



cnesco
conseil national
d'évaluation
du système scolaire

L'évaluation des élèves par les enseignants dans la classe et les établissements : réglementation et pratiques.

Une comparaison internationale dans les pays de l'OCDE

Décembre 2014

Cnesco
31-35 rue de la Fédération
75 015 Paris
Sec. 01 55 55 02 09
secretariat.cnesco@education.gouv.fr

Pour plus d'informations :
cnesco.fr

Rejoignez-nous :   

Ce rapport est publié par le Conseil national d'évaluation du système scolaire, avec les contributions de Nathalie Mons, professeure de sociologie à l'Université de Cergy-Pontoise, présidente du Cnesco, de Marie-Christine Toczek, professeure en sciences de l'Éducation à l'Université Blaise-Pascal de Clermont-Ferrand, directrice du laboratoire Activités, connaissance, transmission, éducation (Acté), membre du Cnesco, de Anne West, professeure en politiques éducatives à la London School of Economics, directrice du groupe de recherche éducation au département politiques sociales, membre du Cnesco, et de Sadiya Barkouss, chargée de missions au Cnesco.

Les opinions et arguments exprimés n'engagent que le Cnesco.

Disponible sur le site du Cnesco : <http://www.cnesco.fr>

Publié en décembre 2014
Conseil National d'évaluation du système scolaire
Carré Suffren - 31-35, rue de la Fédération
75 015 Paris
© Cnesco

SOMMAIRE

Introduction	7
I. La complexification des cadres réglementaires en matière d'évaluation des élèves dans la classe dans les pays de l'OCDE	9
A. Des réglementations nationales des évaluations internes de plus en plus contraignantes	12
B. Une plus grande variété d'évaluations imposées aux enseignants dans les pays de l'OCDE	15
C. Des usages légaux de l'évaluation multiples	22
II. Des pratiques d'évaluation différenciées dans les pays de l'OCDE	26
A. La variété des évaluations mobilisées par les enseignants dans la classe	26
B. Les collaborations entre enseignants autour de l'évaluation des élèves dans l'établissement	34
C. Le besoin de formation continue des enseignants en matière d'évaluation des élèves	37
Conclusion de la présidente du Conseil national d'évaluation du système scolaire	38
Références	40

Introduction

Parmi les administrations publiques, l'école est une des premières productrices d'évaluations. Car l'institution scolaire, responsable des apprentissages des plus jeunes, produit de l'évaluation pour suivre quotidiennement les acquis des élèves dans la classe et communiquer avec les parents sur les progrès de leur enfant. Industrie de présélection sociale, elle produit aussi des évaluations dans le cadre des diplômes qu'elle décerne et contribue ainsi partiellement à définir les niveaux d'emplois et donc les destinées sociales.

L'école produit en conséquence des évaluations d'élèves variées que les experts qualifient soit de **formatives** – lorsque les évaluations aident l'élève en continu dans ses apprentissages - soit de **sommatives** - lorsque les évaluations sont des repères pour les décisions de l'institution en matière de carrière et d'orientation scolaires, par exemple pour un passage dans la classe supérieure, un redoublement ou un choix de filière. Dans de nombreux pays, ces deux dimensions s'entrecroisent.

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, plus particulièrement depuis les années 1990, dans les pays de l'OCDE une panoplie d'évaluations externes - c'est-à-dire qui relèvent d'acteurs extérieurs à l'établissement scolaire - 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.

Le constat est donc celui d'un très fort développement des évaluations dans l'univers scolaire (Mons (2007), Dupriez et Malet (2013)), qu'elles concernent les élèves, les établissements et bien sûr les enseignants comme équipe pédagogique, dans certains pays, ou à titre individuel, à travers des inspections comme en France. Cette amplification des évaluations dans l'école permet ainsi de remplir le spectre large des objectifs que Cardinet (1989) leur assignait dès les années 1980 : 1) améliorer les décisions relatives aux apprentissages de chaque élève ; 2) informer sur sa progression l'enfant et ses parents ; 3) décerner les certifications nécessaires à l'élève et à la société ; enfin, 4) informer pour améliorer la qualité de l'enseignement en général.

Assez étrangement, ce sont les nouvelles évaluations externes - celles pratiquées par des institutions externes à l'établissement scolaire - qui ont donné lieu à des recherches récentes qui en décrivent les développements. **Les évaluations internes par les enseignants dans la classe, bien qu'au cœur historiquement de l'institution scolaire, et analysées, à travers différents champs scientifiques (les sciences de l'éducation, la psychologie...), par des chercheurs s'intéressant à la pédagogie (en particulier aux effets des évaluations sur les apprentissages des élèves), ont suscité peu d'états des lieux des pratiques des enseignants dans leurs classes. Plus particulièrement, les comparaisons internationales portant sur les pratiques évaluatives des enseignants sont très peu nombreuses.**

C'est, en conséquence, dans cette direction que s'est orientée l'étude du Cnesco, qui est mandaté par la « loi n° 2013-595 du 8 juillet 2013 d'orientation et de programmation pour la refondation de l'école de la République » pour mener, notamment, des enquêtes à dimension internationale. **L'investigation présentée ici, conduite dans le temps très limité imparti par la « Conférence nationale sur l'évaluation des élèves », dresse un tout premier bilan des cadres réglementaires qui régissent les**

évaluations par les enseignants dans leur classe et leur établissement, et de la réalité de ces pratiques pédagogiques dans les pays de l'OCDE.

Qu'exigent des enseignants les réglementations de chaque pays en matière d'évaluation des élèves par dans leurs classes, et que font-ils concrètement en la matière ? L'intérêt, ici, est de faire se répondre réglementations et pratiques de classe car, certes, les premières encadrent les secondes ou, du moins, cherchent à les encadrer, mais, en retour, les pratiques collégiales elles-mêmes, notamment celles des collectifs pédagogiques dans les établissements, peuvent générer des normes non légales qui conduisent à réguler les pratiques des enseignants dans leur classe. Les deux points de vue - réglementation et pratiques - sont donc complémentaires pour comprendre comment les enseignants évaluent les élèves dans leurs classes.

Le périmètre de l'étude couvre, selon les informations disponibles, le primaire – l'équivalent de notre école élémentaire - ainsi que le secondaire - l'équivalent de notre collège et de notre lycée -. Le champ géographique de ce rapport couvre les pays de l'OCDE¹, un ensemble de pays économiquement et socialement comparables à la France, mais diversifiés en termes d'aires géographiques (Europe, Amérique, Asie). À ce stade, il s'agit pour le Cnesco d'une *toute première* analyse, rendue ardue par la faiblesse des données de comparaisons internationales, que le Cnesco enrichira par la suite, ses activités intégrant la thématique des pratiques évaluatives au sens large (cf. *Programme d'orientations stratégiques 2014-2017*, juin 2014).

Même si les données réunies ici ne permettent pas de produire une vue exhaustive des réglementations et pratiques en matière d'évaluations relevant de la responsabilité des enseignants, **la force de la stratégie de recherche fondée sur la comparaison internationale apparaît clairement. Par détour, la comparaison internationale révèle les caractéristiques saillantes du cas français, c'est-à-dire à la fois ses particularités – « l'exception française », notamment la faiblesse de la réglementation autour de la notation jusque dans des années récentes - et clairement son inscription dans un mouvement de réformes internationales depuis une quinzaine années.** C'est le cas notamment avec le développement, au moins légal, du « nouveau socle commun de connaissances, de compétences et de culture (article 13 de la loi du 8 juillet 2013) » qui définit de nouveaux critères d'évaluation des élèves par les enseignants. La France n'est donc pas ce village gaulois qui passerait à côté du raz de marée international des réformes en éducation.

Pour présenter ces analyses, le rapport s'articule autour de deux parties principales. La première partie est dédiée à une analyse des cadres légaux qui régissent l'évaluation des élèves relevant de la responsabilité des enseignants, dans un échantillon réduit et diversifié de pays de l'OCDE. Plusieurs dimensions réglementaires sont auscultées pour porter un regard complet sur les différentes facettes de l'évaluation : les pratiques de notation sont-elles encadrées légalement, les enseignants doivent-ils harmoniser leurs évaluations au sein de leur établissement, les palmarès d'élèves sont-ils autorisés, comment les familles doivent-elles être informées des progressions scolaires de leurs enfants, comment tient-on compte légalement des résultats scolaires de l'élève dans le redoublement... ? En complément de l'analyse de ces cadres légaux, **la seconde partie du rapport s'attache à l'étude des pratiques des enseignants dans les classes et les établissements, en mobilisant différentes enquêtes internationales menées sur des échantillons larges de pays de l'OCDE.**

¹ Organisation de coopération et de développement économiques.

I. La complexification des cadres réglementaires en matière d'évaluation des élèves dans la classe dans les pays de l'OCDE

L'évaluation des élèves par les enseignants se révèle une pratique complexe qui articule deux dimensions qui peuvent paraître comme antagonistes (Mialaret, 1979). Tout d'abord, l'évaluation interne est un *jugement* de l'enseignant, sur les progrès scolaires de son élève, produit dans un contexte de classe, requérant une forme de liberté pédagogique sans laquelle l'acte d'enseigner ne peut être adapté au contexte local et aux élèves que l'enseignant doit accompagner. Elle est également, le plus souvent en parallèle - car évaluation sommative et formative se croisent dans nombre de pays - l'appréciation de la valeur de la production scolaire sur la base de critères *objectifs*, compte tenu des enjeux des évaluations, déjà soulignés en introduction - à savoir la sélection scolaire et la certification. Subjectivité et nécessaire objectivité, liberté et contraintes s'entre-mêlent sans cesse dans le travail d'évaluation de l'enseignant dans sa classe.

Aussi, traditionnellement, les réglementations scolaires qui encadrent les évaluations internes ont oscillé entre reconnaissance de la liberté pédagogique des enseignants et contraintes externes, légales, imposées aux enseignants en la matière. Ces contraintes peuvent porter sur plusieurs dimensions que le Cnesco a choisi d'étudier dans ce rapport : la nature des évaluations, la notation, la fréquence des épreuves administrées aux élèves, l'harmonisation des évaluations dans les établissements, leur communication en direction avec les familles...

Ne pouvant se fonder sur une recherche existante portant sur la comparaison internationale des réglementations de l'évaluation interne, le Cnesco a développé une méthodologie propre pour cette analyse, en identifiant à la fois des critères précis de comparaison de politiques d'évaluation interne et un échantillon de pays qui bien que restreint puisse informer à la fois par la diversité des zones géographiques qu'il comporte et la variété des politiques nationales d'évaluation qu'il implique (cf. encadré 1 sur la méthodologie de notre enquête ci-dessous²).

² Dans ce rapport, nous n'étudions pas, pour tous les pays de notre échantillon, l'évaluation dans l'enseignement professionnel. Cette dimension sera intégrée dans une étude ultérieure qui sera publiée, par le Cnesco, sur « L'évaluation dans la classe et l'établissement ».

Encadré 1. Méthodologie de l'enquête Cnesco sur la comparaison internationale des réglementations d'évaluation interne

Dans le cadre de notre étude sur la réglementation en matière d'évaluation interne des élèves, la France est comparée à un échantillon restreint de pays majeurs de l'OCDE. Ils ont été choisis à la fois pour leur diversité géographique (Europe, Asie, Amérique), la variété des modèles de gouvernance scolaire (décentralisés comme l'Allemagne ou plus centralisés comme la Corée) ainsi que pour la variété des modèles d'évaluation qu'ils incarnent, soit situés dans des cadres de compétition scolaire vive (comme la Corée), soit positionnés dans un refus explicite de la comparaison ou du classement entre élèves (la Finlande ou la Suède). Certains pays ont aussi été sélectionnés parce qu'ils sont emblématiques des politiques innovantes en matière d'évaluation formative développées depuis les années 2000, comme le Canada ou l'Angleterre. L'échantillon de pays sélectionnés permet ainsi à la fois de mettre en évidence certaines tendances générales qui ont sous-tendu les évolutions en matière d'évaluation interne dans le concert international des réformes de l'évaluation depuis trois décennies mais aussi de présenter des modèles pionniers d'évaluation qui peuvent, non pas être transposés dans le cadre français mais au moins interroger la situation française.

