Les évaluations nationales de CM2 : modes de passation et spécificités disciplinaires

National assessments at primary school: Implementation and discipline characteristics

Elisabeth Menouar et Daniel Bart

Édition électronique

URL : <https://journals.openedition.org/ree/8368>

DOI : 10.4000/ree.8368

ISSN : 1954-3077

Éditeur

Université de Nantes

Référence électronique

Elisabeth Menouar et Daniel Bart, « Les évaluations nationales de CM2 : modes de passation et spécificités disciplinaires », *Recherches en éducation* [En ligne], 19 | 2014, mis en ligne le 01 juin 2014, consulté le 02 juillet 2021. URL : <http://journals.openedition.org/ree/8368> ; DOI : <https://doi.org/10.4000/ree.8368>

Recherches en éducation est mise à disposition selon les termes de la Licence Creative Commons Attribution - Pas d'Utilisation Commerciale - Pas de Modification 4.0 International.

Les évaluations nationales de CM2 : modes de passation et spécificités disciplinaires¹

Elisabeth Menouar & Daniel Bart²

Résumé

Cet article rend compte de recherches sur les évaluations nationales de CM2 en français et en mathématiques et plus précisément sur leur passation effective en classe. En nous appuyant sur une première modélisation des séances d'évaluation, sur des observations en classe, des entretiens avec des enseignants et l'étude des prescriptions, nous analysons la mise en œuvre de ces épreuves et leurs spécificités disciplinaires. Nous mettons ainsi en évidence l'instabilité du cadrage disciplinaire de ces évaluations, la variabilité des modes d'actualisation du protocole de passation et les tensions relatives à la gestion temporelle des séances. Nos conclusions reviennent sur des questions méthodologiques et des perspectives de recherche sur les évaluations institutionnelles.

Depuis la fin des années 1980, l'École française a vu la mise en place de différents types d'évaluations institutionnelles plus ou moins durables, touchant l'ensemble des élèves d'un niveau donné du système scolaire. Ainsi, la Loi d'orientation pour l'avenir de l'École de 2005 a instauré de nouveaux programmes et de nouvelles évaluations nationales en français et en mathématiques pour les classes de CE1 et CM2. À la différence des précédentes évaluations CE2 et 6^e qui fonctionnaient depuis 1989, ces nouvelles évaluations CE1/CM2 ont été présentées par le ministère de l'Éducation nationale (2011a, p.1) plus comme des outils de mesure des acquis des élèves, d'information et de pilotage du système que comme des moyens de remédiation de la difficulté scolaire. Les objectifs plus ou moins congruents de ces évaluations ont néanmoins généré une tension chez les enseignants, partagés entre le souhait d'identifier le niveau et les difficultés des élèves, et le souci d'obtenir les meilleurs résultats possibles à communiquer à l'institution. Dès lors, les modalités de codage des exercices ou la période de passation de ces évaluations ont été progressivement modifiées par le ministère. Cependant, quatre ans à peine après leur généralisation, Vincent Peillon et George Pau-Langevin annonçaient, dès leur nomination, une nouvelle réforme des évaluations CE1/CM2³ et, dans l'attente, l'arrêt de la centralisation nationale des résultats.

C'est dans ce contexte institutionnel instable que nous nous sommes intéressés aux évaluations CE1/CM2 et plus particulièrement aux pratiques effectives de passation en fin de primaire. En effet, si la conception (Brézillon, 2008), le contenu (Rémond, 2005) ou l'utilisation (Grangeat, 1996) des évaluations nationales ont été étudiés, peu de travaux semblent avoir été consacrés au moment des épreuves en classes. Or, au-delà des protocoles institutionnels qui visent à normaliser l'effectuation de ces évaluations, que sait-on des conditions réelles de leur mise en œuvre ? Les enseignants adaptent-ils ces prescriptions, que ce soit dans l'explicitation des consignes et des contenus évalués, ou dans le temps de réponse accordé à leurs élèves ? Mais encore, y a-t-il des spécificités disciplinaires dans ces possibles variations de passation ? Autant de questions qui nous ont conduits à mener des recherches exploratoires sur les séances d'évaluation institutionnelle en CM2.

¹ Nous remercions Yves Reuter pour sa contribution précieuse à l'écriture de cet article.

² Elisabeth Menouar, doctorante - Daniel Bart, maître de conférences, Théodile-CIREL, Université de Lille 3.

³ Lettre du 26 juin 2012 de V. Peillon et de G. Pau-Langevin aux personnels de l'Éducation nationale, publiée dans le *Bulletin officiel* n°26 du 26 juin 2012.

Ce sont les principaux résultats de ces travaux que nous présentons dans la suite de cet article, après avoir exposé l'approche de l'évaluation et la méthodologie d'enquête⁴ qui nous ont permis de rendre compte de leur actualisation en salle de classe.

1. Une modélisation⁵ des évaluations pour explorer leur mise en œuvre en classe

Afin de mener notre travail d'exploration des pratiques évaluatives, nous avons élaboré une modélisation de celles-ci en macro-phases chronologiques agencées par rapport à leur effectuation en classe. Comme nous le montrons dans cette section de l'article, ce temps de passation est constitué, au vu des actions enseignantes, comme macro-phase centrale et objet majeur de nos investigations empiriques.

■ Une structuration des évaluations en macro-phases

Nous construisons l'évaluation comme articulation de trois macro-phases principales de préparation, de passation puis de correction-communication.

