquête : l'écart de score associé à la variation d'une unité de l'indice socio-économique a augmenté depuis 2003 de 14 points, soit l'augmentation la plus importante de tous les pays de l'OCDE, pour passer à 57 points, qui devient en 2014 l'écart plus important parmi les pays de l'OCDE, et renseigne donc sur des inégalités à l'école d'origine socio-économique très fortement marquées en France. .

Enfin, la France compte parmi les pays où les élèves ont le moins confiance en eux concernant leurs compétences et font le moins preuve de persévérance pour résoudre des problèmes. Selon PISA 2012, la France se classe toujours parmi les pays de l'OCDE où le niveau d'anxiété est le plus élevé, ce niveau n'ayant pas changé depuis 2003 (Cnesco, 2016).

L'ensemble de ces résultats conduit donc à des constats alarmants, d'autant plus que le cadre de référence des enquêtes TIMSS comme celui de Cedre est fondé sur les programmes scolaires contrairement à celui de PISA qui s'en émancipe délibérément. Cela implique nécessairement des lectures différentes des résultats de ces deux enquêtes, mais le fait que ces résultats se rapprochent constitue une alerte sérieuse...même si certaines limites liées aux évaluations elles-mêmes existent :

- La nature des items, et en particulier le format QCM, qui oriente les réponses fournies par les élèves selon les réponses possibles (par exemple, selon que, parmi les résultats proposés au calcul $3,8 + 1,2$ un distracteur fortement attractif comme 4,10 figure ou non, les choix des élèves peuvent différer). Mais, au-delà du format des items, des contraintes fortes liées à la nécessité de pouvoir construire une échelle de performance impose d'écarter certains items estimés non satisfaisants au regard de ce modèle, après un test expérimental (caractère unidimensionnel du modèle statistique utilisé appelé « modèle de réponse à l'item »).
- Dans le même temps, le fait que les items ne soient que très partiellement libérés ne permet pas d'apprécier sur quel nombre d'items et sur quelles tâches sont fondées les échelles de compétences, et donc d'avoir une connaissance effective de ce qui est mesuré.
- Les questions de la langue, de traduction (PISA, TIMSS) et le fait que dans les tests PISA, le champ de connaissances ou le domaine qu'il mobilise ne soit pas spécifié aux élèves dans les exercices proposés, peuvent constituer des facteurs supplémentaires de difficulté pour des élèves habitués à résoudre des tâches bien identifiées comme relevant d'une matière donnée.
- Les évaluations standardisées comme PISA ou Cedre se déroulent en mai. Ces tests sont présentés aux élèves comme non notés. Il est tout à fait vraisemblable que des élèves, en situation scolaire difficile, ne fassent pas leur maximum pour les réussir au mieux ... En particulier, le taux élevé de non-réponses des élèves français a pu être attribué à la peur de mal faire. Il est peut-être dû aussi à un manque de motivation pour un travail à faible enjeu dans un environnement scolaire où la note est prépondérante.

Pour davantage d'information sur les contenus de ces évaluations, on pourra se référer aux travaux nombreux de Bodin sur PISA notamment (2006, 2008), à l'article de Roditi et Salles qui ont effectué une nouvelle analyse des items de PISA 2015 en s'intéressant aux niveaux de mise en fonctionnement des connaissances mathématiques provoquées par les tâches proposées aux élèves (2015), à la thèse de Grapin (2015) qui s'est intéressée à la validation des tests utilisés par Cedre fin

d'école, et enfin au rapport du Cnesco consacré à la fois à l'analyse des cadres de référence des enquêtes PISA et TIMSS et à leurs contenus (Cnesco, 2016).

Quoi qu'il en soit, il nous paraît tout à fait essentiel de ne pas occulter ces résultats, ne serait-ce que par qu'ils résonnent avec d'autres résultats qui sont autant de signaux inquiétants pour le système scolaire et l'avenir de la France : la difficulté de recruter des enseignants en mathématiques au second degré, d'affecter ceux du premier degré dans les écoles difficiles (Cnesco, 2016) ; un sentiment de compétence perçue/niveau de satisfaction professionnelle mitigé, notamment chez les enseignants du 1^{er} degré (TIMSS 2015), l'inadéquation déclarée de la formation initiale avec les conditions réelles d'exercice et le déficit de formation continue des enseignants (TALIS, 2013). Notre point de vue est donc que ces évaluations méritent d'être prises en considération, parmi d'autres éléments producteurs d'informations sur l'état des connaissances des élèves, et qu'elles doivent l'être, à tous les niveaux, aussi bien aux niveaux national, académique et départemental qu'à celui des établissements et des classes.