Afin d'apprécier la spécificité et la singularité de chaque réglementation en matière d'évaluation des élèves dans les pays sélectionnés, l'étude dresse un panorama de certains critères parmi lesquels les normes de références utilisées pour l'évaluation des élèves dans la classe par les enseignants, le nombre d'acteurs et la part de responsabilité de chacun dans le processus de l'évaluation, les modalités d'évaluation (formes et fréquence), le système de notation en vigueur, le degré de coopération dans l'école ou l'établissement en matière d'évaluation, la communication avec les parents et certains des usages faits de l'évaluation (classement des élèves, critères pour le redoublement).

L'analyse se fonde sur la base des sources primaires émanant des ministères de l'éducation concernés quand ces données étaient disponibles en langue anglaise ou sur des données collectées dans la base de données de comparaison des systèmes scolaires, produite par l'agence européenne Eurydice, la base Eurypedia (2010-2013), qui peut parfois présenter des décalages dans le temps avec les législations les plus récentes. Des experts nationaux ont aussi été mis à contribution pour choisir les pays à intégrer dans notre comparaison internationale et vérifier les informations nationales. Dans la suite du texte, nous utilisons les termes consacrés par la classification internationale des niveaux d'enseignement (CITE) : le primaire - niveau CITE 1 dans cette classification - le collège - CITE 2 et le lycée - CITE 3³.

³ Dans ce rapport, nous n'étudions pas, pour tous les pays de notre échantillon, l'évaluation dans l'enseignement professionnel. Cette dimension sera intégrée dans une étude ultérieure qui sera publiée, par le Cnesco, sur « L'évaluation dans la classe et l'établissement ».

Nous articulons la présentation des résultats de cette enquête de comparaisons internationales sur l'évaluation interne autour de trois parties. La première section s'intéresse à l'existence d'une architecture nationale conditionnant les évaluations en classe et permettra d'analyser l'autonomie des enseignants dans ce domaine. Nous verrons ainsi que, depuis trois décennies, les réglementations sont devenues plus complexes en la matière, restreignant progressivement la liberté pédagogique des équipes enseignantes dans les pays étudiés. La deuxième section, dédiée à une analyse fine des cadres légaux relatifs aux modalités d'évaluation (nature des épreuves, fréquences, définition de systèmes de notation...), laissera voir la très grande variété des dispositifs mis en place dans les pays de l'OCDE. Enfin, la dernière section présente les cadres légaux relatifs aux différents usages de ces évaluations internes (classement des élèves, décision de redoublement ou de passage...).

A. Des réglementations nationales des évaluations internes de plus en plus contraignantes

L'analyse des pays sélectionnés dans notre échantillon met en évidence un fort développement des réglementations en matière d'évaluation par les enseignants dans leurs classes dans tous ces contextes nationaux depuis les années 1970, contraintes réglementaires qui se sont encore amplifiées depuis les années 1990.

Ces normes nationales ou infra-nationales sont présentées dans le tableau 1 ci-dessous.

Tableau 1 : Normes nationales sur l'évaluation des élèves par les enseignants dans la classe et autonomie des enseignants dans une sélection de pays de l'OCDE – Niveaux CITE 1, 2 et 3⁴ (2010-2013)

Critères réglementaires	France	Angleterre, Pays de Galles, Irlande du Nord	Allemagne	Suède	Danemark	Finlande	Québec	Corée
Normes de référence utilisées pour l'évaluation	<p>CITE 1, 2 ET 3 : Objectifs nationaux des programmes.</p> <p>CITE 2 et 3 : Certains critères d'évaluation des examens nationaux sont utilisés comme référence.</p>	<p>CITE 1, 2 et 3 : Normes dictées par l'organisme "The Office of Qualifications and Examinations Regulation".</p>	<p>CITE 1, 2 et 3 : Objectifs curriculaires au niveau des Länder.</p>	<p>CITE 1, 2 et 3 : Normes nationales édictées en termes de syllabus précisant par matière des critères d'évaluation attachés aux niveaux de l'échelle de notation.</p>	<p>CITE 1, 2 et 3 : Objectifs curriculaires au niveau national.</p>	<p>CITE 1, 2 et 3 : Objectifs curriculaires au niveau national.</p>	<p>CITE 1, 2 et 3 : Le cadre réglementaire dicte les normes à respecter notamment en termes de programmes et de modalités d'évaluation. Les cadres d'évaluation prescrits pour chaque matière précisent les apprentissages à évaluer et leur importance relative et les critères d'évaluation.</p>	<p>CITE 1, 2 et 3 : Programmes curriculaires au niveau national.</p>
Acteurs de l'évaluation dans la classe, l'école et l'établissement	<p>CITE 1, 2 et 3 : Large autonomie et liberté pédagogique des maîtres et enseignants. L'EN dans le primaire et le chef d'établissement dans le secondaire sont responsables d'assurer la conformité avec les normes de référence précitées.</p>	<p>CITE 1, 2 et 3 : Les chefs d'établissement s'assurent que les normes sont respectées. Autonomie limitée des enseignants.</p>	<p>CITE 1, 2 et 3 : Autonomie des enseignants dans le cadre des formes d'évaluation proposées.</p>	<p>CITE 1, 2 et 3 : Les enseignants disposent d'une autonomie limitée dans le choix des modalités évaluatives fixées nationalement.</p>	<p>CITE 1, 2 et 3 : Autonomie des enseignants dans le cadre des types d'épreuves proposées.</p>	<p>CITE 1, 2 et 3 : Proposition quant à l'utilisation de l'auto-évaluation, des examens blancs, et des évaluations écrites.</p>	<p>CITE 1, 2 et 3 : Politique d'évaluation ministérielle. Responsabilité partagée entre différents acteurs (Ministère, commissions scolaires et écoles). Autonomie des enseignants dans le choix des moyens d'évaluation (le COMMENT évaluer).</p>	<p>CITE 1, 2 et 3 : Politique d'évaluation des élèves au niveau de chaque établissement. Les enseignants suivent les directives.</p>

Source : Cnesco, 2014

⁴ Dans ce rapport, nous n'étudions pas, pour tous les pays de notre échantillon, l'évaluation dans l'enseignement professionnel. Cette dimension sera intégrée dans une étude ultérieure qui sera publiée, par le Cnesco, sur « L'évaluation dans la classe et l'établissement ».

En effet historiquement, les normes curriculaires - comme les programmes scolaires français - ont formé le cadre principal des évaluations qui, en conséquence, portaient strictement sur ces contenus d'enseignement prescrits. Aux côtés de ces pays, caractérisés par des programmes scolaires, dont la France est emblématique, dans les autres États - numériquement peu nombreux jusque dans les années 1990, les établissements scolaires et donc les équipes pédagogiques bénéficiaient d'une large autonomie tant sur le choix des contenus d'enseignement que sur les évaluations des élèves. La liberté pédagogique des équipes ne signifiait pas pour autant une absence de contraintes pour l'enseignant dans son travail évaluatif. Au contraire, **ces pays à forte autonomie des établissements, comme l'Angleterre dans notre échantillon, mais aussi la Belgique ou les Pays-Bas étaient caractérisés par des collectifs-enseignants locaux très puissants, et un travail en équipe orientant, de fait, le travail individuel de chaque enseignant.** Mais, peu de normes nationales encadraient le travail de l'enseignant dans la classe. L'Angleterre a attendu la décennie 1980 pour imposer le premier *National curriculum* dans un ensemble de matières.

Cette liberté pédagogique fut corsetée doublement à partir de 1990. Dans nombre de pays sans programmes scolaires nationaux, comme la Suisse, l'Allemagne, la Belgique ou certains États américains et australiens, des normes curriculaires sont imposées à travers des *standards* – qui sont des objectifs d'apprentissages imposés aux niveaux clefs de la carrière scolaire de l'élève. Ces *standards*, de fait, se sont imposés au travail de l'enseignant dans la classe.

Un peu plus tard, les pays traditionnellement attachés aux programmes scolaires vont eux-aussi développer des cadres nationaux sous forme d'objectifs pédagogiques, dont le socle commun en France est emblématique.

Ces programmes scolaires par objectifs que l'on retrouve très développés dans deux pays de notre échantillon - l'Angleterre et la Suède - encadrent de façon très détaillée les évaluations internes. Par exemple, en Suède, les différents niveaux de la notation mise en place à la rentrée 2012-2013 - (les lettres A, C et E, voir section suivante sur les systèmes de notation et le schéma 1 ci-dessous) correspondent de façon très précise à des critères de résultats et de compétences des élèves pré-définis dans les syllabus (les programmes scolaires suédois par objectifs). Dans un ensemble de pays de notre échantillon, ces normes sont définies dans des guides qui orientent l'activité d'évaluation interne des enseignants.

Aux côtés de ces nouvelles politiques curriculaires qui définissent précisément les *objectifs* des évaluations des élèves à mettre en place par les enseignants dans la classe, plus récemment, certains pays comme le Québec ou l'Angleterre ont développé, aussi, des orientations spécifiques en matière d'évaluation formative - c'est-à-dire liée au suivi quotidien des apprentissages des élèves dans la classe, encadrant encore plus les *pratiques* des enseignants comme nous le verrons dans la suite de ce document.

Schéma 1. La nouvelle échelle de notation par lettres entrée en vigueur en 2012-2013 en Suède (de A « excellent » à F (échec au cours suivi))



Au total, malgré la réaffirmation constante dans nombre de pays de l'OCDE du principe de la liberté pédagogique des enseignants, le développement de nouvelles normes, sous forme de programmes scolaires formulés en objectifs pédagogiques ou en *standards* et souvent accompagnés de recommandations concrètes sur les pratiques d'évaluation, a progressivement encadré l'activité évaluative des enseignants dans la classe et dans les établissements. Cette multiplication des normes a conduit à des modalités d'évaluation, imposées aux enseignants, très diverses selon les pays comme nous le présentons dans la section suivante.

B. Une plus grande variété d'évaluations imposées aux enseignants dans les pays de l'OCDE

Cette deuxième section, dédiée à une analyse fine des cadres légaux relatifs aux modalités d'évaluation (nature des épreuves, fréquences, définition de systèmes de notation...), laisse voir la très grande variété des dispositifs progressivement mis en place dans les pays de l'OCDE (cf. tableaux 2 et 3).

Un développement de la variété des évaluations internes

L'analyse du tableau 2 met en évidence un fort développement des recommandations ou obligations en termes de variété des natures des épreuves que les enseignants doivent mobiliser dans leur activité évaluative en classe. Notamment dans certains pays pionniers, l'introduction de dispositifs d'évaluation innovants est suggérée ou imposée (cf. encadré 2 pour le cas de l'Irlande).

Ainsi, alors que certains pays comme la France privilégient les formes traditionnelles d'évaluation comme les devoirs écrits dont le contenu est laissé à la discrétion de l'enseignant, d'autres pays, comme l'Angleterre ou le Québec, préconisent ou imposent, dans les textes officiels, de mobiliser de nouvelles formes d'évaluations comme l'auto-évaluation, l'évaluation par les pairs ou le suivi individualisé des élèves.

Certains pays sont aussi attentifs à « protéger » les élèves de certaines évaluations trop formelles durant les premières années du primaire. Ainsi, en Allemagne, durant les deux premières années de la scolarité, les épreuves écrites sont interdites et les enseignants ne peuvent évaluer leurs élèves que par l'observation directe.

La fréquence des évaluations, bien différemment, se révèle, elle, peu contrainte par les réglementations.

Les mécanismes de collaboration entre enseignants, eux aussi peu réglementés se développent cependant. Ainsi certains pays imposent soit des harmonisations d'évaluation internes à l'établissement, soit la prise en compte par l'enseignant de références externes comme les résultats aux tests nationaux notamment dans les classes d'orientation comme en Suède. En Finlande et au Danemark, des guides de notation sont aussi disponibles aux niveaux CITE 2 (collège) et 3 (lycée) pour orienter les pratiques évaluatives des enseignants, sans obligation formelle de les suivre pour autant.