La macro-phase de préparation de l'évaluation se situe en amont du déroulement des épreuves en classe. Dans le cadre des évaluations nationales, l'enseignant n'a pas à concevoir son épreuve. D'origine externe (Cardinet, 1988), les tests papier-crayon en français et en mathématiques, les protocoles de passation et de codage des réponses élaborés par des « experts » sont fournis par le ministère quelques jours avant la date d'effectuation. Le travail préparatif n'est cependant pas absent de ce type d'évaluation : les enseignants que nous avons interviewés pour notre enquête⁶ s'approprient ainsi l'évaluation et son organisation sur leur temps dit « de préparation de classe ». Ils prennent par exemple connaissance du dispositif à travers les documents institutionnels, relèvent les temps de passage des différentes séquences du test, les consignes à donner ou le matériel à préparer.

L'aval de la passation se caractérise par des activités de correction et de notation des épreuves (ici, leur codage), et de communication et commentaire des résultats sous des modalités plus ou moins formelles (rendu des tests aux élèves, saisie et envoi des résultats à la hiérarchie, bulletins, réunions, etc.). Ce temps a été historiquement l'objet de nombre de travaux en psychologie (par exemple, Noizet & Caverni, 1978) ou en sociologie (par exemple, Duru, 1984), qui ont cherché à en étudier les ressorts et les tensions. De notre côté, nous nous interrogeons sur la pertinence de décliner ce moment en deux macro-phases : l'une de correction, où l'enseignant serait seul⁷ face aux copies, et l'autre de diffusion, où il rendrait publics les résultats dans et hors de l'espace restreint de la classe (parents, collègues, institution⁸, Internet⁹). Toutefois, les recherches ont montré la force des relations entre ces deux activités. Bélair et Dionne (2009) ont par exemple rendu compte des difficultés des pratiques de notation des enseignants québécois devant les variations des prescriptions institutionnelles sur le format de communication des résultats, tandis que Merle (2007) a étudié le poids des enjeux de réputation de l'enseignant, de la classe ou de l'établissement dans l'attribution des notes. Lors de notre

⁴ Nous nous appuyons essentiellement sur la démarche et les données de travaux de recherche menés dans le cadre d'un Master (Verfaillie Menouar, 2011 & 2012). Les travaux exposés ici s'inscrivent plus largement dans une recherche sur le « Vécu disciplinaire » des élèves, financée par La Sauvegarde du Nord.

⁵ Nous employons ici le terme de modélisation pour insister sur notre perspective de construction et de mise en discussion d'un outil d'investigation et d'analyse des fonctionnements concrets des évaluations scolaires et non pour signifier une visée d'élaboration d'un modèle particulier d'évaluations qu'il s'agirait de mettre en œuvre.

⁶ Il s'agit de huit enseignants de l'académie de Lille. Les différents matériaux empiriques collectés et analysés pour cette recherche sont précisés dans la partie « Les stratégies de construction et d'observation de la passation ».

⁷ Sachant que les corrections des évaluations nationales peuvent donner lieu, comme certaines évaluations locales, à des séances de travail collectif des enseignants concernés.

⁸ Rappelons qu'un certain nombre de mouvements d'enseignants ont pu appeler à ne pas communiquer les résultats de leurs élèves à l'institution (sur les enjeux professionnels et institutionnels de ces évaluations, voir Harlé-Giard & Josso-Perrochaud, 2013).

⁹ Voir par exemple le site web de cette école de l'Isère où sont publiés depuis le 1^{er} juin 2012, les résultats (non nominatifs) des élèves aux évaluations de l'année : <http://www.ac-grenoble.fr/ecoles/hg/spip.php?article3862> (consulté le 22/12/2012).

enquête, les enseignants interviewés ont, de même, dit se sentir au moins partiellement responsables des résultats de leurs élèves – « *S'ils se plantent tous, c'est de notre faute* » (MLE¹⁰) – et ils ont témoigné du malaise – « *Quand on corrige les évaluations, moi je sais, qu'à chaque fois je suis très mal* » (CF) – ou de la satisfaction – « *c'est un travail de trois ans qui est... pas récompensé... mais un petit peu quand même [...] on est fier de soi* » (EF) – que cela peut générer. Nous avons donc choisi de ne conserver qu'une seule macro-phase de correction-communication dans ce travail exploratoire.

Composante centrale de notre approche de l'évaluation, la passation correspond au temps de classe où le maître met en œuvre l'épreuve prévue. Si cette effectuation a été plus rarement étudiée que les deux phases précédentes, certaines de ses caractéristiques ou tensions spécifiques ont été mises au jour. Ainsi Cardinet (1988) décrit l'oscillation des pratiques des enseignants, entre aide aux élèves en difficulté et attitude propre à une évaluation « *froide et objective* » (p.207) exigée institutionnellement. De son côté, Chevallard (1986, p.39) s'est intéressé à la « *négociation didactique* » se jouant dans « *le secret et l'intimité de la classe* ». Ce processus renvoie aux multiples transactions implicites entre enseignant et élèves concernant les évaluations (leur fréquence, leur forme, leur difficulté, etc.). Par exemple, les conseils oraux distillés aux élèves en cours de passation peuvent permettre aux enseignants de revoir à la baisse une épreuve trop difficile, et d'éviter de la sorte des conflits ou le découragement de la classe, tout en maintenant l'exigence apparente du travail demandé à l'écrit, qui représente ce que les élèves devraient être capables de faire. Toutefois, comme nous le montrons dans la suite, nous pensons que, pour décrire plus systématiquement les pratiques concrètes de passation des évaluations institutionnelles et leurs possibles variations disciplinaires, il est nécessaire d'en spécifier les différentes dimensions.

