

COLLOQUE INTERNATIONAL - *INTERNATIONAL SYMPOSIUM*

# Évaluation en mathématiques : dispositifs, validités & pratiques

Mathematics assessment :  
devices, validity & teachers' practice

21 - 22  
NOV  
2016

**PROGRAMME**



# INFORMATIONS PRATIQUES

## Dates et heures

- Lundi 21 novembre 2016, de 9h à 18h (*accueil à partir de 8h30*)
- Mardi 22 novembre 2016, de 9h à 17h (*accueil à partir de 8h30*)

## Contact

*evaluationmaths@sciencesconf.org*

## Lieu

Université Paris-Est Créteil (UPEC)  
Centre Mail des Mèches - Bâtiment d'Économie  
Rue Poète et Sellier - 94000 Créteil  
*Accès : Métro ligne 8 « Créteil Université »*

## Salles

Amphithéâtre 1  
Salles 301, 302, 314 et 315

## WiFi

1. Choisir le réseau « Invites-Paris12 » (sans configuration particulière)
2. Après connexion au réseau « Invites-Paris12 » ouvrir un navigateur Web ; vous serez redirigé vers le portail d'authentification :

*Identifiant : sceco\_invite*

*Mot de passe : t8zHFD*

## Déjeuners

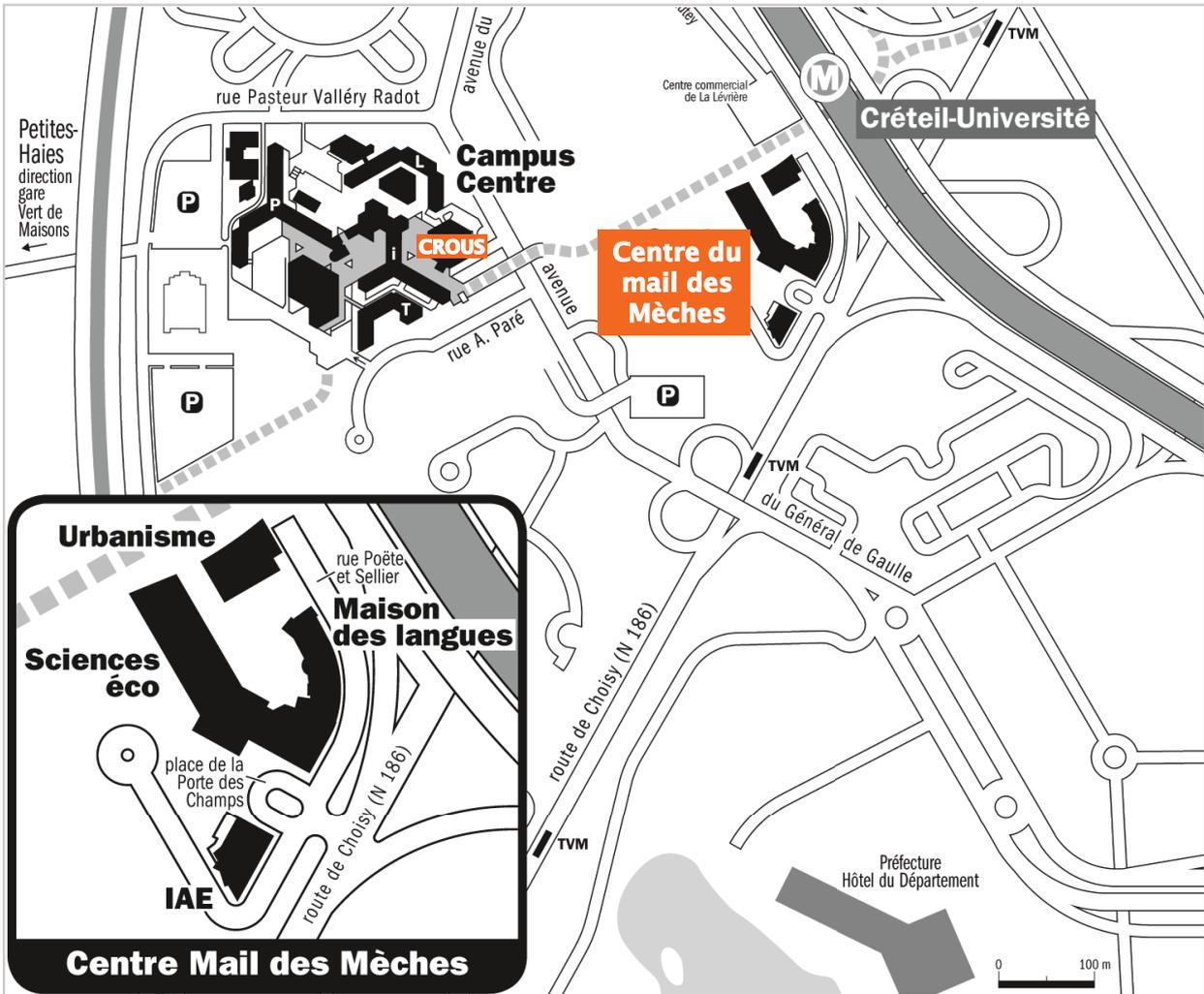
Restaurant Universitaire de la Pointe Jaune (CROUS)  
61, avenue du Général De Gaulle  
94000 Créteil  
*Voir plan n°2 ci-contre*

## Dîner festif

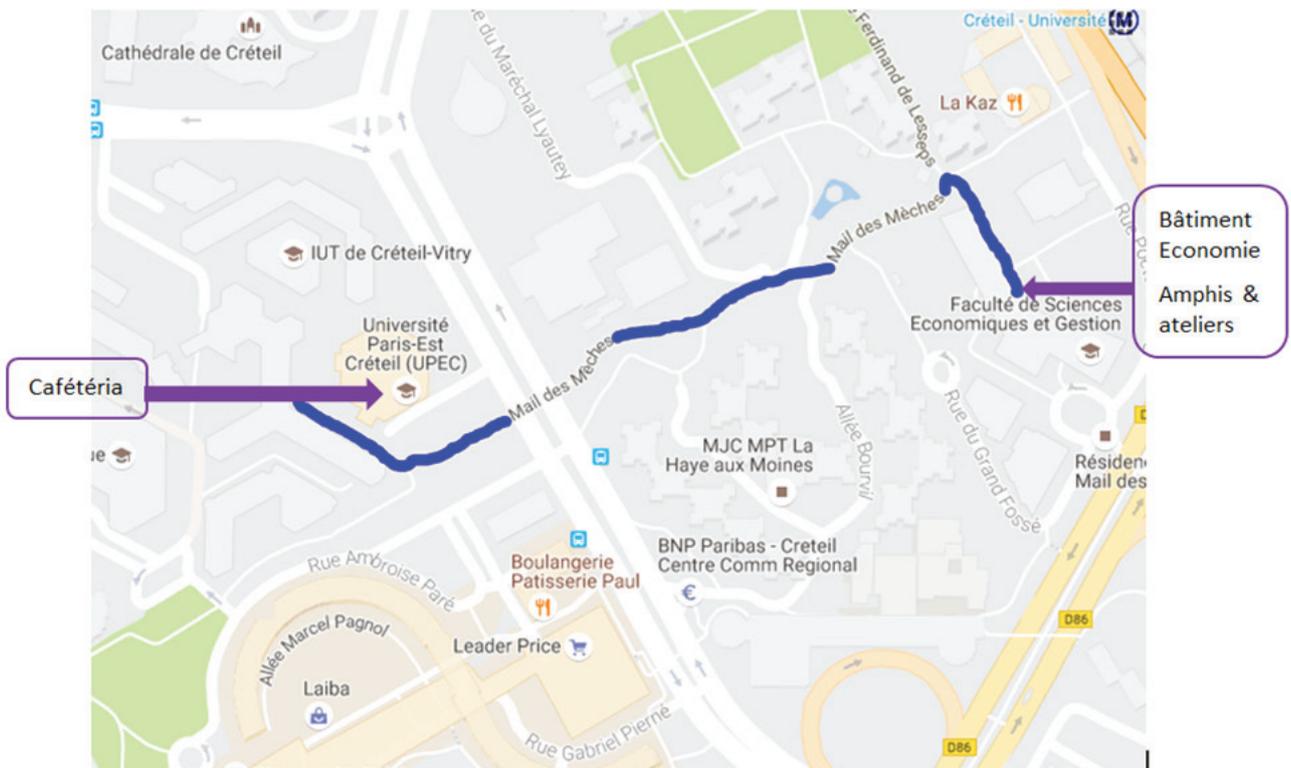
Lundi 21 novembre 2016  
Restaurant « Le Train Bleu »  
Gare de Lyon, place Louis Armand  
75012 Paris  
*Accès : Métro lignes 1 et 14 « Gare de Lyon »*

# PLANS UPEC et Mail des Mèches

## 1. Université Paris-Est Créteil (UPEC)



## 2. Trajet du Centre du Mail des Mèches au restaurant de la « Pointe Jaune » (CROUS)



# INTRODUCTION

## Colloque organisé par l'ESPE de l'académie de Créteil – Université Paris-Est Créteil (UPEC)

Avec le soutien du Laboratoire de Didactique André Revuz (LDAR, Universités Paris Diderot, Cergy Pontoise, Paris-Est Créteil, d'Artois, de Rouen), du Laboratoire Éducation Discours Apprentissages (EDA, Université Paris Descartes), du Laboratoire d'informatique LIP6 (Université Pierre et Marie Curie), de l'Agence Nationale pour la Recherche (ANR), de l'Association des Directeurs d'IREM (ADIREM), de la MAIF et de la CASDEN.

### Présentation

La question de l'évaluation des apprentissages des élèves, notamment en mathématiques, est actuellement une préoccupation majeure des différents acteurs de l'éducation, tant au niveau national qu'international. Les évaluations externes se développent avec des objectifs divers. Sur ces questions d'évaluation, le colloque a pour ambition de conjuguer des problématiques de recherche à des problématiques de formation. Au delà des conférences plénières, différents groupes de travail ou ateliers associant chercheurs issus de divers champs (didactique, psychologie, psycho-édumétrie, etc.) et formateurs, intervenant en formation initiale ou continue, à tous les niveaux d'enseignement (primaire, secondaire, supérieur), permettront des échanges et des débats.

Ce colloque propose d'étudier toutes ces questions liées à l'évaluation des élèves, sous ses différentes formes, et à ses conséquences sur l'enseignement. Elles seront traitées dans des groupes de travail et des ateliers répartis selon trois axes thématiques :

1. les différents dispositifs d'évaluation et leurs apports ;
2. l'étude de la validité des dispositifs d'évaluation et de leur contenu ;
3. l'analyse des pratiques d'évaluations en classe.