Tableau 2 : Réglementation sur les caractéristiques des dispositifs d'évaluation par les enseignants dans la classe (nature, fréquence et harmonisation des épreuves) dans une sélection de pays de l'OCDE – Niveaux CITE 1, 2 et 3 (2010-2013)

Critères réglementaires		France	Angleterre, Pays de Galles, Irlande du Nord	Allemagne	Suède	Danemark	Finlande	Québec	Corée
Modalités d'évaluation	Nature des épreuves	CITE 1, 2 : Devoirs écrits, contrôles de leçons, évaluation orale, observation de l'élève selon des grilles d'observation critériées particulièrement en maternelle. CITE 3 : Devoirs écrits, contrôles surveillés, tests oraux, etc.	CITE 1, 2, 3 : Auto-évaluation des élèves, l'évaluation par les pairs, les tests réalisés par les enseignants, le suivi individualisé des élèves, etc. (Cf http://ofqual.gov.uk/qualifications-and-assessments/national-assessments/)	CITE 1 : Obligation de limiter l'évaluation à l'observation directe des élèves (deux premières années du primaire), épreuves écrites réalisées par les enseignants dans certaines matières (mathématiques, allemand) à partir de la 3 ^{ème} année. CITE 2 et 3 : Contrôles écrits, oraux, travaux pratiques, etc.	Non réglementé.	CITE 1, 2, et 3 : Proposition d'utiliser l'auto-évaluation des élèves (orale et écrite), carnet de bord, portfolio, entretiens écrits réalisés par les enseignants, etc.	CITE 1, 2 et 3 : Proposition d'utiliser auto-évaluation, simulation de tests, évaluations écrites.	Non renseigné	CITE 1, 2 et 3 : Dépend de la matière évaluée. Généralement, les documents officiels citent les évaluations écrites, les évaluations de performances (auto-évaluation, rapports, travaux de recherches, etc.).
	Fréquence	CITE 1, 2 : Evaluations périodiques à la discrétion de l'enseignant. CITE 3 : Contrôles écrits selon une fréquence variable.	CITE 1, 2, 3 : Explicitement non réglementé mais très fréquente étant donné la diversité de la nature des épreuves d'évaluation existante.	CITE 1 : Très fréquente. CITE 2 et 3 : Evaluations périodiques plus soutenues en vue de la préparation des examens standardisés.	Non réglementé.	Non réglementé.	Non réglementé.	Non réglementé.	Non réglementé.
Coopération dans l'établissement et entre les établissements en matière d'évaluation interne des élèves	Mécanismes pour assurer la fiabilité de notation dans l'évaluation sommative	CITE 1 : Conseils des maîtres et conseils des cycles. CITE 2 : Conseils de classe et conseils pédagogiques. CITE 3 : Conseils de classe. CITE 1, 2 et 3 : Réunions entre enseignants à la demande du chef d'établissement.	CITE 1 : Pas de réglementation. CITE 2 et 3 : Coordination interne dans les	Pas de réglementation.	CITE 2 : en 9 ^{ème} année, les enseignants doivent prendre comme références les résultats aux tests nationaux dans leur CITE 3 : idem	CITE 2 ET 3 : Disponibilité de guides de notation des performances des élèves.	CITE 1, 2 ET 3 : Disponibilité de guides de notation des performances des élèves.	Recours à des références communes prescrites par les cadres d'évaluation (le Quoi évaluer). Existence d'un cadre de référence en évaluation (guide général) et guides de correction des épreuves provinciales.	Non réglementé.
	Existence de contrôles communs	CITE 1 : Aucun CITE 2 : Recommandation d'épreuves communes au niveau de l'établissement pour la préparation des examens nationaux (choix du chef d'établissement).	Non réglementé	Non réglementé	Non réglementé	Non réglementé	Non réglementé	Non renseigné	Non renseigné

Source : Cnesco, 2014

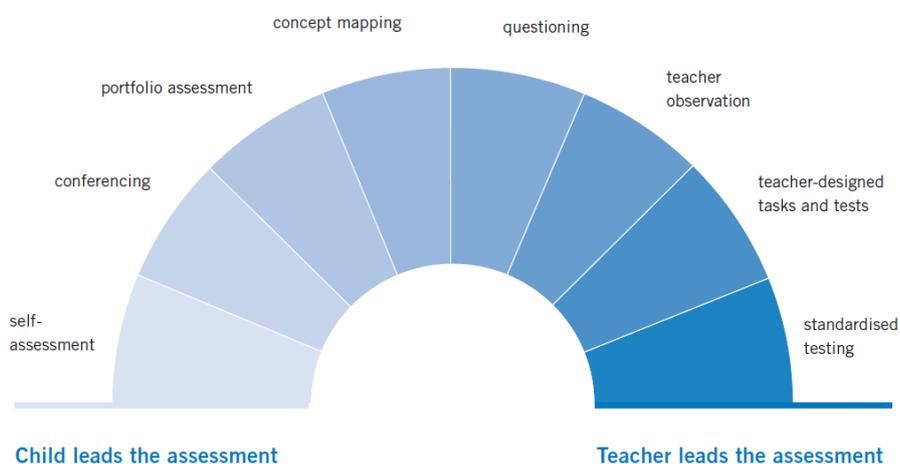
Encadré 2. Un continuum d'épreuves à disposition des enseignants au primaire

Le cas de l'Irlande

Certains pays de l'OCDE définissent des stratégies d'évaluation des élèves dans la classe dans le cadre de leur politique éducative nationale. C'est le cas de l'Irlande qui promeut l'évaluation formative au même titre que l'Angleterre. Les enseignants irlandais du primaire sont invités à utiliser le guide national pour l'évaluation des élèves dans la classe, « *Assessment in the Primary School Curriculum : Guidelines for Schools* », lancé par le *National Council for Curriculum and Assessment* (2007) qui met en avant une variété d'épreuves d'évaluation que les enseignants sont invités à mobiliser dans leurs classes. En effet, le guide national définit un continuum d'épreuves d'évaluation (cf. présentation 2 ci-dessous sur ce « *continuum of assessment methods* ») en distinguant différentes natures d'évaluation selon les degrés d'implication et de responsabilisation tant de l'élève que de l'enseignant dans le processus d'évaluation. Il existe huit natures d'évaluation, positionnées à gauche si l'élève est responsable de sa propre évaluation et plus vers la droite lorsque l'enseignant joue un rôle majeur dans l'évaluation de l'élève.

Pour chaque nature d'épreuve, il existe une méthode et des exemples concrets pour une meilleure compréhension du guide par les enseignants, car certaines d'entre elles peuvent être difficilement compréhensibles sans explicitation. Parmi les épreuves d'évaluation dans la classe, on compte ainsi : (1) l'autoévaluation où l'élève évalue ses propres progrès et ses objectifs d'apprentissage en développant des compétences métacognitives ; (2) le « *conferencing* » qui renvoie aux situations d'évaluation périodique entre l'élève et l'enseignant, et de réunions des enseignants avec les parents ou encore des enseignants entre eux (l'élève connaît alors ses forces et faiblesses pour mieux progresser) ; (3) le « *portfolio* » est un moyen de centraliser tous les travaux de l'élève et de suivre ses compétences et ses connaissances acquises pendant une période ; (4) le « *concept mapping* » est un moyen innovant d'appréhender et d'évaluer les acquis méthodologiques de l'élève en lui faisant réaliser des schémas relationnels (ou « *semantic networking* ») sur une idée ou un sujet ; (5) les questions orales par l'enseignant dans la classe permettent d'évaluer spontanément l'élève sur sa compréhension et ses connaissances ; (6) l'observation par les enseignants permet d'avoir un retour immédiat et objectif sur des situations d'apprentissage dans la classe tout au long de l'année et *a fortiori* d'orienter l'élève dans ses tâches d'apprentissage ; (7) les épreuves conçues par les enseignants peuvent prendre la forme de tests écrits ou oraux dans le cadre de l'évaluation formative ; (8) enfin les tests standardisés qui sont une forme d'évaluation externe.

Présentation 2 : Continuum d'épreuves d'évaluation dans la classe recommandé en classe au primaire, Irlande



Source : National Council for Curriculum and Assessment, 2007

À chaque pays son système de notation

Les cadres réglementaires des pays sont aussi très variés quant au système de notation (cf. tableau 3 sur la notation).

**Tableau 3 : Caractéristiques réglementaires des systèmes de notation et d'annotation dans une sélection de pays de l'OCDE
Niveaux CITE 1, 2 et 3 (2010-2013)**

	France	Angleterre, Pays de Galles, Irlande du Nord	Allemagne	Suède	Danemark	Finlande	Québec	Corée
Critères réglementaires	Mentions littérales	CITE 1 : Appréciations littérales, « acquis, non acquis, en voie d'acquisition, non acquis », « oui/non ». CITE 2 : Commentaires des résultats.	CITE 1 : Appréciations littérales (progrès, forces et faiblesses de l'élève). CITE 2 et 3 : Appréciations littérales accompagnant les notes.	CITE 1 : non réglementé.	CITE 1 et 2 : Jusqu'en 7 ^{ème} année, les notes sont remplacées par un bilan qualitatif personnel	CITE 1, 2 : Jusqu'en classe de 5 ^{ème} au collège, les enseignants peuvent utiliser les notes et/ou les appréciations CITE 3 : voir échelle de notes (combinées à des appréciations écrites).	CITE 1, 2 et 3 : Commentaires accompagnant les notes.	Non renseigné.
	Système de notation/ annotations	CITE 1 : la notation peut être utilisée	CITE 1, 2 et 3 : Note de 1 (très bien) à 6 (très insuffisant). Note de 0 à 15 en dernière année CITE 3.	CITE 1 : absence de notation.	CITE 1 et 2 : Pas de notes jusqu'en 8 ^{ème} année	CITE 1 et 2 : Jusqu'en classe de 5 ^{ème} , le choix est laissé aux enseignants d'utiliser une notation chiffrée ou littéraires (ou une combinaison des deux).	CITE 1, 2 et 3 : Notation en pourcentage (sur 100%).	CITE 1, 2 et 3 : Les notes vont de A (plus de 90% de réussite) à E (moins de 60% de réussite).
	Echelle de notes ordinaire ou numérique	CITE 2 et 3 : Pas de réglementation (généralement notes chiffrées 0-10 ou 0-20). Existence de notes en dessous de la moyenne (5/10 ou 10/20). Cite 2 : Expérimentation de classes sans note.	CITE 1, 2 et 3 : Note de 1 (très bien) à 6 (très insuffisant). Note de 0 à 15 en dernière année CITE 3.	CITE 2 et 3 : notation à partir de la 6 ^{ème} année (auparavant 8 ^{ème} année), sous forme d'une échelle de lettres comprenant 6 niveaux (A à F) dont un (F) correspond à un cours non acquis. Les syllabus des cours comprennent une définition très détaillée des compétences requises correspondant aux niveaux A-C-E. Il existe une notation « - » dans le cas d'une impossibilité de notation.	CITE 3 : Notes chiffrées de -3 à 12 à partir de la 8 ^{ème} année.	CITE 2 et 3 : Notes de 4 (moins bonne note) à 10 (meilleure note). La note 4 signifie l'échec.		

Source : Cnesco, 2014

Depuis les années 1970, les cadres réglementaires qui définissent les notations ont fortement évolué.

Ainsi, certains pays comme, dans notre échantillon, la Suède ou le Danemark, ont éliminé les notations chiffrées durant les premières années de la scolarité, notamment au primaire et dans les premières années de leurs collèges⁵. En Finlande, la législation n'impose l'introduction des notes, par les enseignants, qu'à partir de la classe de 4^{ème} de collège. Dans les classes précédentes, le choix leur est laissé entre une évaluation chiffrée, une évaluation strictement qualitative (sans note) ou la combinaison des deux. En pratique, les enseignants finlandais, dans leur très grande majorité, font le choix de ne pas noter numériquement les élèves jusqu'à 13 ans. A partir de 13 ans, légalement les élèves finlandais sont soumis à une notation sur une échelle allant de 4 (signifiant la non-acquisition des connaissances) à 10 (niveau scolaire excellent).

Le fait d'éliminer les notes ne signifie pas que les enseignants sont exempts de contraintes en termes d'évaluation dans la classe. Ainsi, au Danemark, les notes sur les copies ont été remplacées réglementairement par le développement d'un « programme de l'élève » qui, aux côtés d'une analyse détaillée qualitative de ses résultats scolaires et de son développement social et personnel, doit intégrer des objectifs pédagogiques que l'enseignant fixe à l'élève ainsi que les moyens pédagogiques de l'atteindre. **Dans la majorité de ces contextes nationaux et, hors de notre échantillon, dans d'autres États comme la Suisse, des débats vifs, notamment liés au contexte politique, voient le jour régulièrement autour de la pertinence de la réintroduction des notes à l'école** (cf. encadré 3 sur les cas suisse et suédois).

⁵ Dans les pays nordiques, l'école obligatoire constitue un bloc de scolarité unique, dite *école basique*, sans distinction entre le primaire et le collège.