■ **Les dimensions principales de la macro-phase de passation**

La mise en œuvre d'une évaluation constitue un temps particulier de la vie d'une classe. Entouré d'une certaine solennité, ce moment se distingue du temps d'enseignement et d'apprentissage, mais aussi de l'exercice quotidien, parce qu'« *on introduit un peu plus de cérémonie, de stress et d'équité formelle* » (Perrenoud, 1998/1999, p.81). Dans le but de travailler plus précisément à l'analyse des séances d'évaluation, nous proposons leur découpage en trois phases structurées au regard des actions et des interventions de l'enseignant.

La phase d'*installation* de la situation d'évaluation (dite phase 1) marque le lancement d'une séance. Elle est généralement initiée par l'annonce et la mise en place par l'enseignant, de ses conditions matérielles de réalisation. Cela peut nécessiter de réorganiser partiellement l'espace de la classe, de déplacer des élèves ou des tables, de ranger ou tenir à disposition certains matériels, etc. Mais c'est plus particulièrement aussi le temps de la communication des diverses consignes de passation aux élèves. « *Système complexe* » (Lahanier-Reuter & Reuter, 2004, p.80), combinant consignes écrites et/ou orales, celles-ci touchent aux tâches à effectuer et au cadre de l'effectuation (par exemple, le temps disponible ou les modalités de circulation et d'échanges durant la passation).

Lors de la phase de *contrôle* de la situation d'évaluation (phase 2), l'enseignant laisse *a priori* les élèves réaliser la tâche demandée de manière autonome et dans un temps imparti. Dans une relative extériorité par rapport à ses élèves, celui-ci chercherait à maintenir un ordre propice à l'activité au regard des règles établies dans la phase précédente ou dans les habitudes de la classe. Nous montrerons plus loin les tensions que génère cette composante de la passation chez les enseignants, par exemple entre « *intervenir ou laisser faire* » les élèves face à la consigne (Zakhartchouk, 2000, p.71) ou encore entre préserver le silence de la classe et réguler le déroulement de l'épreuve par de multiples interventions orales.

La phase 3 est la phase d'*achèvement* de la séance. Si nous considérons que cette phase se termine par le ramassage de la dernière copie, nous pensons néanmoins qu'elle ne s'y limite

¹⁰ Les codages de ce type sont placés à la suite des extraits d'entretien de recherche ou des extraits de retranscription de passation pour indiquer quels enseignants sont les auteurs des discours cités.

pas. Celle-ci débute selon nous plus avant, dès lors que des élèves cessent de composer, de manière plus ou moins concordante avec la fin prévue de l'épreuve. Cette période, que nous détaillons plus loin, s'avère ainsi complexe puisque l'enseignant va être aux prises avec deux groupes d'élèves aux effectifs variant de manière inverse, aux modalités d'activités non congruentes et pour une durée difficile à anticiper. Ceci est plus particulièrement observable lorsqu'une activité « d'attente » (par exemple, lire) est autorisée pour les élèves qui ont stoppé leur épreuve.

Le modèle que nous venons d'exposer est construit en référence aux différents types d'évaluation (institutionnelles, ordinaires, etc.). Il s'agit d'une première construction heuristique pour penser et mettre au jour l'agencement particulier des évaluations concrètement mises en œuvre. Il ne dessine donc pas pour nous une structure et une temporalité de la phase 1 à la phase 3 fixes. Par exemple, à la différence des évaluations locales, l'articulation phase 1 (*installation*) et phase 2 (*contrôle*) dans les évaluations institutionnelles se répète autant de fois qu'il y a d'exercices puisque consignes orales et réponses des élèves s'enchaînent successivement, exercice par exercice. La phase 3 y prend également une forme différente : les enseignants que nous avons observés n'autorisent généralement aucune autre activité que celle d'attendre l'exercice suivant en silence.

■ **Les stratégies de construction et d'observation de la passation**

L'objectif de notre démarche de recherche est de décrire et comprendre les pratiques effectives de passation des évaluations institutionnelles de CM2 et leur spécificité disciplinaire. Pour ce faire, nous avons croisé différentes méthodes de construction de données : observations¹¹ outillées de dix séances des évaluations 2011 (quatre en français et six en mathématiques) enregistrements audio et retranscriptions de ces séances, entretiens avec les huit enseignants concernés, études documentaires multiples (cahiers des évaluations, documents destinés aux maîtres, textes officiels).

Les résultats présentés par la suite reposent majoritairement sur l'analyse des dix passations retranscrites. L'unité de base des retranscriptions que nous avons adoptée est le tour de parole, que nous avons défini comme une unité de discours continu d'un locuteur, délimité soit par un changement de destinataire ou de locuteur, soit par une interruption de plus de cinq secondes du discours. Chaque tour de parole saisi est en outre renseigné par l'action ou le déplacement conjoint de l'enseignant, un repérage des élèves concernés et le minutage de ces « évènements ». Cette technique de découpage nous a permis de matérialiser de manière dynamique l'initiateur et le ou les destinataire(s) du discours : de la classe entière à un élève en particulier en passant par l'échange avec un élève évoluant en discours collectif (Bourgeois, 2007, p.66). Comme nous le montrons dans la seconde partie de l'article, l'analyse de ces formes d'échange en fonction des différentes phases de passation et de la discipline nous a permis de caractériser des modalités de pratiques de l'évaluation.