### Responsables scientifiques

- Brigitte GRUGEON-ALLYS (LDAR, Université Paris-Est Créteil)
- Eric RODITI (EDA, Université Paris Descartes)
- Nathalie SAYAC (LDAR, Université Paris-Est Créteil)

### Comité scientifique

- Michèle ARTIGUE (LDAR, IREM, Université Paris Diderot)
- Lalina COULANGE (LACES, Université de Bordeaux)
- Françoise CHENEVOTOT (LDAR, Université d'Artois)
- Lucie DEBLOIS (Université Laval, Québec)
- Christophe DIERENDONCK (Université de Luxembourg)
- Jean-Luc DORIER (Université de Genève, Suisse)
- Annick FAGNANT (Université de Liège, Belgique)
- Annie FEYFANT (Ifé, ENS Lyon)
- Rémi GOASDOUÉ (EDA, Université Paris Descartes)
- Vanda LUENGO (LIP6 Université Pierre et Marie Curie)
- Aline ROBERT (LDAR, Université de Cergy Pontoise)
- Fabrice VANDEBROUCK (LDAR, IREM, Université Paris Diderot)

### Comité d'organisation

- Nadine GRAPIN (LDAR, Université Paris-Est Créteil)
- Mariam HASPEKIAN (EDA, Université Paris Descartes)
- Julie HOROKS (LDAR, Université Paris-Est Créteil)
- Julia PILET (LDAR, Université Paris-Est Créteil)
- Evelyne SCARON (Université Paris Diderot)
- Stéphane SIREJACOB (LDAR)
- Marc VANTOUROUT (EDA, Université Paris Descartes)

### Responsables des ateliers

**Axe 1** : Pascale MASSELOT (LDAR, Université Cergy Pontoise), Frédérick TEMPIER (LDAR, Université Cergy Pontoise), Laura WEISS (Université de Genève).

**Axe 2** : Rémi GOASDOUÉ (EDA, Université Paris Descartes), Nadine GRAPIN (LDAR, Université Paris Est Créteil), Marc VANTOUROUT (EDA, Université Paris Descartes).

**Axe 3** : Mariam HASPEKIAN (EDA, Université Paris Descartes), Julie HOROKS (LDAR, Université Paris-Est Créteil), Julia PILET (LDAR, Université Paris-Est Créteil).

## Lundi 21 novembre 2016, matin

9h	Ouverture officielle par Madame Béatrice GILLE, Rectrice de l'académie de Créteil et chancelière des universités, et par Madame Brigitte MARIN, Directrice de l'ESPÉ de l'académie de Créteil-UPEC.
9h15 - 10h15	Conférence d'ouverture de Sylvie COPPÉ « <i>Questions posées par l'évaluation dans l'enseignement</i> »
10h15 - 10h45	Conférence d'Antoine BODIN « <i>Comparaison d'évaluations internationales en mathématiques</i> »
10h45 - 11h15	<i>Pause</i>
11h15 - 12h45	GROUPES DE TRAVAIL - SESSION 1 (groupes 1, 2, 3 en parallèle) :  <b>Axe 1 « Dispositifs d'évaluation »</b> <b><i>Dispositifs d'évaluation dans des filières spécifiques</i></b> <ul style="list-style-type: none"><li>● Réussir en calcul professionnel en passant par l'enseignement situé - Nadine KIPFER, Ursula SCHARNHORST, Hansruedi KAISER</li><li>● Sur un type d'évaluation en mathématiques : les interrogations orales en C.P.G.E. - Marc LALAUDE-LABAYLE</li><li>● Un dispositif d'évaluation des compétences des candidats à l'apprentissage à Genève (EVA) - Weimar AGUDELO, Laura WEISS</li></ul> <b>Axe 2 « Validité des dispositifs »</b> <b><i>Numérique, littératie et validité</i></b> <ul style="list-style-type: none"><li>● Transition vers le numérique de l'enquête CEDRE Maths fin de d'école - Jean-Marc PASTOR</li><li>● L'utilisation de l'environnement numérique MathemaTIC pour développer les compétences dans le domaine de la résolution de problèmes arithmétiques - Ben HAAS, Vincent KOENIG, Yves KREIS, Romain MARTIN</li><li>● De la lecture dans les exercices PISA de culture mathématique - Franck SALLES</li></ul> <b>Axe 3 « Pratiques d'évaluation »</b> <b><i>(Re)configurations et pratiques d'évaluation</i></b> <ul style="list-style-type: none"><li>● Analyse des pratiques d'évaluation formative supportées par les technologies - Monica PANERO, Gilles ALDON</li><li>● Variations des pratiques enseignantes de l'évaluation en mathématiques, la place des contenus disciplinaires - Elisabet MENOVAR, Dominique LAHANIER REUTER</li><li>● Tâche(s) complexe(s) : un outil pour évaluer quelle(s) compétence(s) ? François AUBIN, Lalina COULANGE</li></ul>
12h45 - 14h	<i>Déjeuner</i>

## Lundi 21 novembre 2016, après-midi

**14h - 15h** Conférence de Paul DRIJVERS (Axe 1)  
« *Outils d'évaluation et automatisation* »

**15h - 16h30** GROUPES DE TRAVAIL - SESSION 2 (groupes 1, 2, 3 en parallèle) :

**Axe 1 « Dispositifs d'évaluation »**  
***Dispositifs d'évaluation à grande échelle***

- Transition vers le numérique : quelles implications pour l'évaluation des élèves et leurs apprentissages ? L'exemple de l'environnement MathemaTIC au Luxembourg - Philippe ARZOUMANIAN, Amina KAFAIEL-KHORASSANI, Ben HAAS
- Autoévaluation des élèves de l'école primaire en mathématiques à l'aide d'un outil assisté par ordinateur-tablette - Denise VILLANYI, Antoine FISCHBACH, Philipp SONNLEITNER, Christina SIRY, Romain MARTIN
- 20 ans d'évaluations à la DEPP : les acquis des élèves en proportionnalité - Stéphane HERRERO

**Axe 2 « Validité des dispositifs »**  
***Différentes approches de la validité***

- Validité d'une évaluation : une approche épistémologique-didactique - Nadine GRAPIN, Brigitte GRUGEON-ALLYS
- Étude de la robustesse d'une ressource à partir de l'évaluation des connaissances des élèves : quelle méthodologie ? Pour quels résultats ? - Éric MOUNIER, Nadine GRAPIN
- Analyse de validité psycho-didactique d'épreuves d'évaluations internes en classe de Cours Préparatoire - Christophe BLANC
- Validation des épreuves externes en mathématiques dans le domaine des figures et solides - Natacha DUROISIN, Marc DEMEUSE

**Axe 3 « Pratiques d'évaluation »**  
***Analyses de pratiques d'évaluation en classe et pour la classe***

- Analyser les pratiques évaluatives d'enseignants du 2<sup>nd</sup> degré en algèbre: Méthodologie et premiers résultats - Julia PILET, Julie HOROKS
- Soutien aux enseignants dans leur pratique d'évaluation en mathématiques : effets sur les compétences des élèves - Michiel VELDHUIS, Marja van den HEUVEL-PANHUIZEN, Xiaoyan ZHAO
- Les pratiques d'évaluation en mathématiques des professeurs des écoles en France : quelques résultats à partir d'une recherche collaborative - Nathalie SAYAC
- L'évaluation formative dans les interactions en classe de mathématiques: une approche didactique - Michella KIWAN, Mariam HASPEKIAN, Éric RODITI

**16h30 - 17h** *Pause*

**17h - 18h** Conférence de Nathalie LOYE (Axe 2)  
« *Le concept de validité entre évolution, définitions et opérationnalisation* »

## Mardi 22 novembre 2016, journée

9h - 10h	Conférence de Lucie MOTTIER-LOPEZ (Axe 3) « <i>Quand des enseignants de l'école primaire évaluent les résolutions de problèmes mathématiques de leurs élèves</i> »
10h - 10h30	Conférence d'André ANTIBI « <i>De la constante macabre à l'évaluation par contrat de confiance (EPCC)</i> »
10h30 - 11h	<i>Pause</i>
11h - 13h	<b>GROUPES DE TRAVAIL - SESSION 3 (groupes 1, 2, 3 en parallèle) :</b>  <b>Axe 1 « Dispositifs d'évaluation »</b> <b><i>Dispositifs de formation et d'évaluation</i></b> <ul style="list-style-type: none"><li>● L'évaluation des professeurs des écoles débutants dans le master MEEF. État des lieux et pratiques de formateurs en mathématiques - Valentina CELI, Pascale MASSELOT, Frédérick TEMPIER</li><li>● Dispositif de formation pour les enseignants du 1<sup>er</sup> degré autour des évaluations en mathématiques - Nathalie LECLERC, Sophie RIARD</li><li>● Des pistes pour préparer les élèves à la réussite du CE1D en mathématiques - Christine GÉRON, Emmanuelle MASSON</li><li>● Un dispositif de formation basé sur l'EPCC et ses différentes variantes - Grégory QUIQUEMPOIS</li></ul> <b>Axe 2 « Validité des dispositifs »</b> <b><i>Validité des évaluations à grande échelle</i></b> <ul style="list-style-type: none"><li>● Déterminants des performances scolaires. Zoom sur les élèves issus de l'immigration en France et au Royaume Uni - Mohammed BIJOU</li><li>● Transition vers le numérique de l'enquête CEDRE Maths fin de troisième: quelles évolutions au service de l'apprentissage des élèves et des pratiques enseignantes ? Philippe ARZOUMANIAN</li><li>● Évaluation des compétences des jeunes en numératie lors de la Journée Défense et Citoyenneté - Stephane HERRERO, Thomas HUGUET</li></ul> <b>Axe 3 « Pratiques d'évaluation »</b> <b><i>Évaluation des apprentissages de contenus ciblés</i></b> <ul style="list-style-type: none"><li>● Évaluations entre pairs en mathématiques dans le cadre d'un problème complexe : analyse de l'évolution de réponses d'élèves - Marianne MOULIN, Sylvie COPPE</li><li>● La fabrique de l'évaluation. Étude de cas d'une séance de mathématiques dans l'enseignement agricole - Laure MINASSIAN</li><li>● Un dispositif organisé à partir de l'identification des besoins d'apprentissage des élèves (14-15 ans) pour réguler l'enseignement en algèbre élémentaire - Brigitte GRUGEON-ALLYS, Françoise CHENEVOTOT</li><li>● Méthode d'apprentissage et évaluation en résolution de problèmes d'arithmétiques : outils progressifs pour développer la flexibilité des stratégies en résolution - Calliste SCHEIBLING-SÈVE, Arbya EICHI, Elena PASQUINELLI, Emmanuel SANDER</li></ul>
13h - 14h30	<i>Déjeuner</i>
14h30 - 16h	<b>Table ronde</b> Questions aux conférenciers des trois axes.
16h - 17h	<b>Conférence de clôture</b> , en lien avec les résultats du projet NéoPraEval Brigitte GRUGEON-ALLYS, Éric RODITI et Nathalie SAYAC

# RÉSUMÉ DES CONFÉRENCES

## « Questions posées par l'évaluation dans l'enseignement »

**Sylvie COPPÉ, Université de Genève, FPSE, équipe DiMaGe**

Les travaux sur l'évaluation dans ses différents aspects sont nombreux, les entrées sont diverses (docimologie, évaluations internationales, évaluation formative, effets sur la motivation, liens avec l'orientation des élèves, etc). Dans cette conférence, nous nous centrerons sur les liens entre évaluations et apprentissages dans le cadre de l'enseignement primaire ou secondaire en prenant le point de vue de la didactique des mathématiques. Plus précisément, il s'agira d'une part de déterminer comment une entrée par la didactique des mathématiques avec ses cadres théoriques, ses outils et méthodes peut trouver une place dans les analyses ou éclairer d'un jour nouveau les résultats produits dans le cadre des recherches sur l'évaluation. D'autre part, nous chercherons comment des questions posées par l'évaluation dans ses différents aspects peuvent renouveler certains questionnements didactiques.

## « Pour une lecture didactique des enquêtes internationale »

**Antoine BODIN, IREM de Marseille**

Les systèmes éducatifs et en particulier les curriculums sont de plus en plus sous l'influence des enquêtes internationales. Avec l'écrit, les mathématiques constituent la discipline la plus concernée par ces enquêtes et celle qui est la plus objet de rapports, de commentaires et de recherches complémentaires.