Encadré 3. Les classes sans notes au primaire : un débat récurrent dans certains pays

Le cas de la Suisse et de la Suède

En France si le débat et les expérimentations des classes sans notes au primaire et au collège est d'actualité, il ne l'est pas moins dans un ensemble de pays. C'est le cas de la Suède ou de la Suisse qui sont emblématiques de revirements nombreux sur les notations à l'école.

Ainsi, en Suisse, dans une majorité de cantons romands, le débat autour des notes et de la sélection scolaire est très présent. Dans les années 1990, ces systèmes scolaires très sélectifs, fondés sur une orientation précoce à l'allemande sous forme de filières dès le collège, remettent en cause les pratiques de notation très pointilleuses en cours et repousse l'évaluation plus tardivement dans la scolarité.

La rénovation du système scolaire dans le canton de Genève prévoit le remplacement des notes chiffrées par des appréciations qualitatives concernant les compétences des élèves ainsi que des objectifs à atteindre au primaire sauf pour la dernière année (Gilliéron et Tessaro, 2009).

Cependant dans les années 2000, le débat reprend autour de la pertinence de réintroduire des notes. C'est le canton de Genève qui relance le mouvement du retour aux notes. Des associations de parents d'élèves et des enseignants réclament le retour des notes, symbole de qualité pour l'évaluation des élèves selon ces derniers. En 2007, les notes sont réintroduites à Genève par référendum (76 % des Genevois ont voté pour le retour des notes (République et canton de Genève (2006)).

Le canton suisse de Vaud a connu une histoire similaire. Jusqu'en 1997, les élèves étaient évalués selon un système de notation quantitatif très fin. Sur la base de leur moyenne générale, les élèves en fin de cursus primaire (fin de 5^{ème} année en Suisse correspondant à l'âge de 11 ans) se voyaient orientés dans certaines filières du secondaire. Une réforme est instituée alors pour mettre fin au stress généré par l'évaluation, invoqué notamment par les parlementaires et les parents (EVM pour École Vaudoise en Mutation, Blanchet, 2005) ; les notes chiffrées et les moyennes par discipline sont supprimées pour l'ensemble du primaire d'abord, puis progressivement jusqu'à la 9^{ème} année et remplacées par une échelle ne comprenant que cinq niveaux de compétences (Gilliéron et Tessaro, 2009). Or une enquête menée en 2003 montre que les enseignants considèrent que « les notes et les moyennes offrent une meilleure protection de l'enseignant face aux contestations des parents » (CERU, 2010). Après un long processus de négociation, notamment entre politiques, enseignants et parents, une solution a été trouvée en 2005, les notes chiffrées et les calculs des moyennes par discipline sont rétablis à partir de la 5^{ème} année. En 2011, une initiative « Ecole 2010 » qui soutenait la réintroduction des notes dès la 1^{ère} année du primaire ne sera pas rejetée par les citoyens.

En Suède, le débat autour de l'évaluation chiffrée est aussi vif, et la réglementation a évolué ces dernières années non sans lien avec les alternances politiques. En effet, l'ancien système de notation qui prévalait lors de la présidence du parti social-démocrate depuis 1994 était un système comprenant un ensemble limité de mentions : insuffisant (IG), passable (G), bien (VG) et très bien (MVG). Quelques années après, lors de l'arrivée au pouvoir, en 2006, de l'alliance des partis de centre-droit, la réflexion sur un projet de réforme de l'École (*Skollagen*) aboutit, en 2011, à un nouveau système de notation selon une échelle élargie de A à F, E étant la note limite de passage et F insuffisant. Par ailleurs, la notation chiffrée qui ne commençait qu'en 8^{ème} année - l'équivalent de notre 4^{ème} -, est réintroduite dès la 6^{ème} année de l'enseignement obligatoire, qui correspond à la fin du primaire français (Ministère de l'éducation et de la recherche suédois, 2014).

Dans les pays qui ont conservé la notation, les systèmes d'évaluation sont très variables selon les pays et, même, au sein de ces pays selon les niveaux d'enseignement, primaire ou secondaire, et selon les voies d'enseignement, notamment générale ou professionnelle, dans le secondaire supérieur, l'équivalent de notre lycée.

D'un pays à l'autre, les échelles de notation qui peuvent être soit numériques, comme les notes en France, soit ordinales (par niveau, par exemple sous forme de lettres), diffèrent, et le nombre de notes pouvant représenter l'échec peut aussi varier.

Certains pays sont, tout d'abord, restés fidèles à la notation selon une échelle numérique. Ainsi, en Allemagne, alors qu'au primaire et au secondaire inférieur (niveau collège), les enseignants sont tenus d'évaluer les élèves avec une échelle de 1 (très bien) à 6 (très insuffisant), au lycée l'échelle de notation diffère selon les voies d'enseignement : au *gymnasium*, de 0 à 15 ; dans les voies professionnelles, de 1 à 6. En Finlande, de même, l'échelle est moins fine dans la voie professionnelle où trois niveaux de résultats scolaires sont repérés seulement, alors que, durant tout le cursus finlandais, l'échelle s'étire de 4 à 10. Les pays d'Amérique du Nord - les États-Unis et le Québec dans notre échantillon - ainsi que la Corée utilisent des échelles beaucoup plus amples sur une base de 100.

D'autres pays n'ont plus recours à l'échelle numérique continue, mais à une échelle présentant plusieurs niveaux ordonnés, le plus souvent sous forme de lettres. Dans la majorité des cas, la rupture avec l'échelle numérique a incité les pays à définir des réglementations plus détaillées décrivant les connaissances/compétences équivalant à chaque niveau de performances. Le passage à une échelle non numérique marque donc le plus souvent un encadrement plus resserré de l'activité évaluative des enseignants dans la classe. Sur ces échelles ordinales, le nombre de niveaux de performance est souvent l'objet d'importants débats publics et un motif de réforme de l'évaluation. Ainsi, en Angleterre, l'échelle de A à G (complétée de la note U représentant l'échec à un cours) a été étendue en 2011 pour inclure un niveau d'excellence A* censé mettre un frein à l'« inflation des notes » ; le nombre croissant de notes élevées ne permettant pas aux universités les plus sélectives d'assurer un recrutement fiable des étudiants talentueux. Récemment également l'échelle de notation en Suède a été étendue de A à F, F représentant l'échec dans un cours.

La précision dans l'étendue des notes ou niveaux qui représentent l'échec à un cours varie aussi selon les pays. Les échelles numériques permettent souvent une appréciation plus fine du caractère non acquis d'un cours, tandis que les échelles ordinales bien souvent le limitent à une seule lettre, comme en Angleterre ou en Suède. Peu de pays prévoient comme la Suède d'attribuer un symbole spécifique et donc de reconnaître les cas où les matériaux d'évaluation seraient trop peu nombreux pour juger des résultats des élèves. Sur le bulletin de l'élève suédois cette situation est notée du symbole « - ».

Au total, les pays présentent des réglementations en termes d'échelle de notation très variées, pour ne pas dire singulières. Si les échelles numériques continues (la notation chiffrée) demeurent le lot commun, un nombre non négligeable de pays ont évolué vers des échelles ordinales (lettres...) dont l'amplitude est source de débat et de réformes successives qui reflètent souvent les visions philosophiques et politiques de l'école. Les projets politiques d'école compétitive et élitiste sont davantage attachés, notamment dans les filières d'élite, à conserver des échelles numériques à l'amplitude étendue, justifiée notamment par la nécessité d'identifier et de différencier finement les résultats des élèves de l'élite scolaire.

On retrouve aussi des visions de l'école différentes au travers des réglementations relatives aux usages des notes, dans le cadre de la communication des résultats scolaires aux parents, mais aussi au sein des classes et des établissements, avec notamment le jeu des classements et palmarès d'élèves.

C. Des usages légaux de l'évaluation multiples

La majorité des pays de l'OCDE réglemente les modalités de communication des résultats scolaires aux parents, en imposant à la fois des fréquences minimales de rencontres enseignants/parents ainsi que des supports codifiés des bilans d'apprentissage. Dans tous les pays étudiés, les parents sont tenus au courant des acquis des apprentissages de leurs enfants via des bulletins au moins deux à trois fois par an (cf. tableau 4).

Tableau 4. Réglementations relatives à certains usages des évaluations par les enseignants dans la classe dans une sélection de pays de l'OCDE
Niveaux CITE 1, 2 et 3 (2010-2013)

Critères réglementaires	France	Angleterre, Pays de Galles, Irlande du Nord	Allemagne	Suède	Donemark	Finlande	Québec	Corée
Fréquence des rencontres parents/enseignants	CITE 1, 2 et 3 : Au moins trois fois par an ; à la demande des parents.	Non réglementé mais rencontres complémentaires à la demande des parents	CITE 2 : Rencontres deux fois par an (milieu et fin d'année) CITE 1 et 3 : Non renseigné	CITE 1, 2 : Entretiens individuels parents/élèves au moins une fois par trimestre CITE 3 : idem La réglementation insiste sur l'échange avec l'étudiant sur l'évaluation mais précise que les notes ne font pas l'objet d'un appel.	CITE 1 et 2 : Deux fois par an CITE 3 : Aucune réglementation	CITE 1, 2 : Entretiens d'évaluation formel avec les parents ; à la demande des parents.	CITE 1, 2 et 3 : Au moins trois bulletins scolaires avec les notes en pourcentage, commentaires	CITE 1, 2 et 3 : Selon la politique des établissements
	Communication avec les parents	CITE 1 : Livret scolaire remis aux parents. Dans le premier degré, bulletins périodiques (cinq au total) CITE 1 et 2 : Livret personnel de compétences, bulletins trimestriels. CITE 3 : Bulletins trimestriels. Livret scolaire présenté au jury de l'examen final.	Non renseigné.	CITE 1 : lors deuxières années, rapport général fin d'année, rapports trimestriels. CITE 2 : Bulletins semestriels. CITE 1 et 3 : Non renseigné.	Non réglementé.	CITE 1 et 2 : Le livret de suivi des élèves « Student plan » compile les résultats des évaluations continues au moins deux fois par an.	CITE 1, 2 : Attestation d'apprentissage à la fin des cycles. CITE 1, 2 et 3 : le bulletin demeure un outil de communication cyclique entre l'école et les parents. Combiné à une information régulière, par exemple, au moyen de rencontres avec les parents, de portfolios, de feuilles de route, d'annotations sur les travaux, de l'agenda scolaire ou d'autres documents.	CITE 1, 2 et 3 : le bulletin demeure un outil de communication cyclique entre l'école et les parents. Combiné à une information régulière, par exemple, au moyen de rencontres avec les parents, de portfolios, de feuilles de route, d'annotations sur les travaux, de l'agenda scolaire ou d'autres documents.
Classement	Pas de réglementation	Pas de réglementation	CITE 1 (à partir du grade 3) : comparaison des résultats élèves par rapport moyenne de classe. CITE 2 : Redoublement possible. Avoir 4 dans chaque matière assure de passer dans la classe supérieure.	CITE 2 et 3 : interdiction formelle d'une comparaison des notes entre élèves.	Non réglementé.	Non réglementé.	Non réglementé.	CITE 1, 2 et 3 : Classement des élèves dans chaque bulletin scolaire.
	Redoublement	CITE 1 : Avis du conseil des maîtres, CITE 2 et 3 : Avis du conseil de classe avec prééminence du chef d'établissement.	CITE 1, 2 et 3 : Pas de réglementation centrale, autonomie complète du chef d'établissement (avec consultation des parents de l'élève).	Il n'existe pas un cadre de notation particulier pour le redoublement.	CITE 1 : autonomie totale de l'établissement. CITE 2 et 3 : sur proposition du conseil de classe.	CITE 1 : redoublement envisagé pour résultats inférieur ou égal à 4/10. CITE 2 et 3 : appréciation globale à la discrétion de l'établissement.	Le seuil de réussite est fixé à 60% pour le passage au cycle supérieur.	Non réglementé

Source : Cnesco, 2014

L'évaluation, notamment quand elle est chiffrée, est aussi mobilisée dans des formes différentes de compétition à l'école ou du moins de comparaison entre élèves. C'est le cas, par exemple, en Allemagne où, dans la majorité des *Länder*, dès la troisième année du primaire, sur le bulletin scolaire, aux côtés de la note de l'élève, doivent figurer des éléments d'appréciation des résultats de la classe pour permettre à l'élève de se positionner. Au contraire, en Suède, toute comparaison entre les résultats scolaires des élèves est formellement interdite par les textes.