Quels acquis des élèves dans le domaine des nombres et du calcul à la charnière école/collège?

Nos propres travaux de recherche, complétés par des analyses que nous avons menées avec Jean-Paul Fischer dans le cadre de la conférence de consensus du Cnesco et de l'Fé/ENS de Lyon en novembre 2015 montrent des difficultés récurrentes et persistantes que manifestent les élèves à l'entrée au collège :

- L'écriture des grands nombres ;
- La connaissance des tables, en particulier des tables de multiplication ;
- Les nombres décimaux (écriture, comparaison, intercalation, encadrement, multiplication par 10 - 100 - 1 000) ;
- Le calcul mental (« sens des nombres », ordre de grandeur d'un résultat) ;
- Le calcul posé ;
- Les résultats sur la résolution de problèmes qui sont très variables.

L'analyse approfondie de ces évaluations (contenus et résultats) nous a conduit à formuler deux hypothèses fortes, qui auraient des incidences fortes en termes de contenus à enseigner, de supports utilisés en classe et de pratiques enseignantes :

- une réussite opératoire sur des tâches portant sur les nombres entiers pourrait masquer une conceptualisation insuffisante de ces nombres ;
- les difficultés repérées sur les nombres décimaux pourraient renvoyer à cette conceptualisation insuffisante des nombres entiers.

TIMSS 2015 confirme les difficultés observées, en les rendant plus flagrantes au regard des résultats d'autres pays :

- dans les trois domaines cognitifs du cadre de référence de TIMSS (connaître, appliquer, raisonner), les scores des élèves français sont bien inférieurs à la moyenne des pays européens participants ;
- les résultats des élèves français sont particulièrement faibles sur les questions portant sur les nombres et le calcul et sur la gestion de données (lecture et interprétation de tableaux et graphiques). ;
- dans l'ensemble des pays participants, 52 % des enseignants s'estiment très satisfaits d'exercer leur métier, alors qu'ils ne sont que 30 % en France.

Vers de nouvelles voies?

Cette question de la pertinence, de la compatibilité des évaluations externes au regard des contextes professionnels des enseignants est investiguée depuis longtemps par les sciences de l'éducation, plus récemment par la didactique.

Une des questions importantes, une fois qu'on a traité de la validité et de la fiabilité de ces évaluations, est posée ainsi par Marcel Crahay (*in* Mottier-Lopez & Crahay, 2009) : « Articuler l'évaluation en classe et le pilotage des systèmes, est-ce possible ? »

« Le défi est de taille. Le « challenge » paraît impossible tant le pilotage par évaluations externes a été présenté comme antinomique de l'évaluation interactive que pratiqueraient les enseignants dès lors qu'ils gèrent la vie d'une classe tout au long de l'année. Pourtant, écrivons-le d'emblée, il nous semble possible de coordonner ces deux pôles d'un faisceau de tensions qui peuvent s'avérer positives dès lors que certaines conditions sont réunies. »

De Ketele (2017), lors sa conférence d'ouverture du dernier colloque de l'Association pour le Développement des Méthodologies d'Évaluation en Éducation (ADMEE), a effectué une tentative de modélisation très intéressante qui intègre les différentes dimensions de l'évaluation. Nous partageons cette vision globale de l'évaluation, et nous pensons qu'il est possible de compléter cette modélisation, en imaginant de nouvelles fonctions et de nouveaux usages pour les évaluations externes des acquis des élèves, qui combinent plusieurs aspects que nous avons évoqués ci-dessus.

Un exemple est celui du dispositif PACEM que nous avons conçu et mis en œuvre entre 2010 et 2012, dans les académies de Créteil et d'Aix-Marseille, et qui a produit des résultats positifs sur les acquis des élèves. Ce dispositif qui repose sur une conception spécifique de l'enseignement des mathématiques et de la formation des enseignants, consiste à partir d'un test standardisé adapté comme amorce d'une formation d'enseignants. Ce test présente les caractéristiques suivantes :

- les items ont été validés statistiquement ;
- il porte sur un champ d'évaluation limité ;
- les tâches proposées sont familières aux élèves et aux enseignants ;
- la durée de passation du test est « raisonnable ».