En 2015, la France a participé à l'enquête PISA (la cinquième depuis l'année 2000) ainsi qu'aux enquêtes TIMSS pour les élèves de CM1 et de terminale S. Les résultats seront rendus publics fin novembre et début décembre et l'on peut déjà s'attendre à des réactions et à des commentaires plus ou moins informés et plus ou moins orientés. On peut aussi s'attendre à ce que les formateurs d'enseignants et l'encadrement soient incités à s'intéresser de près à ces résultats.

Avec le CNESCO, nous préparons une étude assez fournie sur ces enquêtes, étude qui s'éloigne de l'observation souvent stérile des échelles et des palmarès pour s'intéresser aux contenus évalués et à la façon dont ils le sont. Par un heureux hasard, ce rapport sera publié au moment du colloque. L'étude analyse et compare les cadres de référence des enquêtes PISA et TIMSS et analyse les exercices d'évaluation utilisés. Elle est pensée pour essayer de faire que ces enquêtes, tout critiquables qu'elles puissent être, soient utiles aux formateurs et aux enseignants. Les outils utilisés pour l'analyse des questions sont résolument empruntés à la didactique des mathématiques (la taxonomie de la complexité, les niveaux de fonctionnement des connaissances etc.) et l'étude montre comment ces outils permettent d'apporter un éclairage nouveau sur les résultats des enquêtes.

La communication sera essentiellement consacrée à une présentation de l'étude.

## « Outils d'évaluation et automatisation »

**Paul DRIJVERS, Institut Freudenthal, Université d'Utrecht, CITO, centre d'évaluation, Arnhem, Pays-Bas**

De plus en plus, l'évaluation des apprentissages des élèves, notamment en mathématiques, se déroule sous forme numérique. Pourtant, cette évolution n'est pas sans problèmes. Par des restrictions techniques, on risque de limiter les élèves à des tâches offrant très peu d'espace pour s'exprimer en mathématiques d'une façon authentique. Quels critères, alors, pour un outil d'évaluation des aptitudes mathématiques? Quelles options pour l'automatisation de l'évaluation? Quel mode d'exploitation pour le retour aux élèves? Ces questions seront abordées en partant des expériences concernant l'évaluation sur support numérique aux Pays-Bas, pour les élèves de 15 ans comme pour les étudiants au niveau universitaire.

### **« Le concept de validité entre évolution, définitions et opérationnalisation »**

**Nathalie LOYE, Université de Montréal**

Dès le début du XXe siècle, Binet, sans toutefois utiliser le terme de validité, présentait une conception pragmatique, utilitaire et empirique de la pertinence d'un test. Depuis, le concept de validité s'est imposé, a évolué et a fait l'objet de très nombreux écrits. Il se décline ainsi au fil du dernier siècle en une variété d'acceptions, certaines en tout ou en partie théoriques, et d'autres plutôt opérationnelles. Des auteurs, tels par exemple Cronbach ou Messick, ont grandement influencé l'évolution du concept de validité. Cette conférence vise d'abord à explorer de manière critique cette évolution au fil du temps à la lumière des auteurs clés. Dans un second temps, elle propose une présentation des facettes les plus actuelles du concept de validité ainsi que des démarches pour les explorer en pratique lorsqu'il s'agit d'assurer la qualité de l'évaluation.

### **« Quand des enseignants de l'école primaire évaluent les résolutions de problèmes mathématiques de leurs élèves »**

**Lucie MOTTIER-LOPEZ, Université de Genève**

Après un bref rappel des principaux choix théoriques que nous retenons pour étudier les pratiques évaluatives des enseignants en contexte de classe, nous centrerons notre propos sur la question sensible de « ce par rapport à quoi » l'enseignant apprécie les productions mathématiques de ses élèves. Ce faisant, nous interrogerons plus particulièrement la référentialisation au cœur de l'évaluation des apprentissages des élèves, en tant que processus de construction de significations s'appuyant sur des référents formels et des référents contextuels émergents qui se spécifient. Nous nous interrogerons alors sur les liens et les ruptures possibles entre l'évaluation formative, ancrée dans des normes et pratiques de microcultures de classe construites plus ou moins explicitement entre l'enseignant et ses élèves, et les contraintes de l'évaluation certificative quand il s'agit pour l'enseignant d'attribuer des notes chiffrées attestant des connaissances et compétences mathématiques apprises par ses élèves dans sa classe.

### **« De la constante macabre à l'évaluation par contrat de confiance »**

**André ANTIBI, Université Paul Sabatier, Toulouse**

Actuellement en France et dans quelques pays qui s'inspirent du modèle français, le système éducatif est paralysé par un très grave dysfonctionnement : sous la pression de la société les enseignants se sentent obligés, inconsciemment, de mettre un certain pourcentage de mauvaises notes, une constante macabre en quelque sorte, même dans les classes de très bon niveau, pour que leur évaluation et leur enseignement soient crédibles. Ce dysfonctionnement est actuellement reconnu par le Ministère et par pratiquement tous les partenaires de notre système éducatif, dans l'enseignement public et dans l'enseignement privé. Un système d'évaluation destiné à éradiquer ce phénomène, très facile à mettre en place, a été expérimenté pendant trois ans, de 2005 à 2007. Il s'agit du système d'évaluation par contrat de confiance (EPCC). Il est déjà mis en pratique par plusieurs dizaines de milliers d'enseignants. Cette méthode d'évaluation repose sur le principe de base suivant : responsabilisé par un engagement clair et équitable sur ce qui est attendu de lui, l'élève est amené à prendre conscience du fait que les efforts qu'il fournit ne sont pas vains. D'après de très nombreux témoignages, l'EPCC présente d'autres avantages importants : amélioration des relations entre les enseignants et les élèves, amélioration du bien-être des élèves... ; dans le contexte « EPCC » les élèves travaillent beaucoup plus.

## Session 1 - Lundi 21 novembre 2016, 11h15-12h45

### AXE 1 - DISPOSITIFS D'ÉVALUATION

#### *Dispositifs d'évaluation dans des filières spécifiques*

##### **Réussir en calcul professionnel en passant par l'enseignement situé**

**Auteur(s)** : Nadine KIPFER, Ursula SCHARNHORST, Hansruedi KAISER

**Mots clés** : calcul professionnel, prédictivité des tests, enseignement situé

Les tests d'entrée en mathématiques, administrés par les écoles professionnelles dans la première année de formation, visent à évaluer l'aptitude des apprentis à suivre l'enseignement en calcul professionnel. Étant donné les résultats souvent insuffisants, un grand nombre de cours d'appui en calcul sont proposés. Toutefois, une étude pilote de Wüthrich (2015) a suggéré qu'un test standardisé évaluant les connaissances de base en mathématiques sous-estime les préconnaissances des apprentis. En effet, il ne permet pas de prédire s'ils doivent fréquenter un cours d'appui ou s'ils sont capables de suivre l'enseignement du calcul professionnel. Sur les 34 apprentis qui ont passé le test, 27 apprentis n'atteignaient pas la norme pour le groupe en question. Avant de les orienter vers les cours d'appui, l'enseignant a donné en classe une séquence d'enseignement situé en 8 étapes (Kaiser, 2013).

L'évaluation de cet enseignement a montré que 30 apprentis sur 34 ont acquis les compétences envisagées, ce qui remet la prédictivité du test standardisé en cause. L'objectif du projet est donc d'évaluer si une didactique située focalisant des situations d'application typiques de mathématiques dans les métiers concernés est plus appropriée pour évaluer les réelles compétences des apprentis et ainsi mieux sélectionner les individus qui nécessitent des cours d'appui. L'échantillon sera composé de 12 enseignants de différentes écoles professionnelles et de trois métiers différents. Ces enseignants administreront dans un premier temps les tests d'entrée habituels pour obtenir une vue d'ensemble des résultats des apprentis, ensuite ils effectueront une séquence d'enseignement situé dont ils évalueront les effets à la fin. Les résultats aux deux tests seront comparés. La communication portera sur la présentation des concepts théoriques sous-jacents au projet, les détails de l'étude pilote ainsi que les étapes du projet.

##### **Sur un type d'évaluation en mathématiques : les interrogations orales en C.P.G.E.**

**Auteur(s)** : Marc LALAUDE-LABAYLE

**Mots clés** : C.P.G.E., enseignement supérieur, transition, évaluation orale, T.S.D., sémiotique

Les C.P.G.E., acronyme de Classes Préparatoires aux Grandes Écoles, constituent l'une des voies de l'enseignement supérieur scientifique dont l'objectif principal est de préparer les étudiants à intégrer l'une des Grandes Écoles. Au cours des deux années de préparation en C.P.G.E., en plus des cours et T.D., des devoirs à la maison et des devoirs surveillés hebdomadaires, les étudiants sont évalués au cours d'interrogations orales menées par un enseignant « extérieur » à la classe.

Dans un premier temps, nous nous proposons de présenter ces évaluations en mathématiques. Plus précisément, nous détaillerons les profils des équipes d'enseignants qui interviennent, les questions qui peuvent être posées aux étudiants, les façons dont les enseignants peuvent conduire leurs interrogations orales ...

Dans un second temps, devant la complexité des savoirs et connaissances manipulés à ce niveau d'enseignement, l'évaluation et l'analyse des raisonnements produits nécessitent des outils adaptés. Pour ce faire, la T.S.D. et la sémiotique de Peirce constituent deux des cadres que nous articulons via le modèle d'analyse des raisonnements de Bloch et Gibel (2011) et un diagramme d'argumentation sémiotique.

##### **Un dispositif d'évaluation des compétences des candidats à l'apprentissage à Genève (EVA)**

**Auteur(s)** : Weimar AGUDELO, Laura WEISS

**Mots clés** : Certification des acquis, docimologie, évaluation informatisée

Pour permettre aux entreprises formatrices d'apprentis à Genève de choisir parmi les candidats en fonction de leurs compétences scolaires de base, un test EVA a été mis sur pied par l'Office d'orientation professionnelle et continue (OFPC), s'adressant aux jeunes ne sortant pas directement de l'école obligatoire. Ce test portant sur les mathématiques et le français doit garder un parallélisme avec les évaluations communes (evacom, Tronc Commun) que tous les élèves passent à Genève dans leur dernière année d'école obligatoire (degré 9, 15 ans). En outre, pour permettre des passations collectives nombreuses et éviter de faire appel à des correcteurs, EVA doit pouvoir être passé et corrigé par ordinateur.

Pour concevoir et créer ce test de mathématiques, une équipe composée de didacticiens, de docimologues et ponctuellement d'enseignants concepteurs des evacom a été mise sur pied. À partir du plan d'études de mathématiques et de questions des evacom, un très large panel d'exercices a été élaboré et adapté aux contraintes propres au test informatisé. Ces exercices ont été prétestés auprès d'un public d'élèves aux caractéristiques relativement semblables à celles du public cible. Après plusieurs étapes, comportant chacune des analyses statistiques classiques, plus des analyses de la Théorie de la Réponse aux Items (TRI) à deux niveaux (difficulté et discrimination) et de généralisabilité et des discussions avec des experts de didactique et d'évaluation, l'épreuve a été clonée pour en produire trois versions équivalentes. En effet, il ne s'agit pas, au contraire des evacom, de rédiger chaque année un nouveau test EVA, mais bien de combattre les risques de fuites.

Depuis 2012, le test EVA a été passé à Genève par plus de 4000 jeunes. Il donne grande satisfaction aux entreprises formatrices, qui sont toujours plus nombreuses à obliger leurs candidats visant un apprentissage à le passer.