Les notes sont aussi souvent utilisées dans les décisions de redoublement ou de promotion dans la classe supérieure sur des bases extrêmement précises, comme en Allemagne (cf. rapport Cnesco 2014 sur le redoublement).

En conclusion, cette première analyse de la réglementation met en évidence une progression des contraintes légales encadrant l'évaluation depuis trois décennies. Ainsi, en matière de notation, si les réglementations peuvent évoluer dans le temps - révélant que la notation chiffrée est un enjeu de société qui dépasse largement l'école et s'invite régulièrement dans le débat public -, des critères de plus en plus précis sont mis en place. Les invitations à la collaboration entre professeurs et au développement de nouvelles évaluations innovantes se font aussi de plus en plus pressantes et commencent à conditionner partiellement les pratiques enseignantes, comme en attestent les résultats des enquêtes internationales que nous analysons dans la seconde partie de ce rapport.

II. Des pratiques d'évaluation différenciées dans les pays de l'OCDE

Après avoir comparé, dans une première partie, les cadres réglementaires régissant l'évaluation interne des élèves par les enseignants, dans un échantillon restreint de pays de l'OCDE, dans cette seconde partie, nous mobilisons différentes enquêtes internationales de l'OCDE (PISA 2009, TALIS 2013) et de l'IEA⁶ (PIRLS 2012) pour appréhender empiriquement, au-delà des cadres légaux, la réalité des pratiques enseignantes en matière d'évaluation dans leur classe et dans leur établissement (au moins telles qu'elles sont déclarées par les principaux intéressés). Ces enquêtes nous informent principalement sur les pratiques des enseignants intervenant en collège et très ponctuellement sur celles des enseignants en primaire et au lycée, qui sont à ce jour sous-explorées dans les comparaisons internationales.

Dans la majorité des pays de l'OCDE, l'évaluation des élèves par les enseignants présente des formes et des organisations variées. Il existe, ainsi, des différences dans la nature des évaluations, dans leurs fréquences ainsi que dans les pratiques de coopération, qui les sous-tendent, au sein des équipes pédagogiques d'un même établissement.

Dans une première section, nous exposerons la variété des évaluations qu'utilisent les enseignants dans leur classe et la fréquence de leur administration, en France et dans certains pays de l'OCDE. Dans une deuxième section, nous étudierons les pratiques de coopération entre enseignants dans les établissements en matière d'évaluation. Enfin, une troisième section sera consacrée aux besoins en formation continue en matière d'évaluation interne des élèves déclarés par les enseignants.

A. La variété des évaluations mobilisées par les enseignants dans la classe

Une des fonctions majeures de l'évaluation formative est de permettre à l'enseignant de suivre, avec ses élèves, leurs progrès scolaires par divers moyens et à fréquence régulière, en interaction avec leurs parents.

L'enquête internationale sur *L'enseignement et l'apprentissage* conduite par l'OCDE (TALIS 2013, cf. encadré 4) à laquelle la France a participé pour son cycle le plus récent, permet d'appréhender la variété des modalités d'évaluation mobilisées par les enseignants, sur la base de leur déclaration. Elle porte exclusivement sur les enseignants intervenant dans l'équivalent de notre collège.

⁶ The International Association for the Evaluation of Educational Achievement (IEA) (<http://www.iea.nl/>).

Encadré 4. L'enquête internationale sur l'enseignement et l'apprentissage 2013 (TALIS, 2013)

L'enquête internationale sur *L'enseignement et l'apprentissage 2013* (TALIS, 2013) de l'OCDE s'intéresse aux conditions de travail et à l'environnement pédagogique des enseignants dans un certain nombre de pays de l'OCDE incluant la France, la Finlande, l'Angleterre, l'Australie ou encore le Japon. TALIS met en lumière les convictions des enseignants quant aux pratiques pédagogiques efficaces et apporte des éclairages sur leurs pratiques d'évaluation dans la classe.

Quels sont les professionnels de l'éducation interrogés dans TALIS ? La population visée par l'enquête TALIS est celle des enseignants et chefs d'établissement des établissements, publics et privés, de l'enseignement ordinaire du premier cycle du secondaire. Dans chaque pays participant, un échantillon représentatif de 200 établissements et, dans chacun d'entre eux, de 20 enseignants et d'un chef d'établissement, a été sélectionné pour les besoins de l'étude. Plus de 4 millions d'enseignants répartis dans plus de 30 pays et économies participants ont répondu à l'enquête pour le premier cycle du secondaire.

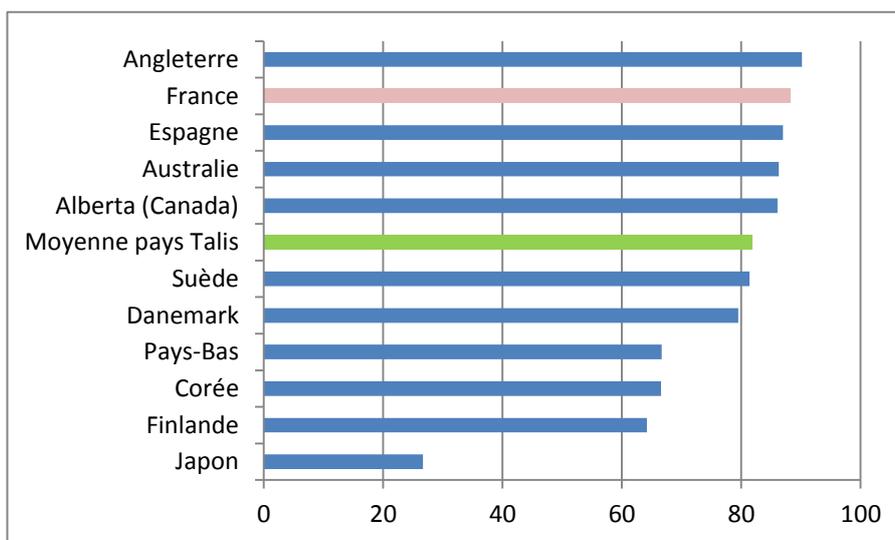
En France, 3 002 enseignants et 204 chefs d'établissement de 204 établissements du premier cycle du secondaire, c'est-à-dire intervenant dans un collège, ont répondu à l'enquête.

Cette enquête procède par questionnaire et appréhende les pratiques déclarées par les enseignants ou les chefs d'établissement, ce qui est différent d'une enquête conduite par observation dans les classes, qui permettrait d'appréhender au plus près la réalité de ces pratiques mais se révélerait très lourde à conduire dans le cadre d'une comparaison internationale.

L'enquête TALIS 2013 met tout d'abord en évidence la variété des évaluations pratiquées par les enseignants, telles que ces derniers les déclarent. C'est particulièrement le cas de la France qui se situe au-dessus de la moyenne des pays de TALIS⁷ en la matière (cf. figure 1) : près de 90 % des enseignants français déclarent mobiliser une diversité d'évaluations de leurs élèves.

⁷ L'échantillon TALIS 2013 ne comprend pas tous les pays de l'OCDE. Aussi dans la suite du texte et notamment dans les graphiques est mentionnée non pas la moyenne OCDE mais la « moyenne des pays TALIS ».

Figure 1. Pourcentage d'enseignants du premier cycle du secondaire déclarant utiliser diverses modalités d'évaluation des élèves



Source : TALIS, 2013

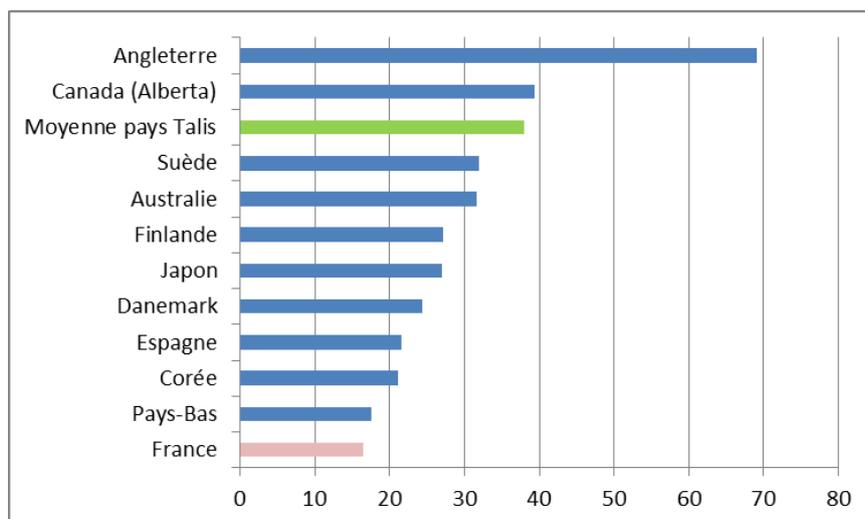
Une analyse plus fine par type d'évaluation (auto-évaluation des élèves, observations du travail de l'élève en direct par l'enseignant, interrogation à l'oral...) permet d'affiner l'analyse des pratiques et révèle que certaines modalités d'évaluation sont encore peu courantes dans les classes françaises.

Les figures des pages suivantes présentent les proportions moyennes d'enseignants du premier cycle du secondaire qui déclarent utiliser « souvent » ou « à chaque séance » différentes méthodes pour évaluer les élèves dans leur classe. Les enseignants déclarent souvent en utiliser plusieurs.

L'auto-évaluation des élèves : une pratique peu développée en France

L'auto-évaluation, pratique en cours de développement dans certains pays de l'OCDE (Angleterre, Québec...), fait ainsi partie des modes d'évaluation encore peu mobilisés par les enseignants français. En moyenne, dans l'échantillon de pays TALIS, 38 % des enseignants du premier cycle du secondaire déclarent demander régulièrement à leurs élèves de s'évaluer eux-mêmes (voir figure 2). Les disparités dans l'usage de cette modalité d'évaluation entre les pays sont importantes puisque l'on constate que les enseignants anglais sont près de 69 % à déclarer utiliser l'auto-évaluation contre 16.5 % des enseignants français. **Parmi tous les pays de l'échantillon TALIS, la France est le pays dans lequel les enseignants sont les moins nombreux à laisser les élèves s'auto-évaluer. L'existence d'une réglementation sur le sujet dans d'autres pays, comme l'Angleterre, explique ainsi que les enseignants, dans ces autres contextes nationaux, utilisent plus souvent cette pratique pédagogique.**

Figure 2. Pourcentage d’enseignants du premier cycle du secondaire déclarant demander à leurs élèves de s’auto-évaluer « souvent » ou « à chaque séance »



Source : TALIS, 2013

Interroger à l’oral les élèves devant la classe : une modalité d’évaluation répandue en France

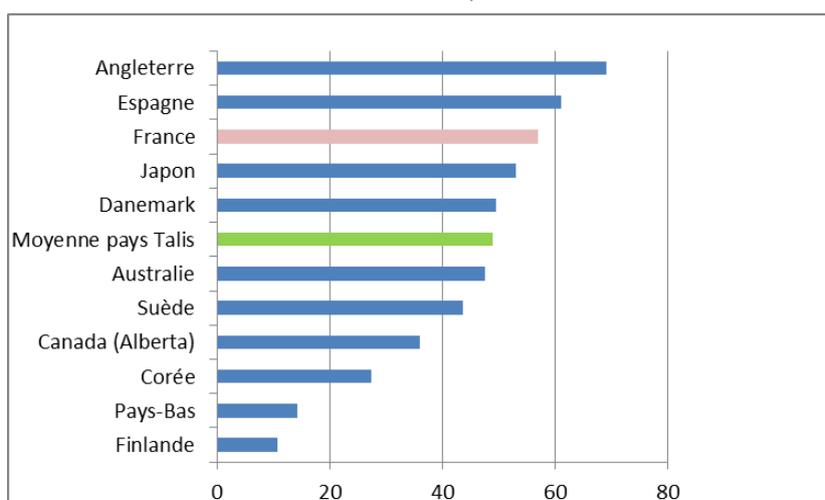
A l’opposé, pratiquer l’interrogation des élèves à l’oral, devant les autres élèves de la classe, est une pratique plus banalisée en France que dans les autres pays de l’enquête.

En moyenne, dans l’échantillon des pays TALIS, 50 % des enseignants du premier cycle du secondaire déclarent faire répondre régulièrement les élèves à des questions devant la classe (voir figure 3).

En France, les enseignants sont 57 % à mobiliser cette pratique.