On obtient ainsi un matériau facilement partageable entre enseignants d'une même école ou d'un même établissement, et entre écoles et établissements. La formation des enseignants qui est ancrée sur ce test, vise explicitement l'amélioration des acquis des élèves en misant sur l'évolution des pratiques professionnelles et contient les éléments suivants :

- un questionnement des processus d'apprentissage des élèves et des pratiques des enseignants;
- des apports mathématiques et didactiques;
- un travail sur les contenus et sur les déroulements pour co-construire et « conscientiser » une palette de « bonnes pratiques possibles », avec un objectif de partage de pratiques.

Ces évaluations standardisées, sous certaines conditions, peuvent donc constituer des éléments parmi d'autres, de développement professionnel des enseignants en mathématiques : elles peuvent en effet générer une réflexion à la fois individuelle et collaborative, intégrée aux contextes et aux enjeux locaux, et prioritairement centrée sur les apprentissages des élèves.

À l'heure où un retour des évaluations standardisées exhaustives des acquis des élèves est possible, il nous semble donc important d'une part de considérer ces évaluations comme des objets de recherche, et d'autre part que tous les acteurs du système éducatif réfléchissent ensemble comment les définir, comment les mettre en œuvre et comment les prendre en compte.

Références

Bodin, A. (2008). Lecture et utilisation de PISA pour les enseignants. *Petit x* ; n° 78, pp. 53-78, IREM de Grenoble.

Chesné, J.-F. (2014). *D'une évaluation à l'autre : des acquis des élèves sur les nombres en sixième à l'élaboration et à l'analyse d'une formation d'enseignants centrée sur le calcul mental*. Thèse de doctorat. Université Paris-Diderot.

Chesné, J.-F. & Fischer, J.-P. (2015). *Les acquis des élèves dans le domaine des nombres et du calcul à l'école primaire*. Rapport rédigé pour le Conseil national d'évaluation du système scolaire (Cnesco).

Cnesco (2015). *Nombres et opérations : premiers apprentissages à l'école primaire*. Dossier de synthèse. <http://www.cnesco.fr/fr/numeration/>

Cnesco (2016). *Acquis des élèves : comprendre les évaluations internationales PISA et TIMSS*. Dossier de synthèse. <http://www.cnesco.fr/fr/comparaison-pisa-timss/>

Cnesco (2016). *Résultats de l'enquête internationale TIMSS : éclairage du Cnesco sur les mathématiques au primaire*. Note d'actualité.

De Ketele, J.-M. (2017). A quoi bon évaluer. *Actes du colloque de l'ADMEE 2017* (à paraître).

Grapin, N. (2015). *Étude de la validité de dispositifs d'évaluation et conception d'un modèle d'analyse multidimensionnelle des connaissances numériques des élèves de fin d'école*. Thèse de doctorat. Université Paris-Diderot.

HCE (2011). *Les indicateurs relatifs aux acquis des élèves. Bilan des résultats de l'école*.

IGEN-IGAENR (2005). *Les acquis des élèves, pierre de touche de la valeur de l'école ?* Rapport n°2005-089.

Mons, N. & Pons (2006). *Les standards en éducation dans le monde francophone. Une analyse comparative*. Neuchâtel : IRDP.

Mons, N. (2007). *Les nouvelles politiques éducatives*. Paris : PUF.

Mottier Lopez, L. & Crahay, M. (sous la direction de) (2009) *Evaluations en tension. Entre la régulation des apprentissages et le pilotage des systèmes*. De Boeck : Bruxelles.

Roditi, E. & Salles, F. (2015). Nouvelles analyses de l'enquête Pisa 2012 en mathématiques. Un autre regard sur les résultats. *Éducation & Formations*, n°86-87. MENESR-Depp.

Trosseille, B. et Rocher, T. (2015). Les évaluations standardisées des élèves. Perspective historique. *Educations & Formations*, n°86-87. MENESR-Depp.