## AXE 2 - VALIDITÉ DES DISPOSITIFS

### *Numérique, littérature et validité*

#### Transition vers le numérique de l'enquête CEDRE Maths fin de d'école

**Auteur(s)** : Jean-Marc PASTOR

**Mots clés** : évaluation, école, CEDRE

L'enquête CEDRE en mathématiques menée par la Direction de l'évaluation, de la prospective et de la performance (DEPP) vise à évaluer les connaissances et les compétences des élèves en fin d'école-primaire, au regard des objectifs fixés par les programmes. Le protocole d'évaluation 2014 est délivré principalement sous forme de tests « papier-crayon ». Néanmoins, une partie « numérique » a été ajoutée, qui comprend des items repris à l'identiques - dématérialisés - sur les deux supports et des items spécifiques au format numérique. Les techniques de l'information et de la communication modifient les compétences devant être mises en œuvre par les élèves. Il est important d'observer ce que les techniques numériques garantissent comme continuité par rapport aux techniques traditionnelles ; ce qu'elles proposent comme nouveautés et ce qu'elles ne peuvent pas remplacer. Cette communication sera scindée en deux parties. La première partie consistera à présenter les résultats de l'évaluation CEDRE Mathématiques en fin d'école de mai 2014 pour sa partie numérique en la mettant en regard des résultats de la partie papier [Etienne Dalibard et Jean-Marc Pastor, Note d'information n°18, mai 2015]. La seconde partie présentera à partir d'items « dématérialisés » une comparaison des performances des élèves sur les deux supports prolongeant ainsi les premiers résultats [Pascal Bessonneau, Education & Formation n°86-87, mai 2015, p.159-179]. En guise de conclusion, nous ferons un premier état des variables qui influent sur la réussite aux items.

#### L'utilisation de l'environnement numérique MathemaTIC pour développer les compétences dans le domaine de la résolution de problèmes arithmétiques

**Auteur(s)** : Ben HAAS, Yves KREIS, Vincent KOENIG, Romain MARTIN

**Mots clés** : numérique, apprentissage, mathématiques, problèmes arithmétiques, compétences

Dans le cadre du « Digital Lëtzebuerg » et suite à l'appel Horizon 2020 par la commission européenne, le gouvernement du Luxembourg a lancé des initiatives pour l'implémentation du numérique dans ses différents ministères. Dans ce contexte, l'éducation nationale a initié le développement d'un environnement numérique pour l'apprentissage des mathématiques à l'école fondamentale pour les élèves de cycle 4, Université du c'est-à-dire âgés de 10 à 12 ans.

Suite au constat de faibles performances en mathématiques aux épreuves standardisées de début de cycle 3 (élèves de 8 à 10 ans), la question d'élargir l'implémentation de cet environnement numérique à ce cycle s'avère opportune.

En effet, les résultats à ces épreuves indiquent que d'une part un quart de ces élèves n'arrive pas à atteindre le niveau socle et que d'autre part les résultats ne se sont que peu améliorés entre 2011 et 2013. En d'autres termes, il y aurait donc un quart de la population scolaire de cycle 3 qui n'aurait pas les compétences nécessaires pour suivre les apprentissages prévus en mathématiques. Les élèves concernés sont majoritairement d'un milieu socio-économique faible et issus de la migration.

Ainsi, une étude est envisagée pour voir dans quelle mesure l'usage de l'environnement MathemaTIC favoriserait le développement des compétences en mathématiques, notamment dans le domaine de la résolution de problèmes arithmétiques.

Cette étude se réalisera entre 2016 et 2018 avec un échantillon de 20 classes (400 élèves). Les items seront construits en partant des niveaux testés dans les Ep.Stan 3.1 et en se basant sur le plan d'études du cycle 3. Pour ce faire, l'approche adoptée est celle dite « centrée utilisateur », issue du domaine des Interactions Hommes-Machines (IHM). Les compétences au début et à la fin de l'utilisation-même de l'environnement MathemaTIC par les élèves de cycle 3 seront mesurées par les épreuves standardisées Ep. Stan, sous le couvert du Lucet.

#### De la lecture dans les exercices PISA de culture mathématique

**Auteur(s)** : Franck SALLES

Les exercices PISA de mathématiques sont construits afin de mesurer la performance des élèves de 15 ans dans le domaine de la « literacy » mathématique. Ce concept va au-delà des acquis scolaires et se concentre sur la mise en œuvre des savoirs dans des tâches issues du monde réel. Ainsi la plupart des exercices (items) du test sont contextualisés. Des situations doivent être introduites aux élèves dans des énoncés textuels.

Les items PISA de culture mathématique embarquent donc une tâche de lecture a priori non négligeable tant quantitativement que qualitativement. Cette étude approfondit la nature du lien entre lecture et mathématiques à travers la comparaison des résultats PISA en mathématiques de deux sous-populations d'élèves de France : les moins bons lecteurs et les meilleurs lecteurs. Les écarts de réussite entre ces deux groupes ont été étudiés item par item. Il s'agit de déterminer quelles caractéristiques d'items sont plus ou moins différenciatrices selon le niveau de lecture des élèves. Ces caractéristiques relèvent aussi bien des mathématiques que d'autres critères dont la longueur et la lisibilité des énoncés. Les élèves les moins performants en lecture réussissent moins souvent les items de mathématiques. Concision et lisibilité des textes réduisent l'écart et cet effet se constate encore selon le contenu mathématique ou le contexte de la situation.

L'étude révèle enfin que l'absence de réponse des élèves à un énoncé ne s'explique pas par sa longueur mais plutôt par la complexité de sa lecture.

Cette étude permet donc d'interroger ce qu'évaluent vraiment les items PISA et plus généralement l'évaluation de la compétence mathématique indépendamment de la lecture. Des précautions sont à prendre lors du développement des items afin de réduire la quantité de lecture et d'augmenter la lisibilité des énoncés.

## AXE 3 - PRATIQUES D'ÉVALUATION (Re)configurations et pratiques d'évaluation

### Analyse des pratiques d'évaluation formative supportées par les technologies

**Auteur(s) :** Monica PANERO, Gilles ALDON

**Mots clés :** évaluation formative, technologie, collègue

Le projet européen FaSMEd a pour objectif d'étudier le rôle des technologies dans la mise en œuvre de stratégies d'évaluation formative dans les classes de mathématiques et de sciences. Dans le cadre de ce projet et dans une perspective de "design-based research", nous avons travaillé conjointement avec les partenaires européens à l'élaboration d'un cadre théorique et avec les écoles partenaires à la mise au point et au test de nouvelles stratégies d'évaluation dans les classes. Dans la deuxième année de ce projet chacun des neuf partenaires a conduit des études de cas dans les écoles associées au projet. En nous appuyant sur les définitions de Black et Wiliam (2009) et en considérant à la fois les acteurs, les stratégies d'évaluation formative et les propriétés des technologies, nous avons construit un modèle théorique tridimensionnel permettant d'analyser les séquences d'enseignement construites sur les principes de l'évaluation formative. Dans cette présentation nous proposerons le cadre théorique et nous l'illustrerons par son utilisation sur l'analyse d'études de cas à la fois en mathématiques et dans une perspective co-disciplinaire en mathématiques et physique. Dans les deux cas, l'objet de l'étude a été la notion de représentation graphique mais dans le premier cas dans le cadre des mathématiques et dans le second cas comme objet frontière entre mathématiques et sciences physiques. Les analyses conduites dans des contextes différents et des niveaux de classe distincts ont cependant permis de relever des invariants permettant de mettre en évidence le rôle de la technologie dans la mise en œuvre de situation fondée sur l'évaluation formative.

### Variations des pratiques enseignantes de l'évaluation en mathématiques, la place des contenus disciplinaires

**Auteur(s) :** Elisabet MENOVAR, Dominique LAHANIER REUTER

**Mots clés :** évaluation, pratiques enseignantes, vécu disciplinaire

Nous proposons une étude des pratiques enseignantes de l'évaluation en Mathématiques à l'école primaire. Nous souhaitons montrer les spécificités disciplinaires des pratiques évaluatives en Mathématiques, mais également les variations en fonction des contenus disciplinaires.

Les données sont importantes et variées, issues d'une part d'une thèse (Verfaillie Menouar, en cours) et d'autre part de travaux autour du concept de vécu disciplinaire (Reuter, 2013) : des observations de séances d'évaluation, des questionnaires, des entretiens et des recueils de documents. Il s'agit de décrire et de comparer les pratiques effectives des enseignants, de mettre en lumière ce qui est évalué (discipline, domaines, contenus) et de comprendre la manière dont les élèves vivent les évaluations. Nous montrons que les pratiques d'évaluation en mathématiques occupent une place plus conséquente et prennent des formes plus diverses que dans d'autres disciplines. Puis, nous montrons les effets positifs et négatifs de ces pratiques sur les élèves : image de soi, bien être ou mal-être, plaisir ou souffrance... Enfin, par l'étude des intitulés des documents d'évaluation, nous rendons compte de la configuration de la discipline Mathématiques dans la manière dont les enseignants la construisent quand ils évaluent leurs élèves, en interrogeant la place plus ou moins importante de certains contenus. Nous concluons en avançant que l'étude des pratiques enseignantes de l'évaluation par une entrée disciplinaire permet de penser différemment l'évaluation, non seulement d'un point de vue global, non seulement pour en montrer les effets négatifs, et aussi pour en penser des dispositifs de formation, puisqu'il semble exister des difficultés pour les enseignants à mettre en œuvre l'évaluation de certains contenus disciplinaires.

### Tâche(s) complexe(s) : un outil pour évaluer quelle(s) compétence(s) ?

**Auteur(s) :** François AUBIN, Lalina COULANGE

**Mots clés :** tâche complexe, compétences et connaissances mathématiques, contrat didactique

Le concept de « tâche complexe » est un objet institutionnel dont l'apparition en France (en 2006) est liée à la volonté de développer une évaluation par compétences au sein de l'institution scolaire. L'enquête internationale PISA qui semble avoir été à l'origine de cette volonté institutionnelle, pointe les difficultés particulières des élèves de français dans la production de réponses aux items de mathématiques, correspondant à des tâches complexes. Nous nous intéressons d'une part aux origines potentielles des difficultés d'élèves ainsi constatées, d'autre part, aux définitions multiples de l'objet « tâche complexe » dans l'institution, en lien avec l'évaluation par compétences en mathématiques. Notre recherche qui se centre sur cet objet « tâche complexe » comprend dès lors deux volets. Un premier volet repose sur une expérimentation conduite dans une classe de troisième française (élèves de 14-15 ans) autour d'un problème correspondant à un item libéré de l'enquête PISA de 2003, le « problème du menuisier ». Cette expérimentation démontre que ce problème recouvre des ruptures de contrat didactique diverses, liées à la fois à la forme « textuelle » de l'énoncé et aux connaissances supposées disponibles sur une notion donnée, et nous renseigne plus avant sur le traitement d'une telle tâche complexe par les élèves et leur enseignant dans la classe de mathématiques. Un deuxième volet consiste en une enquête institutionnelle sur l'objet « tâche complexe » et sur les définitions multiples de cet objet liées à sa transposition au sein des différentes ressources (enquête internationale PISA, sites académiques, Projet européen LEMA, manuels scolaires) destinées aux enseignants français. Les transformations ou redéfinitions de l'objet « tâche complexe » constatées nous semblent à même de peser sur l'enjeu institutionnel qui en motive a priori la présence dans les pratiques enseignantes : l'évaluation des compétences en mathématiques.