En Angleterre, les enseignants sont plus nombreux à pratiquer cette modalité d’évaluation avec un taux de 70 %. En revanche, ils sont seulement 11 % à le déclarer en Finlande.

Figure 3. Pourcentage d’enseignants du premier cycle du secondaire déclarant interroger les élèves devant la classe « souvent » ou « à chaque séance »



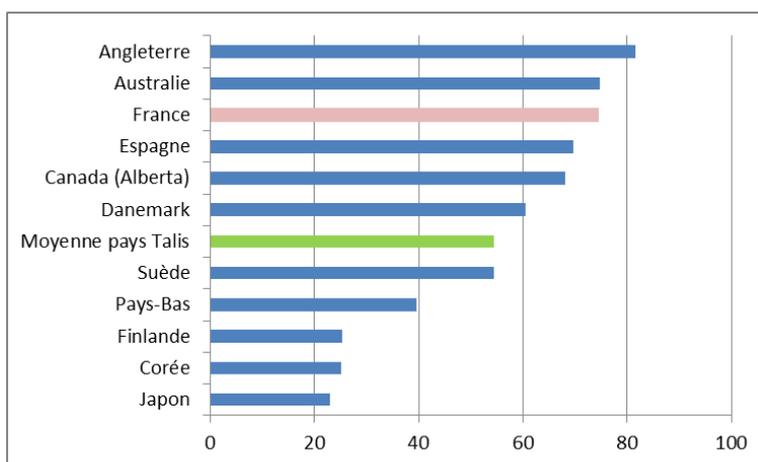
Source : TALIS, 2013

Ajouter un commentaire écrit à la note chiffrée : une pratique courante chez les enseignants français

Autre pratique répandue en France, la notation chiffrée est le plus souvent (mais pas toujours) accompagnée d'une analyse qualitative de l'évaluation de l'élève.

Ainsi, en moyenne, dans l'échantillon de pays TALIS (voir figure 4), 54 % des enseignants du premier cycle du secondaire déclarent ajouter régulièrement un commentaire écrit à la note chiffrée ou à l'appréciation des travaux des élèves, alors qu'en France, les enseignants sont près de 75 % à le déclarer. Les disparités de réponses sont importantes selon les pays. Alors que plus de 80 % des enseignants en Angleterre déclarent ajouter un tel commentaire à la note ou à l'appréciation attribuée à l'élève, ils sont seulement 25 % en Corée et 23 % au Japon. Au Canada (Alberta), la proportion est de 68 % et de 61 % en Belgique flamande. Ces fortes disparités entre pays pourraient s'expliquer par le fait que, comme nous l'avons montré dans la première partie de cette analyse, les cadres réglementaires nationaux en matière d'évaluation des élèves imposent rarement cette évaluation qualitative, au niveau du collège, à l'exception des cas où la notation quantitative n'est pas autorisée.

Figure 4. Pourcentage d'enseignants du premier cycle du secondaire déclarant ajouter régulièrement un commentaire écrit à la note chiffrée



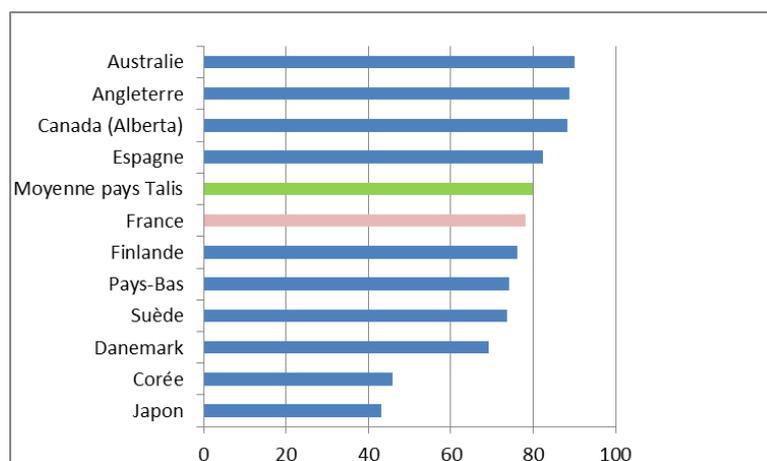
Source : TALIS, 2013

L'observation des élèves pendant qu'ils effectuent en classe une tâche particulière et le commentaire immédiat de leur travail

L'évaluation rapprochée par un suivi d'observation de l'élève est aussi très répandue dans les classes françaises, même si les enseignants de certains autres pays de l'OCDE la pratiquent plus couramment. En moyenne, dans l'échantillon de pays TALIS (voir figure 5), 80 % des enseignants du premier cycle du secondaire déclarent observer régulièrement les élèves pendant qu'ils effectuent une tâche particulière en classe afin de leur apporter un retour évaluatif immédiat.

La proportion d'enseignants déclarant observer les élèves pendant qu'ils effectuent en classe une tâche particulière en leur faisant un commentaire varie d'un pays à l'autre : s'ils sont près de 78 % en France, un niveau proche de la moyenne des pays de TALIS, la pratique apparaît plus banalisée encore dans les pays anglo-saxons de l'échantillon (Angleterre, Canada, Australie).

Figure 5. Pourcentage d’enseignants du premier cycle du secondaire déclarant observer les élèves pendant qu’ils effectuent en classe une tâche particulière et apporter un commentaire immédiat sur leur travail « souvent » ou « à chaque séance »

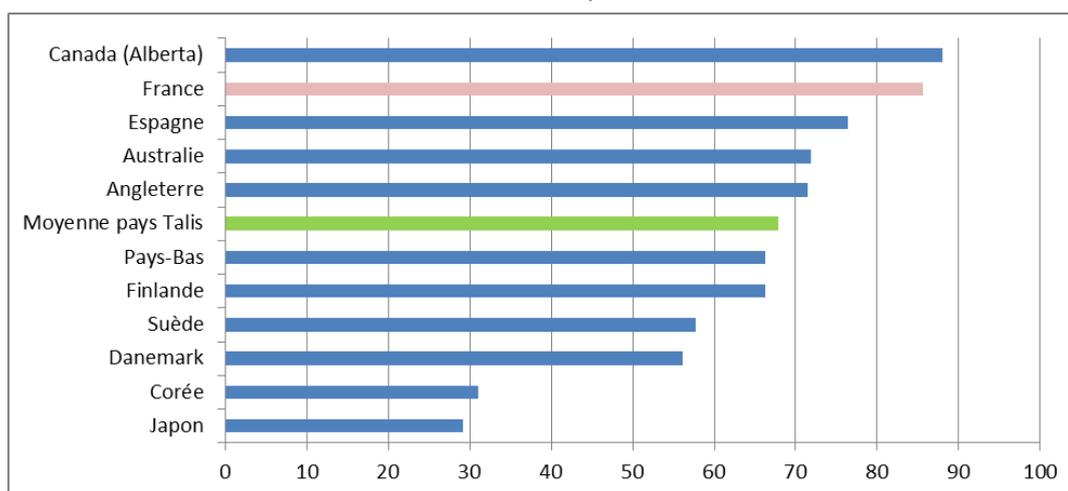


Source : TALIS, 2013

L’élaboration de contrôles écrits par l’enseignant lui-même est au cœur des pratiques d’évaluation des enseignants français.

En moyenne, dans l’échantillon de pays TALIS (voir figure 6), 68 % des enseignants du premier cycle du secondaire déclarent élaborer régulièrement leurs propres épreuves pour évaluer les acquis des élèves dans la classe. Les données témoignent d’une autonomie des enseignants français en la matière puisque près de 86 % d’entre eux déclarent élaborer leurs propres tests. En Angleterre, en Espagne ou encore en Australie, la majorité des enseignants déclarent également élaborer leurs propres épreuves. A l’opposé, en Corée ou au Japon, les enseignants sont seulement 30 % à déclarer souvent élaborer eux-mêmes les épreuves de leurs élèves.

Figure 6. Pourcentage d’enseignants du premier cycle du secondaire déclarant réaliser leurs propres contrôles écrits « souvent » ou « à chaque séance »



Source : TALIS, 2013

La fréquence des évaluations dans certains pays de l'OCDE

Au-delà de la variété des modalités d'évaluation, l'enquête PISA 2009 (OCDE, voir encadré 5 ci-dessus) permet d'appréhender les fréquences de certaines formes d'évaluation mobilisées par les enseignants des élèves âgés de 15 ans, c'est-à-dire dans le cas français, des enseignants intervenant au lycée ou au collège (pour les élèves ayant redoublé). Dans PISA, ces pratiques sont appréhendées indirectement à travers un questionnaire administré au chef d'établissement. Malgré l'absence de la France dans cette partie de l'enquête - la France n'administrait pas encore pour PISA 2009 le questionnaire « chef d'établissement » - l'intérêt de ces données réside dans le fait d'apporter quelques repères internationaux sur la fréquence d'utilisation de certaines modalités d'évaluation.

Encadré 5. L'étude PISA 2009

Le Programme International pour le Suivi des Acquis des élèves (PISA) est une évaluation internationale menée par l'Organisation de coopération et de développement économiques (OCDE), qui vise à tester les compétences des élèves de 15 ans. L'enquête est menée au niveau international de manière cyclique tous les trois ans en compréhension de l'écrit et cultures scientifiques et mathématiques. En 2009, PISA a entamé son deuxième cycle en testant de nouveau en priorité les compétences des élèves en compréhension de l'écrit. Fondées sur les déclarations des chefs d'établissement en ce qui concerne les pratiques pédagogiques enseignantes, PISA 2009 nous renseigne notamment sur les pratiques d'évaluations des enseignants dans la classe. En particulier, nous nous intéressons ici au pourcentage d'élèves scolarisés dans un établissement pratiquant certaines formes d'épreuves (épreuves établies par les enseignants, portfolios, etc.) dans l'ensemble large des pays de l'OCDE ayant administré le questionnaire. Bien que la France fasse partie de l'échantillon de pays participant à l'étude PISA 2009, le questionnaire « chef d'établissement » n'était pas encore administré dans cette enquête. (PISA, 2009)

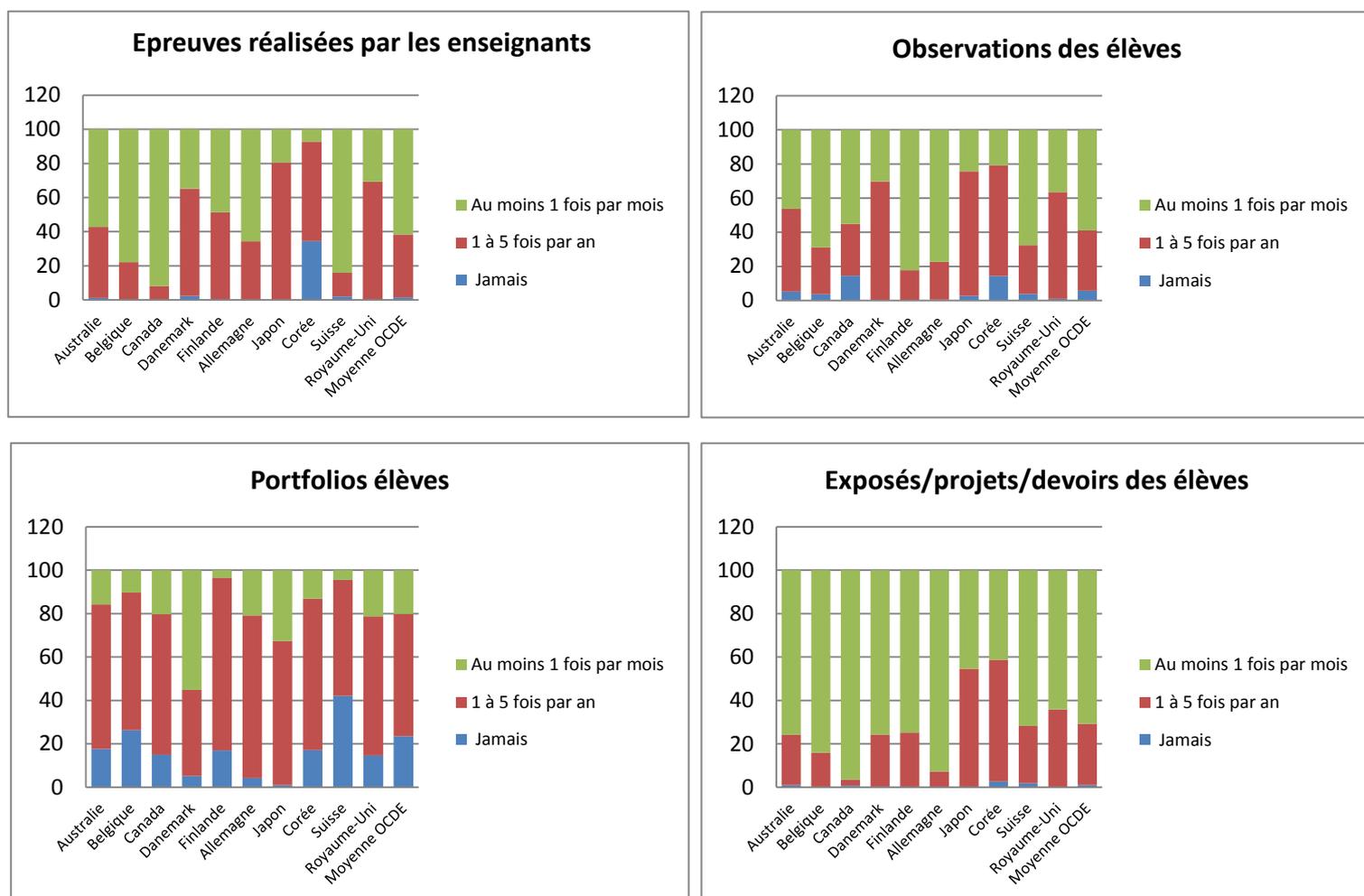
Les données PISA, présentées en figure 7, permettent d'analyser quatre modalités d'évaluation : les contrôles réalisés par les enseignants eux-mêmes, l'évaluation par observations de l'élève, la constitution de portfolios par les élèves⁹ ainsi que les exposés, projets ou devoirs des élèves qui peuvent impliquer une part de travail personnel en dehors des cours.