Nous avons également considéré que le livret des évaluations nationales destiné à l'enseignant (MEN, 2011a) constituait une base pertinente pour cerner la pratique prescrite par l'institution. En nous appuyant sur les règles et les conditions de l'évaluation qui y sont précisées, sur les consignes de réalisation des exercices que l'enseignant doit lire aux élèves et sur les temps de réponse qu'il doit leur allouer, nous avons donc reconstruit ces pratiques que nous qualifions de *protocolaires* en français et en mathématiques. Par là même, nous souhaitons nous donner les moyens d'observer les différences entre instructions et actualisations¹². En effet, nous considérons que les modalités (objet, fréquence, etc.) de ces écarts fournissent des indicateurs pertinents pour caractériser des types de mises en œuvre enseignantes de l'évaluation. C'est à l'analyse des principaux résultats obtenus dans le cadre de cette démarche que nous allons maintenant nous consacrer.

¹¹ Nos travaux de recherche nous ont également conduits à observer treize séances d'évaluations ordinaires (sept en français et six en mathématiques).

¹² Précisons que si nous étudions les écarts entre les prescriptions et leurs réalisations, nous les considérons ici moins comme des biais méthodologiques qu'il s'agirait de limiter que comme des indicateurs des tensions, contraintes, opportunités, incertitudes, etc. de la mise en œuvre effective des protocoles.

2. Les modes de passation des évaluations nationales

Alors que le projet même d'évaluations nationales suppose une homogénéité des modes de passation des épreuves que représente le protocole prescrit, il est difficile de se prononcer sur le niveau d'harmonisation des pratiques effectives. C'est ce que nous allons voir, en analysant successivement la mise en œuvre du protocole de l'évaluation, les modalités d'intervention des enseignants et leur gestion temporelle des épreuves. Mais nous verrons d'abord qu'il existe une variabilité non négligeable et quelque peu problématique entre les programmes d'enseignement officiels, les cadres disciplinaires identifiables dans les évaluations nationales en français et en mathématiques et leur actualisation dans la classe.

■ *Le cadrage disciplinaire des épreuves*

En mathématiques, la comparaison entre le protocole de l'évaluation nationale (MEN, 2011a) qui correspond aux contenus à évaluer et les instructions officielles¹³ c'est-à-dire les contenus à enseigner, montre une différence légère. Excepté le découpage du domaine « Nombres et calculs » des programmes en deux domaines distincts dans les évaluations, les trois autres domaines disciplinaires majeurs (« Géométrie », « Grandeurs et mesures », « Organisation et gestion des données ») se retrouvent désignés à l'identique dans le protocole du test national.

En revanche, en français, l'organisation disciplinaire entre les programmes et les évaluations semble moins concordante. Toutefois, même au sein des programmes officiels de l'école primaire, les découpages de la discipline varient entre la partie descriptive des contenus (p.21-23) et la partie consacrée à la progressivité des enseignements (p.34-37). Dans la partie descriptive, les contenus sont rattachés à trois domaines (« Langage oral », « Lecture, écriture », « Étude de la langue française ») réunissant des sous-domaines, tandis qu'ils sont déclinés selon des sous-domaines juxtaposés tels des disciplines (Chervel, 1988 ; Reuter, 2004) dans les progressions proposées. Le domaine « Étude de la langue » n'apparaît plus dans ces dernières et les sous-domaines qui le composaient dans les programmes décrits y sont situés dans le domaine « Lecture, écriture ». De plus, dans ce domaine « Lecture, écriture » des progressions, un sous-domaine « Écriture » est ajouté à côté du sous-domaine « Rédaction ». Dans les évaluations nationales, les catégories se présentent encore autrement, puisque, dans la synthèse des connaissances et compétences évaluées (MEN, 2011a, p.3), seuls les sous-domaines « Vocabulaire », « Grammaire » et « Orthographe » sont rattachés à un domaine (« Étude de la langue »), lui-même juxtaposé à deux autres ensembles, « Lire » et « Écrire », qui ne sont pas spécifiés ainsi dans les textes officiels. Enfin, le « Langage oral », domaine des programmes décrits et sous-domaine des progressions, n'est pas visé par les évaluations nationales, ce choix n'étant véritablement explicité à aucun moment.

L'élaboration et l'organisation prescrite des épreuves ne semblent qu'indirectement découler de cette actualisation des disciplines et elles matérialisent un cadrage disciplinaire peu apparent. Par exemple, les six séquences¹⁴ disciplinaires de l'évaluation comportent entre quatre à douze exercices relevant de différents sous-domaines sans logique explicite de regroupement et d'enchaînement. Les exercices peuvent également comporter des items renvoyant à différents sous-domaines, voire aux deux disciplines dans le cas de l'exercice 19 de la troisième séquence de français. Cet exercice touche ainsi à la fois au sous-domaine « Lire » du français et au domaine « Organisation et gestion de données » rattaché aux mathématiques. Aucune information orale prescrite et/ou indication écrite n'est d'ailleurs prévue par l'institution pour préciser cette particularité aux élèves (et aux enseignants). Cela est plus généralement le cas, puisque les (sous-)domaine(s) évalué(s) ne sont pas précisés dans les exercices et que les

¹³ *Bulletin Officiel*, Hors-série n°3 du 19 juin 2008, « Horaires et programmes d'enseignement de l'école primaire ».

¹⁴ Que ce soit en français ou en mathématiques, les évaluations nationales CM2 sont structurées en trois séquences disciplinaires regroupant de quatre à douze exercices.

disciplines concernées ne le sont que très faiblement, principalement au moyen de la séparation en deux parties distinctes du cahier des élèves¹⁵ (MEN, 2011b).