# Session 2 - Lundi 21 novembre 2016, 15h-16h30

## AXE 1 - DISPOSITIFS D'ÉVALUATION

### *Dispositifs d'évaluation à grande échelle*

#### **Transition vers le numérique : quelles implications pour l'évaluation des élèves et leurs apprentissages? L'exemple de l'environnement MathemaTIC au Luxembourg**

**Auteur(s)** : Philippe ARZOUMANIAN, Amina KAFAIEL-KHORASSANI, Ben HAAS

**Mots clés** : numérique, apprentissage, mathématiques

Afin de favoriser le développement des compétences du 21<sup>ème</sup> siècle, l'apprentissage personnalisé et l'engagement des élèves, l'École luxembourgeoise se met progressivement sur le rail de la technologie. Si cette tendance globale est de plus en plus marquée, force est de reconnaître qu'améliorer la qualité de l'enseignement et de l'apprentissage, en changeant le processus pédagogique à travers le numérique, est une entreprise ambitieuse. Cette présentation portera sur MathemaTIC, une initiative nationale, qui a pour but de développer et de fournir un environnement d'apprentissage numérique adapté aux socles de mathématiques luxembourgeois pour les élèves de Cycle 4 (5<sup>ème</sup> et 6<sup>ème</sup> années d'études) de l'enseignement fondamental. Nous traiterons la richesse et les défis qu'offre l'environnement MathemaTIC et les enjeux pédagogiques et didactiques de la construction des items le composant. Nous mettrons notamment en avant les différents moments d'apprentissage (test diagnostique, parcours personnalisé, test bilan) ainsi que l'accompagnement des enseignants clé de la réussite du projet.

#### **Autoévaluation des élèves de l'école primaire en mathématiques à l'aide d'un outil assisté par ordinateur-tablette**

**Auteur(s)** : Denise VILLANYI, Antoine FISCHBACH, Philipp SONNLEITNER, Christina SIRY, Romain MARTIN

**Mots clés** : autoévaluation, enseignement primaire, testing assisté par ordinateur, analyse corrélacionnelle

Même si en théorie il y a consensus sur les effets bénéfiques de l'autoévaluation, celle-ci est peu pratiquée dans les salles de classe et un doute persiste quant aux capacités des élèves de l'école primaire d'évaluer leurs compétences avec une précision acceptable. Un outil innovateur est au centre de notre projet, notamment un outil d'autoévaluation assistée par ordinateur-tablette, conçu pour les élèves âgés de 9 à 10 ans. Cet outil leur permet d'évaluer leurs compétences en mathématiques (définies dans le « plan d'études » du Luxembourg) et il est facilement applicable en salle de classe. Est-ce que les élèves de l'école primaire, âgés de 9 à 10 ans, peuvent évaluer leurs compétences en mathématiques de manière pertinente ? Telle est la question-clé de cette communication. Une évaluation pertinente implique deux choses : (a) l'autoévaluation des élèves coïncide avec leurs compétences « réelles » et (b) les élèves sont capables de reconnaître les difficultés inhérentes des items présentés. L'outil d'autoévaluation a été appliqué à différents moments auprès de 658 élèves, âgés de 9 à 10 ans, issus de 42 groupes-classe de l'enseignement primaire au Luxembourg. Ces mêmes élèves ont participé aux tests standardisés en mathématiques, faisant parties du monitoring scolaire du Luxembourg. (a) Suite à l'analyse corrélacionnelle des scores moyens des autoévaluations des élèves en mathématiques avec leurs scores de tests standardisés en mathématiques, nous trouvons une corrélation moyenne,  $r = .42$ . (b) L'analyse de dispersion, reprenant les scores moyens des autoévaluations en mathématiques en fonction du niveau de difficulté théorique (consensus d'experts) des items, révèle que les élèves reconnaissent le degré de difficulté des items indépendamment de leur niveau de compétences en mathématiques. Nous en concluons que, globalement les élèves âgés de 9 à 10 ans sont capables d'évaluer leurs compétences en mathématiques d'une manière pertinente et valide.

#### **20 ans d'évaluations à la DEPP : les acquis des élèves en proportionnalité**

**Auteur(s)** : Stephane HERRERO

**Mots clés** : proportionnalité, évaluation, DEPP

Une des missions de la Direction de l'Évaluation, de la Prospective et de la Performance (DEPP) est de concevoir et de mettre en oeuvre des évaluations, des enquêtes et des études sur différents aspects du système éducatif français. Un certain nombre d'entre elles portent plus particulièrement sur les acquis des élèves en mathématiques. L'objet de cette communication est de présenter les acquis des élèves sur la proportionnalité depuis 20 ans aux différentes évaluations de la DEPP. Ces évaluations sont de plusieurs natures (diagnostiques, bilan, comparaison temporelle, suivi longitudinal, pour fournir les indicateurs de la LOLF) et portent sur des programmes et des niveaux différents. La comparaison directe des résultats de ses évaluations n'est ainsi pas possible. Elle peut cependant fournir des indicateurs de réussite à un niveau donné, relativement à un type de compétence donné, une évolution dans le temps ou dans la scolarité. La première partie de l'exposé présentera les différentes évaluations retenues dans cette étude en regard des différents programmes de l'enseignement obligatoire depuis 20 ans : leurs objectifs, les épreuves passées par les élèves, les procédures de passation et de correction. La deuxième partie offrira un rapide panorama du traitement de la question de la proportionnalité aux différents niveaux de l'école élémentaire et du collège sur les 20 dernières années ainsi que les procédures de résolution attendues ou privilégiées. La troisième partie explicitera la méthodologie employée lors de cette étude. Enfin, la quatrième partie proposera les principaux résultats de l'étude ainsi que le prolongement envisagé.

## AXE 2 - VALIDITÉ DES DISPOSITIFS

### *Différentes approches de la validité*

#### **Validité d'une évaluation : une approche épistémologique-didactique**

**Auteur(s) :** Nadine GRAPIN, Brigitte GRUGEON-ALLYS

**Mots clés :** évaluation, évaluation externe, validité, psycho-didactique

Déterminer si une évaluation est valide, c'est-à-dire si elle évalue bien ce qu'elle est sensée évaluer (et uniquement cela) est une question vive, partagée par tout concepteur d'évaluation. L'étude de la validité que nous présentons dans cette communication repose principalement sur une approche épistémologique-didactique, même si dans nos travaux (Grugeon & Grapin 2015, Grapin 2015), nous nous appuyons aussi sur des approches psychodidactiques et psychométriques (dans le cas des évaluations externes à grande échelle). Il s'agit de considérer, dans un domaine mathématique donné, l'ensemble des tâches de l'évaluation au regard d'une organisation mathématique de référence définie à partir d'une étude épistémologique et didactique. La donnée d'une telle référence permet d'effectuer une sélection des tâches de l'évaluation (si l'analyse est réalisée en amont) ou d'analyser l'ensemble des items pour déterminer, par exemple, s'ils couvrent le domaine en étant représentatifs de ce dernier, s'ils sont cohérents avec leur objectif d'évaluation, si le codage de la réponse permet de repérer le raisonnement de l'élève, etc. Nous considérons que l'analyse de validité se joue d'abord à un niveau local avec une analyse a priori de chacune des tâches de l'évaluation. Menée en lien avec la définition de l'organisation mathématique de référence, cette analyse conduit ensuite à une étude globale sur l'ensemble des items pour repérer d'éventuels manques ou redondances dans les types de tâches proposés ou encore une répartition peu équilibrée selon la complexité.

L'analyse de la validité menée à ces deux niveaux permet ensuite une réinterprétation des résultats obtenus à l'évaluation. Après avoir décrit cet outil d'analyse, nous l'exploiterons pour étudier les bilans CEDRE école et collège, dans leur contenu et dans les résultats produits, dans deux domaines différents : les nombres entiers (incluant la numération, le calcul et la résolution de problèmes arithmétiques en fin d'école) et l'algèbre en fin de collège.

#### **Étude de la robustesse d'une ressource à partir de l'évaluation des connaissances des élèves : quelle méthodologie ? Pour quels résultats ?**

**Auteur(s) :** Éric MOUNIER, Nadine GRAPIN

**Mots clés :** évaluation, ressource, connaissances numériques

Les recherches menées en didactique des mathématiques sur l'usage des ressources par les enseignants sont menées principalement à partir des pratiques enseignantes et conduisent à montrer une certaine variabilité de ces dernières. Nous avons choisi dans notre travail d'étudier la robustesse d'une ressource en mettant en perspective les pratiques enseignantes avec les connaissances construites par les élèves à la fin du processus d'enseignement ; nous nous plaçons dans une approche comparative entre des classes qui ont suivi un enseignement dispensé à partir d'une ressource spécifique, développée par Mounier (2010) sur l'enseignement du nombre au CP, et d'autres classes qui ont suivi un enseignement construit à partir d'autres manuels. Nous abordons essentiellement dans la communication la méthodologie avec laquelle nous évaluons les connaissances d'élèves de fin de CP et proposons de décrire la façon dont les évaluations ont été conçues pour être valides. Deux types d'évaluation et de passation interviennent : des évaluations «classiques» menées en classe en papier-crayon sur l'ensemble des élèves et des passations individuelles avec le chercheur sur un échantillon réduit d'élèves. La conception du contenu de ces évaluations (tâches, analyse des réponses...) repose sur une analyse a priori qui apporte des preuves de validité didactique de contenu telles qu'elles ont été définies par Grapin (2015). Le deuxième niveau d'observation, avec le chercheur, permet d'accéder avec un grain d'analyse plus fin, aux procédures effectivement mises en jeu par les élèves pour résoudre les tâches et apporte des preuves de validité complémentaires. Nous montrons ainsi la façon dont ces évaluations peuvent permettre à la fois d'évaluer de façon générale les connaissances des élèves de fin de CP et d'être exploitées plus spécifiquement par le chercheur pour l'étude de la robustesse d'une ressource.

#### **Analyse de validité psycho-didactique d'épreuves d'évaluations internes en classe de Cours Préparatoire**

**Auteur(s) :** Christophe BLANC

**Mots clés :** évaluation interne, analyse de tâches, validité psycho-didactique

Notre communication s'inscrit dans une réflexion sur une évaluation qui soit « au service des apprentissages ». La question à laquelle nous cherchons à répondre est celle de la validité de l'évaluation de ce point de vue : le diagnostic que l'évaluation permet est-il bien fondé ? D'un point de vue didactique, l'évaluation pose en effet la question de sa propre validité. Nous nous attacherons dans cette communication à porter notre attention sur les tâches pour évaluer dans le cadre d'évaluations internes conduites en classe de CP (1<sup>ière</sup> primaire), dans le domaine des programmes intitulé « Nombres et calcul ». Particulièrement, nous nous intéresserons à leur validité psycho-didactique. Rechercher à établir la validité psychodidactique d'une tâche revient à estimer si la tâche permet bien d'engager (de par les contenus mis en jeu, de par leur mise en contexte etc.) une activité de l'élève correspondante à ce que l'enseignant cherche à observer (ce qui est formulé par l'item). Deux parties structureront la communication :

- dans une première partie, nous présenterons brièvement le cadre conceptuel d'analyse de la tâche que nous avons construit. Nous présenterons ensuite la méthodologie d'analyse des tâches d'évaluation que nous avons élaborée en référence au cadre précédemment défini.
- dans une deuxième partie, nous présenterons quelques-uns des résultats de nos analyses lesquelles ont été entreprises à partir d'un corpus de 754 épreuves recueillies auprès de 16 enseignants de CP. Nous les organiserons selon certaines composantes de la tâche prescrite aux élèves susceptibles d'avoir une influence sur la validité des épreuves. Nous conclurons notre présentation en soulignant que, au sein des épreuves que nous avons collectées et analysées, les composants mis en lumière semblent insuffisamment pris en compte. De ce fait, la qualité diagnostique de l'évaluation conduite est susceptible d'être faible.