La figure 7 montre tout d'abord de fortes disparités dans la fréquence d'administration des **épreuves réalisées par les enseignants eux-mêmes**. Si en moyenne dans l'OCDE, 40 % des élèves apparaissent scolarisés dans un établissement qui pratique ce type d'évaluation de une à cinq fois par an, ces contrôles sont beaucoup plus fréquents dans certains pays. Ainsi, au Canada, en Belgique et en

⁹ Le portfolio de l'élève est défini par un ensemble de réalisations scolaires et de travaux qui attestent de ses compétences et de ses acquis. Le portfolio permet de garder les traces des réalisations de l'élève de façon continue pour une meilleure évaluation formative.

Allemagne 75 % des élèves sont scolarisés dans un établissement où les enseignants effectuent des contrôles écrits au moins une fois par mois.

Figure 7. Pourcentage d'élèves de 15 ans scolarisés dans un établissement pratiquant les évaluations suivantes : (résultats fondés sur les déclarations du chef d'établissement)



Source : PISA 2009, OCDE 2010

Les données PISA 2009 mettent aussi en évidence que la constitution de **portfolios** par les élèves est une pratique innovante qui tend à se développer : en moyenne dans l'OCDE, seul un cinquième des élèves est scolarisé dans un établissement qui ne pratique pas ce type d'évaluation. Cependant, la fréquence de cette pratique demeure faible.

Les évaluations comme les **exposés, les devoirs à la maison ou les projets** qui peuvent impliquer un travail à la maison sont eux pratiqués sur une base extrêmement fréquente dans les pays de l'OCDE.

Les données sur l'évaluation par l'**observation des élèves** révèlent sa banalisation mais aussi de fortes disparités dans sa fréquence d'utilisation, selon les pays de l'OCDE. Si 60 % des élèves sont scolarisés dans un établissement la pratiquant régulièrement - au moins une fois par mois - la part des élèves

scolarisés dans des établissements où cette pratique est ponctuelle - 40 % en moyenne dans l'OCDE - n'est pas négligeable.

Au total, si dans les pays de l'OCDE et en France, les enquêtes internationales mettent bien en évidence une variété de modalités d'évaluation, dont certaines innovantes ont bien pénétré les classes comme l'auto-évaluation des élèves ou les portfolios, l'étude de leur fréquence permet d'affiner l'analyse. Elle révèle des profils de pays plus avancés que d'autres dans l'évolution des pratiques en matière d'évaluation - les pratiques innovantes y sont plus régulières et plus répandues. L'analyse des fréquences des évaluations met également en évidence une intensité variable dans l'évaluation interne des élèves qui doit certainement aussi être analysée en relation avec les évaluations externes qui se sont très fortement développées dans certains pays, notamment depuis le début des années 1990 (Mons, 2007). Ainsi, la proportion élevée d'élèves anglais scolarisés dans des établissements où les enseignants administrent leurs propres épreuves sur une fréquence relativement faible - c'est-à-dire une à cinq fois par an - doit être rapprochée de l'utilisation très intensive de tests standardisés externes à tous les niveaux de la carrière scolaire des jeunes anglais. Des effets de substitution peuvent donc voir le jour entre les épreuves internes qui ressortent de l'initiative des enseignants et les tests standardisés externes.

Il ressort de ces constats que l'évaluation n'est plus seulement l'affaire d'un enseignant isolé dans sa classe, mais, dans de nombreux pays de l'OCDE, un sujet d'échange au sein de l'équipe pédagogique. Nous étudions dans la partie suivante les pratiques en termes de collaboration entre enseignants autour de l'évaluation des élèves dans l'établissement.

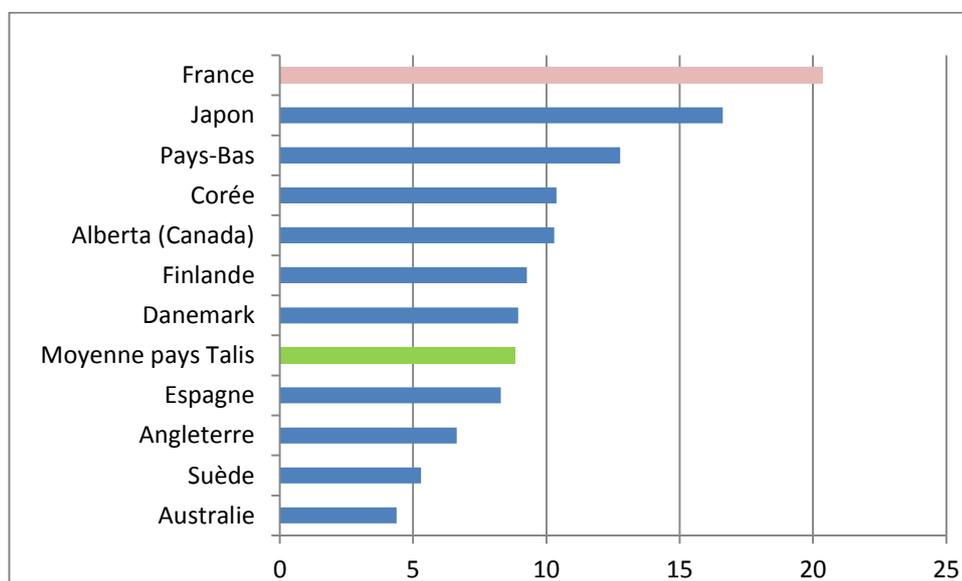
B. Les collaborations entre enseignants autour de l'évaluation des élèves dans l'établissement

De nombreuses collaborations se développent entre enseignants autour de l'évaluation des élèves au sein des établissements. L'enquête TALIS a plus particulièrement interrogé les enseignants sur des échanges autour de barèmes communs servant à évaluer les progrès des élèves.

Les données TALIS 2013 (cf. figure 8) montrent ainsi que seuls 9 % des enseignants des pays concernés par l'étude déclarent ne *jamais* coopérer avec d'autres enseignants dans l'établissement pour appliquer ces harmonisations d'évaluation.

En France, aussi, les pratiques de collaboration apparaissent fréquentes dans ce domaine. Néanmoins, 20 % des enseignants au collège qui déclarent ne jamais coopérer pour l'instauration de barèmes communs, la France est le pays où les enseignants du premier cycle du secondaire sont les plus nombreux à déclarer ne pas coopérer en la matière. A l'opposé, en Australie, en Suède ou en Angleterre, seuls quelques 5 % des enseignants déclarent ne pas collaborer en matière d'évaluation.

Figure 8. Pourcentage d'enseignants du premier cycle du secondaire qui déclarent ne jamais collaborer avec d'autres enseignants pour appliquer des barèmes communs pour évaluer les progrès des élèves

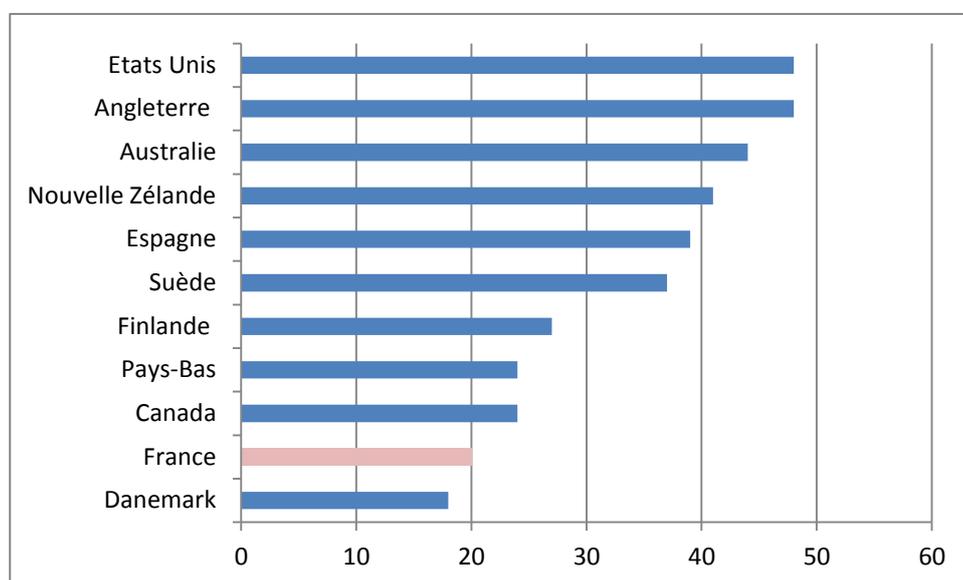


Source : TALIS, 2013

Ce constat d'une faible collaboration pédagogique entre enseignants français au sein de leur établissement dépasse le cadre de l'évaluation. Ainsi, par exemple, 32 % des enseignants français du premier cycle du secondaire déclarent aussi ne jamais assister à des conférences d'équipe.

Ce constat d'une collaboration limitée entre enseignants français peut aussi être posé au primaire, malgré des données plus indirectement liées à l'évaluation dans la classe. Interrogés, dans l'enquête PIRLS 2012 (voir encadré 6) sur différentes dimensions de collaboration au sein de l'équipe pédagogique, les enseignants français qui déclarent un très haut niveau de coopération au sein de leur équipe (cf. figure 9) apparaissent peu nombreux (quelque 20 %) au regard de pays, notamment de la sphère anglo-saxonne, dans lesquels près de 50 % des enseignants déclarent de telles coopérations.

Figure 9. Pourcentage d'élèves du premier degré dont les enseignants déclarent un très haut niveau de collaboration avec les autres enseignants de leur établissement



Source : PIRLS, 2012

Note pour la lecture : Les enseignants dans l'enquête PIRLS 2012 ont été interrogés sur cinq dimensions possibles de collaboration au sein de l'équipe pédagogique de leur établissement (approche d'une discipline scolaire, préparation et planification du matériel pédagogique, observation d'une autre classe pour apprendre davantage sur l'enseignement...) ainsi que sur la fréquence de ces interactions (jamais, une à trois fois par mois, une à trois fois par semaine, quotidienne). Dans notre figure, les élèves dont les enseignants sont ici considérés comme très coopératifs sont ceux qui collaborent avec leurs collègues en moyenne au moins une à trois fois par semaine sur trois des cinq dimensions de collaboration de l'enquête et au moins deux à trois fois par mois sur chacune des deux autres dimensions.

L'ensemble de ces pratiques ne sont pas sans lien avec les besoins de formation continue en évaluation, exprimés par les enseignants, que nous analysons dans la section suivante.