Enfin, au-delà des prescriptions, nous avons également analysé la manière dont ces découpages disciplinaires s'actualisent au cours des passations. La discipline « français » est annoncée trois fois sur quatre par l'enseignant dès les premiers tours de parole de la séquence, tandis que la discipline « mathématiques » ne l'est qu'une fois sur deux. Sur les dix séances observées, seules trois ont par la suite fait l'objet d'un rappel du cadre disciplinaire au cours de l'épreuve, dont deux fois pour justifier une modalité de réponse du type « *On écrit les résultats en chiffres, on est en mathématiques !* » (MLE). Nous n'avons relevé aucune autre indication du rattachement disciplinaire d'un exercice, que ce soit au moyen d'indications écrites ou orales ou *via* des précisions sur les contenus évalués. Il existe cependant d'autres vecteurs de cadrage disciplinaire plus implicites, tels que l'annonce d'une utilisation possible d'outils comme le dictionnaire en français ou le compas et l'équerre en géométrie. L'enseignant peut donc apporter, implicitement ou explicitement, des indications disciplinaires dont on peut penser qu'elles sont adaptées à sa classe. Ces modalités indicatives, sans être homogènes et massives, s'éloignent cependant d'un cadre protocolaire qui précise qu'« *aucune autre indication (que ce qui est prévu) ne doit être donnée* » (MEN, 2011a, p.1). Comme nous le montrons dans la partie suivante, les discours et actions des enseignants montrent encore plus nettement d'autres actualisations des évaluations prescrites.

■ **L'actualisation du protocole**

L'étude du livret d'évaluations nationales nous a permis de construire et poser un mode de passation référent : la pratique protocolaire. Nous proposons de le définir par la mise en œuvre des consignes prescrites à l'enseignant (par exemple dire telle chose aux élèves, leur laisser tant de minutes pour répondre) et l'absence d'initiatives autres que celles liées au maintien des conditions d'évaluation (par exemple demander le silence). Cette pratique protocolaire est rarement observée : seules deux séances de mathématiques peuvent être qualifiées ainsi.

Les interventions enseignantes non prévues sont en effet nombreuses et parfois même signalées aux élèves : « *Je vous explique parce que ça va poser problème [...] normalement, je n'ai pas le droit...* » (CF). Mais celles-ci peuvent aussi se manifester de manière non verbale, par un geste ou un regard signalant à un élève une erreur dans son cahier.

En analysant ces écarts au protocole, nous avons caractérisé quatre autres modes d'actualisation des passations. Nous les détaillons ci-après par ordre croissant d'ouverture possible aux échanges avec les élèves, en principe exclus par les consignes d'effectuation.

La *pratique prudente* est plus particulièrement repérable durant la phase *d'installation* de l'évaluation. L'enseignant adapte alors les consignes du protocole qu'il doit lire aux élèves en fonction des difficultés qu'il prévoit ainsi que l'explique CF en cours d'entretien : « *Comme l'année passée ils se sont plantés dans la consigne, en la lisant j'ai insisté sur entourer et souligner...* ». Cette pratique se retrouve essentiellement en mathématiques. Elle peut être associée au mode de *passation dirigée*. Ce phénomène est surtout visible durant la phase de *contrôle* de la situation d'évaluation. Il s'agit ici, non plus seulement d'avertir collectivement les élèves d'un écueil éventuel comme dans le cas précédent, mais d'intervenir individuellement et/ou collectivement, en fonction d'un dysfonctionnement effectif observé dans un cahier. Cette intervention verbale ou non, ne consiste pas à indiquer la bonne réponse mais vise plutôt à alerter les élèves de la nécessité de revoir une réponse, par une forme d'invalidation implicite et anticipée : comme nous l'explique MLE « *Si on n'est pas derrière [...], ça part vite* ».

¹⁵ Un sommaire précise la pagination des deux parties explicitement attribuées aux séquences de français et de mathématiques. La fin d'une séquence disciplinaire est indiquée après le dernier exercice de celle-ci par la mention : « Fin de la x^e séquence de français (ou de mathématiques) ». Lorsque le cahier est ouvert à une page d'exercices, rien n'indique donc la discipline, à moins qu'on soit sur la dernière page de la séquence. Enfin, le protocole ne demande pas aux enseignants de préciser à l'oral ou l'écrit (tableau) les disciplines évaluées.

Les deux autres modes de passation sont caractérisés par des interactions plus nettes avec les élèves. La pratique *aidante* est celle de l'enseignant qui accepte, sans toujours annoncer que cela est possible, de répondre individuellement ou collectivement aux questions spontanées des élèves. La pratique *accompagnée*, quant à elle, est encore plus explicitement ouverte aux échanges. L'enseignant peut ainsi lui-même questionner ses élèves : « *Alors, la consigne est comprise ?* » (JT) ou chercher à les faire réfléchir en leur renvoyant des questions ou en sollicitant des avis. Ces deux modes de passation sont majoritairement retrouvés en français.

Cette catégorisation provisoire et ses rattachements disciplinaires ne sont cependant pas exclusifs : ils reposent sur l'identification de tendances majoritaires durant les séances d'évaluation et n'excluent donc pas des combinaisons de modes de passation. Comme nous le montrons dans la partie suivante, les actions enseignantes sont en effet omniprésentes et diverses, que ce soit pour un même enseignant, une même séquence ou une même discipline.