## Validation des épreuves externes en mathématiques dans le domaine des figures et solides

**Auteur(s)** : Natacha DUROISIN, Marc DEMEUSE

**Mots clés** : évaluations externes non certificatives, apprentissages scolaires, mathématiques

L'enseignement de la géométrie à l'école primaire renvoie à deux champs de connaissances distincts, à savoir les connaissances spatiales et les connaissances géométriques (Marchand, 2009 ; Berthelot & Salin, 1993). Alors que de nombreux auteurs (i.e. Wai et al., 2009) ont mis en évidence que les connaissances spatiales étaient un prédicteur important de réussite dans différents domaines (mathématiques, scientifiques, technologiques, ingénierie...) et également indispensables à l'exercice de nombreuses professions plus techniques, l'acquisition de ce type de connaissances par les élèves reste problématique.

L'analyse menée à partir de plusieurs évaluations externes non certificatives dans les domaines des « figures et solides » et des « grandeurs » a permis, d'une part, d'identifier les compétences et mécanismes spatiaux les plus difficilement acquis (décentration, visualisation spatiale, rotation mentale) par les élèves de l'enseignement fondamental et du début de l'enseignement secondaire en Fédération Wallonie-Bruxelles. D'autre part, elle a conduit à s'interroger sur la validité des évaluations proposées. Alors que l'objectif est d'évaluer les acquisitions des élèves dans les domaines des « solides et figures » et des « grandeurs », il apparaît que les évaluations ne prennent pas en compte toutes les dimensions de cet apprentissage en mathématiques ; on remarque en effet que les évaluations portent essentiellement sur les connaissances géométriques aux dépens des connaissances spatiales.

## AXE 3 - PRATIQUES D'ÉVALUATION

### *Analyses de pratiques d'évaluation en classe et pour la classe*

#### Analyser les pratiques évaluatives d'enseignants du 2<sup>nd</sup> degré en algèbre : méthodologie et premiers résultats

**Auteur(s)** : Julia PILET, Julie HOROKS

**Mots clés** : évaluation formative, algèbre, oratiques enseignantes

Dans le cadre du projet ANR Néopraéval qui s'intéresse à l'évaluation en mathématiques, nous nous penchons sur les pratiques évaluative des enseignants du 2<sup>nd</sup> degré en mathématiques et cherchons à en construire un cadre d'analyse (Robert et Rogalski, 2002). Notre approche de l'évaluation prend en compte à la fois les spécificités des contenus enseignés, ici l'algèbre élémentaire du collège (Grugeon & al., 2012), et les activités de l'enseignant en classe pour réguler les savoirs (Allal & Mottier-Lopez, 2007). Nous nous intéressons plus spécifiquement à la manière dont ils impliquent les élèves dans leurs apprentissages et dont ils prennent en compte leurs productions. Nous tenons compte aussi du contexte du métier d'enseignant et des contraintes qu'il impose (public ZEP, demandes institutionnelles par rapport à l'évaluation). Ce travail s'effectue en grande partie au sein d'un Lieu d'éducation associé (Léa) dans et autour duquel s'organise un travail collaboratif entre enseignants et chercheurs.

Nous présenterons nos réflexions sur la construction de cette méthodologie de recherche et les questions que nous nous posons pour la développer, ainsi que nos analyses d'une partie des données recueillies pendant les deux premières années d'existence du Léa.

#### Soutien aux enseignants dans leur pratique d'évaluation en mathématiques : effets sur les compétences des élèves

**Auteur(s)** : Michiel VELDHUIS, Marja van den HEUVEL-PANHUIZEN & Xiaoyan ZHAO

**Mots clés** : évaluation formative, workshops, enseignants

Dans le cadre de deux études dans deux pays – les Pays-Bas et la Chine – nous avons étudié les effets sur l'apprentissage des élèves, d'une formation des enseignants en mathématiques sur leur pratique d'évaluation. L'échantillon de l'étude Néerlandaise était composé de 616 élèves et de 30 enseignants de CE2 ; celui en Chine était constitué de 3040 élèves et de 47 enseignants de CE2. Dans les deux pays, les enseignants étaient répartis dans des conditions expérimentales ou de contrôle. Les enseignants des conditions expérimentales participaient à des workshops sur l'utilisation de techniques d'évaluation en mathématiques. Ces techniques avaient pour but de permettre aux enseignants de découvrir les raisonnements des élèves et ainsi d'adapter leurs enseignements aux besoins des élèves. Étant donné que les deux curricula avaient des points de focalisation différents, aux Pays-Bas les techniques portaient sur le domaine des nombres, tandis qu'il s'agissait, en Chine, de la multiplication des nombres à deux chiffres. Des résultats de tests standardisés de mathématiques ont été pris en compte afin de déterminer les effets sur l'apprentissage des élèves des workshops sur les techniques d'évaluation. Aux Pays-Bas, les élèves des enseignants qui avaient participé à trois workshops d'une heure sur l'utilisation des techniques d'évaluation, s'étaient améliorés significativement plus, en terme de compétences mathématiques, que les élèves des autres conditions ( $d = 0.26$ ). En Chine, il n'y a eu, à première vue, aucun effet sur le gain en performance des élèves (les analyses sont en cours). Les résultats de ces deux études montrent que soutenir la pratique d'évaluation des enseignants peut contribuer à l'amélioration des compétences en mathématiques des élèves. Prenant en compte les perspectives culturelles différentes, ceci nous permet de confirmer, qu'il est primordial de porter une attention particulière à la façon dont les enseignants utilisent l'évaluation dans leurs cours de mathématiques.

## **Les pratiques d'évaluation en mathématiques des professeurs des écoles en France : quelques résultats à partir d'une recherche collaborative**

**Auteur(s)** : Nathalie SAYAC

**Mots clés** : didactique des mathématiques, pratiques d'évaluation, tâches

Les programmes de l'école primaire française sont conçus pour permettre l'acquisition de connaissances, et de compétences visant la réussite de tous les élèves, dans leur diversité. Dans ce cadre, les professeurs ont toute liberté pour que ces apprentissages se réalisent et pour évaluer leurs élèves, mais comment les évaluent-ils réellement ? Selon quels critères et quelles modalités ? Pour mieux connaître les pratiques évaluatives des professeurs des écoles nous avons conçu une recherche collaborative (Desgagné, 1997) menée spécifiquement en mathématiques, au niveau de l'école primaire. Nous rendrons compte des résultats de cette étude dans cette présentation.

## **L'évaluation formative dans les interactions en classe de mathématiques : une approche didactique**

**Auteur(s)** : Michella KIWAN, Mariam HASPEKIAN, Éric RODITI

**Mots clés** : didactique des mathématiques, évaluation formative, interactions verbales

La communication présentée ici porte sur les interventions des enseignants à visée d'évaluation formative en classe de mathématiques. Elle propose de montrer les apports, la portée et les limites d'une approche didactique particulière conçue pour leur analyse différentielle. Après avoir indiqué les orientations théoriques auxquelles nous nous référons, nous précisons ce que nous désignons par évaluation formative dans les interactions en classe. Nous présentons alors l'approche didactique développée pour l'analyse de ces évaluations et nous mettons cette approche à l'épreuve d'une étude de plusieurs vidéos d'enseignement de l'algèbre élémentaire en 2<sup>ème</sup> et 3<sup>ème</sup> année de collège en France et au Liban.

# **Session 3 - Mardi 22 novembre 11h-13h**

## **AXE 1 - DISPOSITIFS D'ÉVALUATION**

### ***Dispositifs de formation et d'évaluation***

## **L'évaluation des professeurs des écoles débutants dans le master MEEF. État des lieux et pratiques de formateurs en mathématiques**

**Auteur(s)** : Valentina CELI, Pascale MASSELOT, Frédéric TEMPIER

**Mots clés** : Master MEEF, professeur des écoles, évaluation de la formation.

Dans le cadre du master MEEF (Métiers de l'Enseignement, de l'Éducation et de la Formation 1<sup>er</sup> degré), concevoir des sujets d'évaluation est l'une des tâches du formateur en mathématiques. À partir d'une réflexion sur les connaissances mathématiques et didactiques utiles pour enseigner les mathématiques à l'école, nous avons élaboré et testé un outil d'analyse (Celi et al., 2016) permettant de questionner le travail d'élaboration de sujets originaux et d'ouvrir sur une réflexion plus générale sur les contenus et modalités de formation (Peltier 1995). À l'occasion de ce colloque, nous souhaitons, d'une part, proposer un état des lieux très général sur les modalités de formation des enseignants stagiaires dans le master MEEF (2<sup>ème</sup> année) en France et, d'autre part, montrer les premiers résultats d'une étude de cas sur les pratiques d'évaluation de quelques formateurs de différents statuts et issus de différentes académies : qu'est-ce qui est évalué ? Comment ? Y a-t-il des décalages avec les intentions des formateurs ? Y a-t-il des tensions liées aux modalités d'évaluations imposées par les maquettes de master ? Pour le premier point, nous nous appuyons sur une étude faite par la COPIRELEM lors du colloque de Besançon (2015) à partir de questionnaires envoyés aux formateurs de mathématiques de différentes ESPE (École Supérieure de Professorat et d'Éducation). Pour le second point, nous nous appuyons sur une étude de cas de quelques formateurs. L'outil d'analyse déjà élaboré nous permet une première analyse des contenus des évaluations proposées par ces formateurs pour dégager les types de connaissances évaluées et les manières de les évaluer. Des entretiens avec les mêmes formateurs complètent ce travail dans le but de mieux comprendre leurs choix, leurs contraintes et de dégager des tensions éventuelles entre les deux.

## Dispositif de formation pour les enseignants du 1<sup>er</sup> degré autour des évaluations en mathématiques

**Auteur(s) :** Nathalie LECLERC, Sophie RIARD

A l'heure où l'évaluation est au cœur des prescriptions institutionnelles, la mise en place d'un dispositif de formation autour de l'évaluation en mathématiques permet de (re)placer l'évaluation au centre des réflexions des enseignants et des formateurs. Ce dispositif s'inscrit dans la dynamique d'une recherche collaborative qui étudie des pratiques évaluatives d'enseignants du premier degré et s'appuie sur un outil d'analyse du niveau de complexité de tâches mathématiques (Sayac & Grapin, 2015) issu de travaux en didactique.

Dans le but de faire évoluer les pratiques, le dispositif prévoit l'appropriation, par les professeurs des écoles, de ce même outil pour analyser des réponses d'élèves extraits d'évaluations externes. Ainsi, les formés pourront découvrir les différents facteurs de complexité inhérents aux exercices du point de vue de l'énoncé, de la tâche mathématiques et du niveau de compétence.

Les enseignants seront amenés à réinvestir ce travail d'analyse sur leurs propres évaluations données en classe afin d'identifier leur « profil » d'évaluateur. De son côté, le formateur prendra en compte ces éléments pour rester proche de la Zone Proximale De développement des Pratiques de chacun (Robert, 2010), c'est-à-dire partir de la pratique personnelle pour la faire progresser. Il réunira des formés de profil semblable pour qu'ils puissent concevoir collectivement une évaluation de fin de séquence sur les nombres entiers.