Encadré 6. L'étude PIRLS 2012

Lancée par l'Association internationale pour l'évaluation des acquis scolaires (IEA), PIRLS (Progress in International Reading Literacy - Programme international de recherche en lecture scolaire) est un programme qui évalue la lecture des élèves de quatrième année au primaire (correspondant au CM1 pour la France) dans différents pays du monde (OCDE et pays du Sud). En 2012, 45 pays, dont la France, ont participé à l'enquête internationale. (PIRLS, 2012)

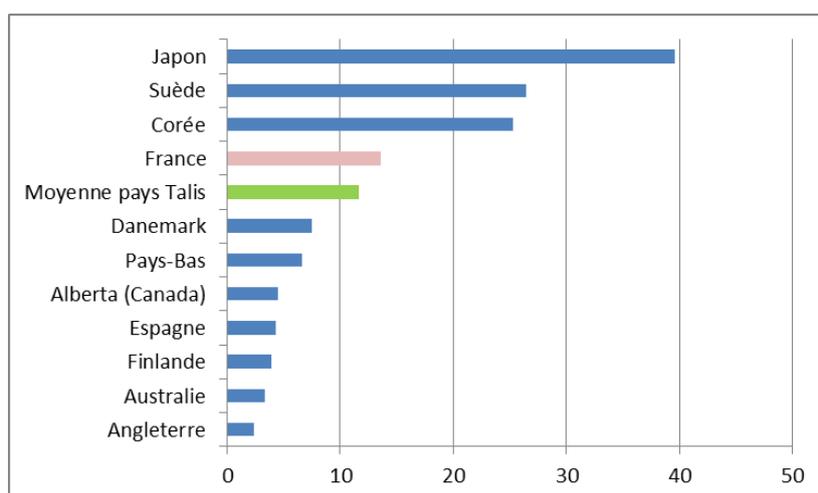
C. Le besoin de formation continue des enseignants en matière d'évaluation des élèves

Dans l'étude TALIS 2013, les besoins de formation continue sont examinés, notamment en matière d'évaluation des élèves dans la classe.

En moyenne, dans l'échantillon de pays TALIS, 12 % des enseignants du premier cycle du secondaire déclarent avoir d'importants besoins de formation continue en matière d'évaluation interne des élèves.

La figure 10 montre qu'en France, la proportion d'enseignants du premier cycle du secondaire déclarant un besoin important de formation continue en matière de pratiques d'évaluation des élèves est légèrement au-dessus de la moyenne des pays de TALIS (13.6 %). Ce taux s'élève à près de 25 % pour la Corée et la Suède. En revanche, ce besoin déclaré de formation ne dépasse pas 5 % en Angleterre, en Finlande ou en Australie. Le développement dans ces pays d'une politique centrée sur l'évaluation dans la classe n'est certainement pas étranger à ce résultat.

Figure 10. Besoins des enseignants en matière de formation continue dans le domaine des pratiques d'évaluation des élèves du premier cycle du secondaire dans les pays couverts par l'étude TALIS (en %)



Source : TALIS, 2013

Au total, l'analyse des pratiques évaluatives par les enseignants dans la classe apporte bien une vision complémentaire de l'étude du cadre réglementaire.

En lien avec la montée de cadres réglementaires plus contraignants, elle met en évidence l'installation de pratiques innovantes qui apparaissent de plus en plus répandues dans un ensemble de pays de l'OCDE, comme la constitution de porte-folios ou l'auto-évaluation. Les enseignants français, du moins ceux intervenant au collège, car les données actuelles permettent peu une analyse des pratiques au primaire et au lycée, paraissent quelque peu en retrait face à ces dispositifs d'évaluation notamment l'auto-évaluation.

A l'opposé, l'étude de la réalité des classes, du moins telle que déclarée par les enseignants, révèle que des pratiques enseignantes peuvent se développer sans base réglementaire forte. C'est le cas notamment des pratiques de coopération entre enseignants autour de la construction de barèmes d'évaluations communes.

Conclusion de la présidente du Conseil national d'évaluation du système scolaire

Cette première analyse du Cnesco sur les évaluations internes des élèves par les enseignants révèle tout d'abord la faible exploration scientifique des réglementations et des pratiques dans les classes en la matière. Notamment dans une perspective comparatiste internationale, les analyses des cadres légaux sont très peu nombreuses et les grandes enquêtes internationales ont peu étudié les pratiques enseignantes dans le domaine de l'évaluation dans la classe. L'enseignement primaire est particulièrement peu couvert par les recherches. La prudence est donc de mise sur ce que nous pouvons conclure au terme de cette comparaison internationale sur les évaluations dans la classe qui relèvent de l'initiative des enseignants.

Si, à ce stade, l'analyse de ces évaluations internes est incomplète, la force de la stratégie de recherche fondée sur la comparaison internationale apparaît clairement. Elle nous révèle tout d'abord que, dans nombre de pays analysés dans notre étude, **l'évaluation des élèves, notamment la notation, est un sujet largement débattu dans l'enceinte de l'école, mais aussi dans et par la société civile. Clairement, le sujet des notes n'est pas exclusivement un dossier technique à dimension pédagogique. C'est un sujet de société qui enflamme régulièrement les débats publics et s'invite dans les campagnes électorales, à la fois parce que la note fait référence à nos systèmes scolaires de sélection, aux modes de construction de nos élites scolaires et donc sociales, mais aussi parce qu'elle est un véhicule de communication entre le monde scolaire et le monde familial. En ce sens, la note, produite par l'école, appartient autant à l'institution scolaire qu'aux familles qui ne manquent pas de le lui rappeler régulièrement.**

La comparaison internationale nous révèle également les caractéristiques majeures de l'évaluation interne à la française. Ses spécificités nous apparaissent clairement : aux côtés de programmes scolaires très complets corsetant l'activité pédagogique dans la classe, l'évaluation a longtemps été en France le refuge de la liberté pédagogique des enseignants. Alors que, dans la grande majorité des pays de l'OCDE, des réglementations strictes imposaient dès les années 1970 des critères d'évaluation des élèves de plus en plus prescriptifs, et que, dans d'autres pays, les collectifs enseignants fortement présents imposaient, de fait, des formes d'harmonisation, les enseignants français ont longtemps bénéficié d'une marge de manœuvre individuelle sur le sujet.

Ainsi, alors qu'à partir des années 1970, dans de nombreux pays de l'OCDE, les réflexions voient le jour autour d'évolutions de la notation numérique traditionnelle soit vers des écoles sans notes, les écoles basiques nordiques notamment, soit vers de nouvelles échelles de notation ordinales, à lettres par exemple, qui permettent de cerner plus précisément les performances attendues et évaluées en résultats scolaires, les enseignants français demeurent longtemps en dehors de tout mouvement prescriptif, à l'exception d'une timide tentative, vite révisée, dans la foulée de mai soixante-huit de repousser les notes hors de la scolarité. De même, les critères du redoublement étaient - et sont toujours - faiblement définis en termes de niveau minimal de résultats scolaires, et les pratiques de classement des élèves, dans les classes ou dans les établissements, peu éclairées par la législation.

Face à ce quasi-immobilisme réglementaire français, les autres pays de l'OCDE ont progressé pas à pas depuis plusieurs décennies. Tout d'abord les réglementations se sont affinées sur les conditions des évaluations, notamment la notation numérique a été questionnée. Il en est résulté des interrogations et plus de prescriptions sur des objectifs de connaissances et de compétences des élèves définis plus précisément aux côtés des programmes scolaires. Ces objectifs plus précis et les progrès de la recherche scientifique sur l'évaluation ont remis en cause, à leur tour, les pratiques des enseignants : de nouveaux guides orientant ces pratiques ont été produits, non pour imposer de nouvelles directions aux enseignants mais pour les conseiller sur des pratiques innovantes peu pratiquées comme les auto-évaluations, les porte-folios, ou d'autres encore. Ces réflexions ont, au final, débouché sur des politiques scolaires concernant l'évaluation, notamment formative, dans certains pays pionniers comme le Québec ou l'Angleterre, politiques dotées d'un volet de formation continue important. **Cependant, depuis une dizaine d'années, la France a clairement rejoint le mouvement de réformes internationales qui rendent les réglementations plus prescriptives. C'est le cas, récemment, avec le développement du « nouveau socle commun de connaissances, de compétences et de culture (article 13 de la loi du 8 juillet 2013) » qui comporte de nouveaux critères d'évaluation des élèves par les enseignants.** La France n'est donc pas ce « village gaulois » qui passerait à côté du raz de marée international des réformes en éducation. De même, si la collaboration entre enseignants demeure en retrait par rapport à d'autres pays qui bénéficient historiquement d'une culture du travail en équipe au sein des établissements depuis longtemps très autonomes, les coopérations entre enseignants français se sont développées.

Entre immobilisme et mouvement, entre attente réglementaire et difficultés de mise en œuvre d'une politique d'évaluation, la France est désormais au milieu du gué, comme dans de nombreux domaines en éducation, hésitant entre deux modèles d'évaluation et les superposant : l'un traditionnel fondé sur la liberté pédagogique dans la classe ainsi que sur la prescription de contenus très cadrée par les programmes scolaires et sur des évaluations certificatives harmonisées, pilotées par l'État comme notre baccalauréat ; l'autre superposant de nouveaux ingrédients avec le modèle ancien, comme des strates sédimentées - le socle commun de connaissances, de compétences et de culture, les expérimentations des écoles sans notes... **C'est cet assemblage, à ce jour hétéroclite, d'ancien et de nouveau qu'il faut mettre en cohérence à l'avenir. C'est la condition de la lisibilité de l'école pour les professionnels de l'éducation qui y œuvrent au quotidien, mais aussi pour les parents, soucieux de suivre et de comprendre les progrès de leurs enfants à l'école.**

Références

Blanchet, A. (2006). Supprimer les notes : regard rétrospectif sur une réforme contestée. Actes du 19^{ème} colloque international de l'ADMEE/SSRE 2005.

Cardinet, J. (1986). *Pour apprécier le travail des élèves*. Bruxelles : De Boeck Université.

CERU. (2010). La suppression des notes à l'école est un échec. Répliques.

Cnesco. (2014). Programme d'orientations stratégiques du Conseil national d'évaluation du système scolaire 2014-2017. Paris.

Dupriez, V., & Malet, R. (2013). *L'évaluation dans les systèmes scolaires : accommodements du travail et reconfiguration des professionnalités*. Bruxelles : De Boeck Université.

Eurypedia. (2010-2013). European Encyclopedia on National Education Systems. Eurydice. Accessible sur <https://webgate.ec.europa.eu/fpfis/mwikis/eurydice/index.php?title=Countries>

Gilliéron, P., et Tessaro, W. (2009). L'évaluation en classe en Suisse romande : des politiques aux pratiques enseignantes. *Mesure et évaluation en éducation*, vol. 32, n° 3, pp. 47-76.

Loi n° 2013-595 du 8 juillet 2013 d'orientation et de programmation pour la refondation de l'école de la République. France.

Ministère de l'éducation et de la recherche suédois. (2014). Skollagen. Accessible sur <http://www.government.se/sb/d/2063>

Mialaret, G. (1979). *Vocabulaire de l'éducation*. Paris : PUF.

National Council for Curriculum and Assessment. (2007). *Assessment in the Primary School Curriculum : Guidelines for schools*. Dublin : NCCA.

Mons, N. (2007). *Les Nouvelles Politiques éducatives*. Paris : PUF.

Mullis, I.V.S., Martin, M.O., Foy, P., & Drucker, K.T. (2012). *PIRLS 2011 : International Results in Reading*. Chestnut Hill, MA : TIMSS & PIRLS International Study Center, Boston College.

République et Canton de Genève. (2006). *Instruction publique, culture et sport, Votation du 24.09.2006, Votation sur les notes*. Accessible sur <http://www.ge.ch/dip/GestionContenu/detail.asp?mod=dossier.html&id=40>

Robert, P. (2009). La Finlande : un modèle éducatif pour la France ? Les secrets de la réussite, Paris : ESF Éditeur.

OCDE. (2014). *Résultats de TALIS 2013 : une perspective internationale sur l'enseignement et l'apprentissage*. Paris: OCDE.

OECD. (2013). *Synergies for Better Learning: An International Perspective on Evaluation and Assessment*. Paris: OCDE.

OECD. (2010). *PISA 2009 Results: What Students Know and Can Do – Student Performance in Reading, Mathematics and Science*. Paris : OECD.

OECD. (2012). *Grade Expectations: How Marks and Education Policies Shape Students' Ambitions*. Paris: OECD.



Liberté • Égalité • Fraternité
RÉPUBLIQUE FRANÇAISE

 **cnesco**
conseil national
d'évaluation
du système scolaire