■ **La diversité et l'omniprésence des actions enseignantes au sein de l'évaluation**

Lecture des consignes, contrôle et rappel du temps, maintien des élèves en situation d'épreuve... la mise en œuvre des évaluations nationales, même hypothétiquement limitée à l'application des procédures protocolaires, ne se réduit jamais à de la « surveillance ». Ainsi, l'installation de l'évaluation suppose une annonce plus ou moins solennelle de l'enseignant, de possibles réaménagements de la disposition de la classe (par exemple, pour séparer les élèves), la vérification et la distribution d'outils, etc. Comme nous l'avons vu plus haut, certains enseignants complètent les consignes pour expliquer, reformuler, apporter des précisions sur la présentation et la forme de la réponse attendue, tandis que d'autres s'informent de la compréhension de la tâche à réaliser en questionnant des élèves.

Le contrôle de la situation est marqué par la circulation des enseignants entre les rangs, durant laquelle une majorité d'entre eux regarde les productions des élèves. Différentes formes de communication sont également observées, que ce soit dans des interventions spontanées de l'enseignant ou des réponses à des demandes d'élèves¹⁶. L'enseignant peut s'adresser à voix haute à tous les élèves pour donner des informations temporelles, s'approcher d'un élève pour un échange individuel à voix basse ou répondre à voix haute en élargissant son discours à l'ensemble de la classe (Bourgeois, 2007). Sous une forme non langagière (Postic, 1977/1992), outre les gestes et mimiques pour le maintien de l'ordre, l'enseignant peut également pointer silencieusement quelque chose dans le cahier d'un élève.

Nos observations montrent que les interventions enseignantes touchent principalement aux consignes ou portent sur des conseils méthodologiques invitant l'élève à bien lire ce qui lui est demandé, à se relire ou à se concentrer, voire forment, comme nous l'avons dit, une validation/invalidation implicite d'une réponse. Nos résultats soulignent également que l'enseignant peut amener les élèves à faire un lien avec ce qui a déjà été abordé en classe (Delannoy-Courdent & Delcambre, 2007). Il peut encore faire référence à une méthode ou démarche habituelle de la classe pour préciser la tâche attendue dans un exercice des évaluations ou pour expliquer un format de question ou de réponse dans le cahier du test. Nous n'avons relevé que de rares interventions centrées sur les contenus disciplinaires. Le cas échéant, il s'agit plutôt de proposer une définition d'un mot de l'exercice ou de la consigne. Plus exceptionnellement encore, un enseignant peut expliciter davantage des contenus de façon très individualisée, à voix basse, avec les élèves en difficulté.

Enfin, nos observations montrent que les enseignants tendent à intervenir pour stimuler les élèves qui ont pu interrompre ou cesser leur travail devant un problème ou par découragement. Durant la phase terminale des exercices, ils ignorent plutôt ou feignent d'ignorer les élèves qui ont terminé leur travail avant la fin du temps imparti. Dans le cas contraire, cela peut amener tout au plus l'enseignant à des rappels à l'ordre de façon à garder le contrôle de la situation

¹⁶ Nos résultats montrent qu'il y a en moyenne 1,5 tour de parole par minute d'évaluation.

d'évaluation dans une ambiance silencieuse. L'approche de la fin du délai de réponse à un exercice est généralement annoncée avant d'être signalée par un « *Posez vos crayons !* ». La clôture des séquences, concomitante avec la fin du temps imparti pour le dernier exercice, est le plus souvent indiquée par une courte annonce de l'enseignant, du type « *C'est terminé, fermez les cahiers* » ou le ramassage de ceux-ci.

D'une manière générale, nous n'avons pas observé d'enseignants semblant éloignés de la situation concrète d'évaluation ou se lançant dans une activité parallèle¹⁷. Cela peut être lié au caractère officiel des évaluations institutionnelles. Sans doute cela peut-il aussi s'expliquer par notre présence dans la classe ou par le fait que les enseignants qui ont accepté d'être observés se distinguent par une certaine implication dans la conduite de ces épreuves. Mais cela est peut-être plus encore lié au fait que le déroulement du protocole d'évaluation exige de l'enseignant un contrôle quasi permanent du « timing » et de l'enchaînement des exercices. C'est ce que nous explorons dans la section suivante.

■ **La temporalité de la passation : enchaînement et contrôle des durées**

Contrairement aux évaluations ordinaires, où le temps de travail alloué aux élèves est généralement défini pour l'ensemble de l'épreuve considérée, chaque exercice des évaluations nationales est à réaliser successivement, dans un temps précisé par la procédure (entre deux et vingt-cinq minutes). Le protocole insiste donc sur la nécessité de contrôler et de signaler régulièrement le temps de réponse aux élèves ainsi que de disposer d'une horloge visible dans la classe (MEN, 2011a, p.1). Si l'importance de cette question se manifeste par l'usage généralisé du chronomètre chez les maîtres observés, aucun d'entre eux n'a cependant invité ses élèves à utiliser cette horloge durant nos observations.

Le contrôle de la temporalité des exercices est toutefois problématique. On remarquera tout d'abord que le cadre protocolaire précise que la durée de passation indiquée pour chaque exercice comprend le temps devant être consacré à la lecture par l'enseignant du texte de consigne. Mais cette délimitation de la durée est contradictoire avec le fait que, pour chaque exercice, l'annonce du temps de réponse dont disposent les élèves est placée¹⁸ à la fin de ces consignes lues par les maîtres, instant où ces derniers lancent le décompte du temps. Cela provoque donc nécessairement un écart entre les durées prescrites et les durées effectives (voir *infra*).