Enfin, cette nouvelle évaluation commune, conçue en formation, sera proposée sur le terrain et lors d'un nouveau temps de regroupement, la comparaison des réponses des élèves de chaque classe permettra aux formés d'interroger la mise en oeuvre de toute leur séquence d'apprentissage.

A son tour, ce dispositif de formation sera évalué par les formateurs. Il s'appuiera d'une part sur l'analyse d'un « écrit intermédiaire », proposé aux formés en début et en fin de formation et d'autre part, sur l'envoi différé d'une future évaluation conçue par les formés.

## Des pistes pour préparer les élèves à la réussite du CE1D en mathématiques

**Auteur(s) :** Christine GÉRON, Emmanuelle MASSON

**Mots clés :** épreuve externe, consignes, variables didactiques.

Le Certificat d'études du 1<sup>er</sup> degré (CE1D) en Fédération Wallonie-Bruxelles est un certificat sanctionnant la réussite d'une épreuve externe organisée pour tous les élèves terminant le premier degré du secondaire<sup>1</sup>, âgés normalement de 14 ans (maximum 16 ans).

Les résultats obtenus lors de cette épreuve ne sont pas toujours satisfaisants, surtout lorsqu'on les compare aux résultats obtenus deux ans plus tôt par ces mêmes élèves au CEB<sup>2</sup>. Le score moyen dégringole de plus de 15%.

Comment expliquer ces différences ? Comment aider les élèves à surmonter leurs difficultés ?

Ces questions ont constitué le fil conducteur d'un contenu de formation destiné aux enseignants du premier degré du secondaire de la Ville de Liège. Notre objectif était de sensibiliser la trentaine de participants à la nécessité d'entamer avec les élèves un travail en amont de l'épreuve et de leur donner des pistes concrètes pour le réaliser. C'est toute la réflexion menée avec les enseignants que nous allons présenter et illustrer par de nombreux exemples.

Notre contribution s'inscrit donc dans le premier axe (les différents dispositifs d'évaluation et leurs apports) dans le sens où elle tente d'aider les enseignants à construire des dispositifs d'apprentissage et d'évaluation interne qui préparent progressivement les élèves à la résolution de problèmes tels qu'énoncés dans l'épreuve externe.

## Un dispositif de formation basé sur l'EPCC et ses différentes variantes

**Auteur(s) :** Grégory QUIQUEMPOIS

Dans l'académie de Créteil, sous l'impulsion d'un inspecteur, une réflexion a été initiée sur des problématiques liées à l'évaluation des apprentissages des élèves et des formations sont proposées avec pour objectif de permettre aux professeurs de s'en emparer.

Ces formations s'organisent en plusieurs phases, quelles que soient les disciplines des professeurs.

Dans un premier temps, tous les stagiaires sont mis en situation d'évaluation préparée et en cohérence avec la phase d'apprentissage (une dictée préparée de Bernard Pivot dans les conditions réelles), puis dans un second temps, ils doivent faire face à une évaluation non préparée (dictée non préparée qui fait suite à la 1<sup>ère</sup>).

A l'issue de la première phase on recueille le nombre de fautes et on obtient une courbe dite en J : la grande majorité des stagiaires fait un faible nombre d'erreurs. Les résultats de la 2<sup>ème</sup> dictée produisent généralement une courbe de Gauss étonnamment parfaite dès lors que le groupe excède 40 personnes. C'est une entrée en matière pour questionner la représentation que se font les enseignants des résultats d'une évaluation, en mathématiques comme dans les autres disciplines, mais nous rendons compte de ce dispositif de formation dans cette discipline.

Fort de ce constat et dans les phases suivantes, c'est sur l'ensemble d'une séquence en mathématiques que les stagiaires sont invités à travailler : comment transformer une tâche simple en tâche complexe en associant des critères de réussites et diverses modalités d'évaluation (auto-évaluation, évaluation par les pairs) pour une réelle évaluation formative qui doit permettre à l'élève d'identifier ses réussites et les points qu'il doit continuer à travailler.

Se pose ensuite la question de l'articulation entre la phase d'apprentissage et la phase d'évaluation, de telle sorte que l'évaluation ne soit pas piégeuse, mais en totale cohérence avec les compétences travaillées en phase d'apprentissage. C'est alors que l'EPCC et ses différentes variantes (dont le CPE) sont présentées aux stagiaires.

<sup>1</sup> Ce premier degré comporte une 1<sup>ère</sup> et une 2<sup>ème</sup> année commune (1C et 2C). Pour les élèves qui éprouvent des difficultés dans le 1<sup>er</sup> degré commun ou dont on suppose qu'ils éprouvent des difficultés, une orientation est possible vers une 1<sup>ère</sup> ou une 2<sup>ème</sup> année complémentaire (1S ou 2S).

<sup>2</sup> Certificat d'études de base, épreuve externe certificative obligatoire pour tous les enfants de la fin de l'enseignement primaire, âgés habituellement de 12 ans.

## AXE 2 - VALIDITÉ DES DISPOSITIFS

### *Validité des évaluations à grande échelle*

#### **Déterminants des performances scolaires. Zoom sur les élèves issus de l'immigration en France et au Royaume Uni**

**Auteur(s) :** Mohammed BIJOU

**Mots clés :** performances scolaires, immigration, multi-niveaux, Mundlak, PISA2012.

Le présent article se propose d'étudier les déterminants des performances scolaires en faisant un zoom sur les élèves issus de l'immigration en France, en comparaison du cas du Royaume Uni. Les résultats des estimations effectuées à travers un modèle multi-niveaux avec l'adoption de l'approche de Mundlak pour traiter le biais d'endogénéité permet en même temps de confirmer et de nuancer les résultats parus dans le récent rapport de l'enquête PISA 2012 sur les élèves français.

Bien que le score des élèves issus de l'immigration en France soit moins bon comparé à celui des élèves autochtones, la différence de performance entre ces deux groupes est bien inférieure à celle publiée par l'enquête PISA 2012.

Par ailleurs, à la différence de ce qui est avancé dans la même enquête, il ne semble avoir aucune certitude sur le fait que les élèves issus de la deuxième génération d'immigrations soient plus performants que leur camarades issus de la première génération d'immigration, une fois contrôlées les variables pertinentes déterminant les performances scolaires de ces élèves.

#### **Transition vers le numérique de l'enquête CEDRE Maths fin de troisième : quelles évolutions au service de l'apprentissage des élèves et des pratiques enseignantes ?**

**Auteur(s) :** Philippe ARZOUMANIAN

**Mots clés :** mathématiques, numérique, évaluer

Aujourd'hui, la place prépondérante prise par l'informatique questionne l'école sur la transition d'un environnement dominé par le papier vers un environnement dominé par le support numérique.

Pour évaluer les élèves, le numérique offre de multiples possibilités en termes de contenus (format des items, données de navigation des élèves, items interactifs, etc.). Dans cette communication, nous présenterons certaines des évolutions envisagées pour l'enquête de la DEPP, CEDRE mathématiques fin de troisième, à l'horizon 2019. Cette communication sera scindée en trois parties.

La première partie consistera à présenter rapidement les forces et les faiblesses de l'évaluation CEDRE Mathématiques en fin de troisième de mai 2014 sous sa forme actuelle papier-crayon.

La seconde partie présentera quelques exemples d'items mettant en œuvre des animations interactives à l'aide d'un logiciel de géométrie dynamique. La troisième partie mettra en évidence quelques exemples d'items mobilisant un tableur-grapheur. Ces items sont conçus pour évaluer certaines connaissances du programme de l'enseignement secondaire français et permettent de vérifier si les compétences de résolution de problèmes sont acquises par les élèves. Nous présenterons également les différentes étapes qui ont jalonné la conception de l'ensemble de ces exemples en décrivant les enjeux pédagogiques et didactiques ainsi que les liens avec les recommandations institutionnelles françaises actuelles. Nous décrirons ensuite les différentes modalités mises en œuvre pour atteindre les objectifs visés. Notamment, nous mettrons en avant comment certains de ces items ont été expérimentés en utilisant les modèles de réponse à l'item \*Thierry Rocher, Education & Formation n°86-87, mai 2015, p.37-59]. En guise de conclusion, nous donnerons quelques pistes de travail pour que les enseignants intègrent ces formats innovants d'items dans leur pratique pédagogique.

#### **Évaluation des compétences des jeunes en numératie lors de la Journée Défense et Citoyenneté**

**Auteur(s) :** Stéphane HERRERO, Thomas HUGUET

**Mots clés :** numératie, innumérisme, évaluation

Depuis sa création en 1998, la Journée Défense et Citoyenneté (JDC, ex. Journée d'appel de préparation à la défense – JAPD) évalue chaque année les performances en lecture d'environ 700 000 jeunes. Trois générations de tests de lecture se sont succédées. Les deux premières s'appuyaient sur un support papier. Depuis 2009, le test est informatisé et vidéoprojeté. Les jeunes répondent à l'aide d'un boîtier électronique. Cette innovation a simplifié l'organisation des tests et rendu possible l'élargissement des champs évalués lors de la JDC. Après une pré-expérimentation sur 1 153 jeunes, une évaluation complémentaire de numératie a ainsi pu être réalisée, en 2013, sur un échantillon de 56 000 jeunes. Elle mesure leur capacité à mobiliser des compétences numériques et calculatoires dans la vie quotidienne. Elle permet aussi une comparaison avec les performances en lecture. Partagée en quatre profils, la population étudiée est décrite tant sur un plan cognitif que conatif. 9,7 % des jeunes ont des difficultés en numératie. Des difficultés en lecture n'impliquent pas nécessairement en numératie et inversement. Globalement, il apparaît qu'environ 14 % des enquêtés présentent des difficultés dans au moins l'un des deux domaines. Un sous-groupe de 4,8 % de jeunes ne maîtrise que très insuffisamment le calcul primaire. Ils peuvent être considérés comme étant en situation d'innumérisme. L'étude décrit leurs acquis ainsi que leurs lacunes, sources de profondes difficultés au quotidien. Couvrant l'ensemble du territoire métropolitain, grâce au réseau des centres du service national, les résultats font apparaître des disparités régionales. En apportant un éclairage inédit sur la non-maîtrise des mathématiques élémentaires à l'entrée de la vie adulte, cette étude pose la question d'une meilleure prise en compte de la numératie dans les dispositifs scolaires de remédiation, dans un contexte où la priorité est donnée à la maîtrise de la lecture.

## Du curriculum à l'évaluation en Suisse romande: comment caractériser des tâches d'évaluation en référence au Plan d'études romand ?

**Auteur(s)** : Shanoor KASSAM

**Mots clés** : plan d'études, validité, accord interjuge

Avec des pratiques évaluatives très hétérogènes (Wirthner&Ntamakiliro, 2008), la mise sur pied d'un système d'évaluation intercantonale demeure un sujet délicat en Suisse romande. Si un curriculum commun - le Plan d'études romand (PER) - a été adopté en vue d'harmoniser les pratiques, c'est bien celles d'enseignement apprentissage qui étaient visées, et la traduction de ces objets en objets d'évaluation constitue une étape qu'il s'agit encore d'investiguer. Cette communication s'inscrit dans le cadre d'une recherche visant à caractériser les pratiques évaluatives des cantons romands et nous aborderons ici une étape de ce projet : la mise en relation (appariement) de tâches d'évaluation issues des épreuves cantonales de mathématiques avec des éléments du PER (Voir Brown, 1996, sur le Content validity). Nous présenterons les résultats d'un exercice de caractérisation de tâches évaluatives confié à cinq enseignants, ayant tous suivi une formation sur le processus de codage, ainsi que sur les catégories utilisées. Force sera de constater que la consistance dans le codage de certaines tâches est faible. En partant des résultats de l'accord entre les participants (accord interjuge), nous explorerons la question suivante : sur la base de quels critères est-il possible d'attribuer un ou des élément(s) du PER à une tâche d'évaluation ou, à l'inverse, comment caractériser les éléments du PER afin de les traduire en activités d'évaluation ?