La gestion de la durée globale des séquences est également délicate. En effet, le protocole ne précise pas la durée de chaque séquence mais il donne seulement des indications générales (cinquante minutes par séquence de français et trente minutes en mathématiques). Or, le résultat obtenu par l'addition des temps attribués aux exercices composant une séquence dépasse généralement de quelques minutes cette indication théorique. Si on peut considérer que ce dépassement est faible (par exemple, trois minutes pour la séquence 3 de français), notons tout de même que celui-ci correspond à la durée de nombreux exercices de ces évaluations. De plus, les indications temporelles relatives aux séquences ne prennent pas en compte l'intervalle de temps qui sépare la passation de deux exercices successifs. Par conséquent, plus une séquence comporte d'exercices (jusqu'à douze dans la séquence 1 de mathématiques), plus le contrôle de sa durée en fonction des indications protocolaires générales peut être problématique.

Si, comme cela est requis par le protocole, on tient compte du temps d'énonciation des consignes dans la mesure des durées de passation, les résultats montrent effectivement que la durée moyenne des séquences en classe est supérieure d'au moins 25% à la moyenne des durées prescrites. En mathématiques, les durées excèdent jusqu'à quinze minutes les trente et une minutes protocolaires, et en français, les durées relevées s'étendent jusqu'à dix-sept

¹⁷ Sauf dans le cas d'un cours double où l'enseignant a prévu une activité autonome pour le groupe non concerné par l'évaluation.

¹⁸ Même si les enseignants peuvent effectivement trouver le temps qu'ils devraient globalement consacrer à chaque exercice dans un tableau de synthèse (MEN, 2011a, p.2).

minutes au-delà des cinquante et une minutes prescrites. Cet accroissement n'est cependant pas homogène et il convient donc de rester prudent dans la manipulation de ces indicateurs : la différence du temps de passation d'une même séquence d'exercices entre deux enseignants peut aller jusqu'à sept minutes.

L'analyse croisée des délais de réponse annoncés par l'enseignant et des durées effectivement accordées aux élèves montre qu'il y a peu de différences. Ces dernières sont même négligeables en mathématiques, où le temps de réponse est parfois écourté lorsque l'enseignant considère et s'assure que tous ses élèves ont répondu avant le délai imparti. En français, la moyenne des durées accordées au cours d'une séquence est supérieure d'environ 10% à la moyenne des durées annoncées (issues du protocole), c'est-à-dire environ cinq minutes par séquence. Notons d'ailleurs que l'enseignant indique parfois à ses élèves cette marge supplémentaire : « *C'est terminé normalement. Julien, tu as fini ? Maxime ? [...] Finis tes mots, vas-y. Allez Maxime, dépêche-toi !* » (CF).

L'allongement des temporalités effectives par rapport aux temporalités prescrites provient donc moins d'une extension de la phase de *contrôle* de la situation d'évaluation que de la phase d'*installation*. Cela est très net en mathématiques, puisqu'il est rare qu'un temps de réponse supplémentaire soit donné. L'allongement de la phase d'*installation* pourrait ici tenir au nombre plus important d'exercices en mathématiques, qui tend à multiplier les temps de consignes. Mais cela doit également être mis en relation avec le *mode prudent* d'actualisation du protocole que nous avons plus particulièrement caractérisé dans cette discipline et où l'enseignant complète les consignes, notamment par des formes d'avertissement aux élèves (voir la partie 2.2).

Ainsi, le non-respect du protocole qui exige de compter l'énonciation des consignes dans le temps de réalisation des exercices concrétiserait, non seulement l'abandon d'une procédure discordante mais également, que les ajustements de contenu et de durée de la phase d'*installation* de la situation sont les préalables nécessaires aux enseignants pour l'application d'un contrôle du temps de réponse relativement ferme.

Conclusion

Nos résultats montrent donc l'importance et la variabilité de l'activité enseignante dans une actualisation des évaluations nationales sensible aux *configurations disciplinaires* (Reuter, 2010). Si nos travaux ne permettent pas de tirer de conclusions sur les effets de cette actualisation en matière de validité méthodologique ou de performances des élèves, ils conduisent néanmoins à s'interroger sur les conditions d'une comparaison de telles évaluations à l'échelon national. Ni bouleversés ni standardisés, les modes de passation nous paraissent plutôt en tension entre l'application d'un protocole et son accommodation à l'histoire, aux habitudes de travail ou aux manières de faire de l'enseignant, des élèves et de la classe. Ces modes de passation renforcent peut-être en cela la validité interne de l'épreuve et l'inscription d'un dispositif externe dans les cadres de fonctionnement pédagogiques et didactiques de la classe. Ces pistes de réflexion gagneront à être méthodiquement interrogées dans nos recherches à venir.

Cette première approche des évaluations nationales nous amène également à poursuivre nos réflexions sur l'affichage du cadre et des contenus disciplinaires évalués (Fluckiger & Bart, 2012). En l'absence de rattachement explicite des objets évalués à des (sous-)domaines, quelle intelligibilité ont les élèves des contenus testés ? D'autant que ceux-ci varient à chaque exercice d'une même séquence et que les explications des enseignants portent rarement sur la discipline concernée ou l'objet disciplinaire évalué. De plus, que peut-on dire de ces interventions enseignantes ? Généralement considérées comme aidantes pour les élèves, ne complexifient-elles pas plutôt la réflexion de certains en alourdissant les consignes ? Ou conduisent-elles les élèves à répondre de manière conforme aux attentes de l'enseignant, c'est-à-dire conforme à l'image que l'enseignant a de leur niveau ?

Enfin, nos investigations empiriques nous amènent à interroger notre modélisation des évaluations. Les incertitudes et questionnements dans la mise à l'épreuve de cet outil d'analyse provisoire ont tout particulièrement porté sur le repérage des limites des phases et macro-phases. Notre approche souple et évolutive de l'évaluation se double donc d'une certaine prudence quant à la portée de nos résultats. Ainsi, les chevauchements et les alternances rapides d'une phase à l'autre, l'imbrication de macro-phases entre elles (par exemple, lorsque la correction est débutée en cours de passation) ne sont pas rares dans les pratiques observées. Les choix de délimitation et de catégorisation sont délicats et discutables, et doivent faire l'objet de travaux complémentaires. Travaux d'autant plus nécessaires que les résultats des évaluations institutionnelles sont bien souvent exposés (Kalali & Jenkins, 2012), commentés ou interprétés en l'absence de connaissances solides sur leurs conditions de construction et de validité.

Bibliographie

- BELAIR L.M. & DIONNE E. (2009), « Évaluation des apprentissages dans le contexte québécois : entre décisions politiques et pratiques en salle de classe », *Mesure et évaluation en éducation*, volume 32, n°3, p.77-100.
- BOURGEOIS I. (2007), « Caractérisation des interventions enseignantes », *Aster*, n°45, p.65-90.
- BREZILLON J. & CHAMPAULT F. (2008), « Les évaluations diagnostiques en France », *Evaluer pour former*, G. Baillat, J.-M. De Ketele, L. Paquay & C. Thélot (dir.), Bruxelles, De Boeck, p.83-98.
- CARDINET J. (1988), *Évaluation scolaire et pratique*, Bruxelles, De Boeck.
- CHERVEL A. (1988), « L'Histoire des disciplines scolaires : réflexions sur un domaine de recherche », *Histoire de l'éducation*, n°38, p.59-119.
- CHEVALLARD Y. (1986/1991), « Vers une analyse didactique des faits d'évaluation », *L'évaluation : approche descriptive ou prescriptive ?* J.-M. De Ketele (dir.), Bruxelles, De Boeck, p.31-59.
- DELANNOY-COURDENT A. & DELCAMBRE I. (2007), « Pratiques langagières du maître et construction d'un rapport aux pratiques d'enseignement scientifique », *Aster*, n°45, p.113-138.
- DURU-BELLAT M. (1986). « Notation et orientation : quelle cohérence, quelles conséquences ? », *Revue Française de Pédagogie*, n°77, p.23-38.
- FLUCKIGER C. & BART D. (2012), « L'introduction du B2i à l'école primaire : évaluer des compétences hors d'une discipline d'enseignement ? », *Questions Vives*, volume 7, n°17, 2012.
<http://questionsvives.revues.org/1006>, consulté le 18 décembre 2012.
- GRANGEAT M. (1996), « Tirer parti de l'évaluation d'entrée en 6^e », *Revue Française de Pédagogie*, n°115, p.89-98.
- HARLÉ-GIARD I. & JOSSO-PERROCHAUD A. (2013), « Les évaluations nationales en CE1 et CM2 en France : entre enjeux institutionnels et tensions professionnelles », *Spirale*, n°51, p.45-58.
- KALALI F. & JENKINS W.J. (éd.) (2012). « PISA, TIMSS : regards croisés et enjeux actuels », *Recherches en Éducation*, n°14.
- LAHANIER-REUTER D. & REUTER Y. (2004), « Décrire une figure géométrique. Éléments d'interrogation à propos de l'écriture en mathématiques et en français », *Les cahiers Théodile*, n°5, p.79-91.
- MERLE P. (2007), *Les notes. Secrets de fabrication*, Paris, Presses Universitaires de France.
- MINISTÈRE DE L'ÉDUCATION NATIONALE, DE LA JEUNESSE ET DE LA VIE ASSOCIATIVE, DGESCO (2011a), Évaluation nationale des acquis des élèves en CM2. Français et mathématiques. Livret de l'enseignant.
- MINISTÈRE DE L'ÉDUCATION NATIONALE, DE LA JEUNESSE ET DE LA VIE ASSOCIATIVE, DGESCO (2011b), Évaluation nationale des acquis des élèves en CM2. Français et mathématiques. Cahier de l'élève.

NOIZET G. & CAVERNI J.-P. (1978), *Psychologie de l'évaluation scolaire*, Paris, Presses Universitaires de France.

PERRENOUD P. (1998/1999), *L'évaluation des élèves. De la fabrication de l'excellence à la régulation des apprentissages*, Paris, De Boeck.

POSTIC M. (1977/1992), *Observation et formation des enseignants*, Paris, Presses Universitaires de France.

REMOND M. (2005), « Regards croisés sur les évaluations institutionnelles », *Repères*, n°31, p.113-140.

REUTER Y. (2010), « Discipline », *Dictionnaire des concepts fondamentaux des didactiques*, Y. Reuter (dir.), Bruxelles, De Boeck, p.85-89.

REUTER Y. (2004), « Analyser la discipline : quelques problèmes pour la recherche en didactique », *Actes du 9^e colloque de l'AIRDF*, Québec.

VERFAILLIE-MENOUAR E. (2011), *Évaluations institutionnelles : conceptions et pratiques d'enseignants de CM2*, Travail d'Études et de Recherche de master 1, sous la direction de Daniel Bart, Sciences de l'Éducation, Université Lille 3.

VERFAILLIE-MENOUAR E. (2012), *Pratiques enseignantes lors de séances d'évaluations écrites en CM2. Étude comparative en français et mathématiques*, Mémoire de master 2, sous la direction de Daniel Bart et Yves Reuter, Sciences de l'Éducation, Université Lille 3.

ZAKHARTCHOUK J.-M. (2000), « Les consignes au cœur de la classe : geste pédagogique et didactique », *Repères*, n°22, p.61-82.