### AXE 3 - PRATIQUES D'ÉVALUATION

#### *Évaluation des apprentissages de contenus ciblés*

#### **Évaluations entre pairs en mathématiques dans le cadre d'un problème complexe : analyse de l'évolution de réponses d'élèves**

**Auteur(s)** : Marianne MOULIN, Sylvie COPPE

**Mots clés** : didactique des mathématiques, évaluation entre pairs, argumentation

Dans le cadre du projet européen de recherche ASSIST ME (Assess Inquiry in Science, Technology and Mathematics Education) dont l'objectif est de concevoir de nouveaux dispositifs d'évaluation formative, d'analyser leur influence sur les apprentissages et les pratiques enseignantes en vue de leur diffusion dans le cadre de démarches d'investigations en sciences, mathématiques et technologie nous avons expérimenté successivement dans le courant des années 2014 et 2015 une séquence de classe de mathématiques dans quatre classes de 6e (élèves de 11-12 ans) de collèges de Lyon, basé sur la résolution d'un problème complexe portant sur les fractions et les aires. Dans cette séquence, deux outils d'évaluation formative étaient proposés : évaluation écrite des différentes réponses produites par les élèves, ce qui se rapproche d'une évaluation entre pairs et débat argumentatif sur ces réponses. L'analyse a priori du problème montre qu'il existe deux grandes stratégies de résolution qui peuvent amener à de multiples réponses, sous différentes formes, justes ou fausses, ce qui était l'objectif et qui justifie les méthodes d'évaluation formative. Nous analysons l'évolution du positionnement écrit de chaque élève sur la validité des réponses, avant et après le débat, ainsi que les types d'arguments avancés.. Grâce à l'analyse a priori nous pouvons interpréter les changements ou les résistances. Nous complétons ces analyses par celles d'extraits de la phase collective de discussion des réponses qui montrent qu'il est difficile pour le professeur et pour les élèves de seulement discuter des réponses sans entrer dans le détail des stratégies.

En conclusion, nous nous questionnons sur les conditions et les contraintes des évaluations formatives entre pairs et par le débat argumentatif. Cela nous permettra également d'interroger la pertinence de dispositifs de classe comme les mises en commun.

#### **La fabrique de l'évaluation. Étude de cas d'une séance de mathématiques dans l'enseignement agricole**

**Auteur(s)** : Laure MINASSIAN

**Mots clés** : évaluation, savoirs concrets, apprentissages

Cette communication vise à analyser des processus évaluatifs à partir des perspectives théoriques ouvertes à la croisée de la sociologie et de la psychologie de l'évaluation et de la didactique (Lafortune et Allal, 2008 ; Merle, 1998). Au-delà de partager l'idée selon laquelle l'évaluation ne consiste pas en une notation stricto sensu d'un test ou d'un devoir (Meirieu, 2005), ces travaux montrent en quoi celle-ci s'exerce plus ou moins informellement dans l'ensemble des activités de la classe (Mottier Lopez et Allal, 2008), comment elle se construit à différents moments au cours desquels les élèves interagissent (Rochex, Crinon ; 2011). Cette manière d'envisager de façon large la notion d'évaluation permet de rendre compte d'éléments qui se combinent. Sur la base d'un suivi ethnographique de deux années d'une même classe de l'ancien BEP agricole, aujourd'hui devenu première année du bac pro en trois ans, on montrera en quoi les façons de vivre et d'interpréter les processus scolaires, du côté enseignant et du côté élèves, contribuent à l'évaluation. En prenant en compte trois types d'échelles temporelles et de contexte : 1/ ce que supposent les enseignants des élèves au travers des exercices proposés ; 2/des situations de classe où les élèves sont confrontés à ces mêmes exercices ; 3/l'évaluation sommative qui fait suite, des interprétations qui sous-tendent les modes d'évaluation se dégagent.

La perspective adoptée relève d'une ethnologie de la classe, c'est-à-dire qu'on s'est gardé d'interpréter les pratiques langagières a priori à l'aune d'une norme du parler en classe au profit d'un dispositif méthodologique permettant aux différents protagonistes d'explicitier leurs propres points de vue sur une situation précise. Un des résultats de l'enquête est que certaines manières d'évaluer informellement les difficultés supposées des élèves n'est pas toujours pertinente, et peut constituer un obstacle à l'appropriation de savoirs disciplinaires.

## Un dispositif organisé à partir de l'identification des besoins d'apprentissage des élèves (14-15 ans) pour réguler l'enseignement en algèbre élémentaire

**Auteur(s) :** Brigitte GRUGEON-ALLYS, Françoise CHENEVOTOT

Nous présentons une recherche visant à étudier l'impact d'un dispositif d'enseignement pour réguler les apprentissages en algèbre élémentaire (élèves de 14 à 15 ans). La recherche a été menée par une équipe pluridisciplinaire de chercheurs travaillant sur les projets Lingot (Delozanne et al. 2010), PépiMeP (Grugeon et al., 2012) et actuellement sur le projet ANR Néopraeval « Nouveaux outils pour de nouvelles pratiques d'évaluation »<sup>1</sup>. Ces projets ont pour objectifs la conception et la diffusion d'outils pour aider les enseignants à gérer l'hétérogénéité des élèves dans leur classe en fin de collège.

Fondés sur une étude épistémologique et didactique du domaine algébrique (Grugeon 1997, Kieran 2007, Bosch et Gascon 2005), ils ont conduit à l'élaboration et l'implémentation de Pépite, une évaluation diagnostique automatisée des compétences des élèves en algèbre élémentaire et à des parcours d'enseignement différencié (Pilet, 2012) sur des plateformes numériques<sup>2</sup>.

Les parcours portent sur un objectif d'enseignement commun à la classe et consistent en des séries d'exercices adaptés aux besoins d'apprentissage des élèves repérés par le diagnostic.

Le dispositif d'enseignement permet d'articuler les différentes fonctions de l'évaluation au service des apprentissages des élèves. Il vise à favoriser la reconnaissance et la prise en compte par les enseignants des procédures et des erreurs des élèves pour organiser les phases de formulation et de validation lors d'épisodes de recherche ou de mise en commun et à adapter en conséquence la nature des feedbacks et des aides à apporter aux élèves.

Nous présentons d'abord les fondements didactiques et l'organisation du dispositif d'enseignement. Puis nous rendons compte des évolutions de l'activité algébrique d'élèves de cinq classes ayant suivi cet enseignement pendant un an, plus particulièrement, l'évolution des modes de calcul, de l'usage de l'outil algébrique pour résoudre des problèmes et de leurs niveaux de raisonnement.

## Méthode d'apprentissage et évaluation en résolution de problèmes d'arithmétiques : outils progressifs pour développer la flexibilité des stratégies en résolution

**Auteur(s) :** Calliste SCHEIBLING-SÈVE, Arbya EICHI, Elena PASQUINELLI, Emmanuel SANDER

Le projet ESCRIMS (Esprit Critique en Mathématiques et en Sciences) vise à développer de nouveaux outils progressifs pour améliorer et évaluer les capacités de résolution de problème et de raisonnement d'investigation en sciences (non abordé ici) à l'école primaire.

Comment évaluer l'acquisition de notions mathématiques par les élèves au-delà de compétences algorithmiques ? Une difficulté est que les connaissances naïves sur les concepts mathématiques restreignent les réussites des élèves aux seuls problèmes cohérents avec cette préconception. Par exemple, les élèves associent spontanément la notion de soustraction à celle de retrait (Fischbein, 1989). La réussite d'un élève à un problème de soustraction introduisant une situation de retrait évalue alors essentiellement ses compétences algorithmiques : elle ne permet pas de dissocier l'application de ses connaissances naïves et celle d'une conception reposant sur la recherche d'un écart. En revanche, un problème décrivant sémantiquement une situation de gain mais nécessitant une soustraction pour être résolu permet d'évaluer le degré conceptuel de la notion développée par l'élève (Hofstadter & Sander, 2013). La notion de distributivité a été travaillée dans cette étude auprès d'élèves de CM1-CM2 de REP+ selon le protocole suivant :

pré-test, 5 séances d'apprentissage pour deux groupes de 4 classes appariés lors du prétest et différenciés selon le matériel pédagogique utilisé, post-test. Une diversité de variables mathématiques (durée, effectif, prix, poids) et de degré de généralisation nécessaire pour fonder sémantiquement la factorisation constituait les tests. Afin d'évaluer leur degré d'abstraction conceptuel, les élèves étaient invités à élaborer plusieurs stratégies de résolution. Les résultats permettent d'envisager des outils d'évaluation progressifs. Le développement conceptuel se mesure par la capacité des élèves à factoriser et proposer les deux stratégies. Les différences significatives en termes de démarches selon les variables et le degré de généralisation des problèmes suggèrent qu'une évaluation différenciée est possible. Un problème s'appuyant sur des préconceptions des élèves (tel que les problèmes de variable Effectif) facilite l'acquisition de la stratégie de factorisation mais n'évalue pas la maîtrise conceptuelle de l'élève. Au contraire, des problèmes faisant intervenir d'autres types de variables et différents degrés de généralisation permettent d'évaluer le degré de dépendance aux connaissances de la vie quotidienne, la faculté de généralisation et la capacité de transfert de stratégie. Cette démarche d'évaluation fondée sur la nature des problèmes et des stratégies effectuées dans des problèmes arithmétiques nous semble constituer une piste prometteuse pour étudier la conceptualisation des notions arithmétiques et leur évolution.

<sup>1</sup> <http://www.ldar.univ-paris-diderot.fr/page/praeval>

<sup>2</sup> LaboMep [www.labomep.net](http://www.labomep.net) est une plateforme d'apprentissage en ligne développée par l'association sésamath et couvrant l'école, le collège et le début du lycée ; WIMs (acronyme de Web Interactive Multipurpose Server) est une plateforme d'apprentissage en ligne, couvrant de nombreuses disciplines de l'école primaire jusqu'à l'université..

# LISTE DES PARTICIPANTS

NOM	Prénom	Institution	Adresse e-mail
ALPHA	Chantal	Académie Créteil	chantal.alpha@ac-créteil.fr
ANDRIEU	Clara	Académie de Paris	clara.andrieu@ac-paris.fr
ANTIBI	André	Université P. Sabatier	andre.antibi@gmail.com
ARTIGUE	Michèle	LDAR	michele.artigue@univ-paris-diderot.fr
ARZOUMANIAN	Valérie	Académie Orléans-Tours	valerie.arzoumanian@ac-orleans-tours.fr
ARZOUMANIAN	Philippe	DEPP	philippe.arzoumanian@me.com
ATAL	Ignacio	CRESS	ignacioatal@gmail.com
AUBIN	François	ESPE Bordeaux	francois.aubin@u-bordeaux.fr
BALLEREAU	Marie-Ange	Académie Orléans-Tours	marie-ange.ballereau@ac-orleans-tours.fr
BENAMAR	Aichai	CRASC	aichabenamar@hotmail.com
BIJOU	Mohamed	Laboratoire d'économie appliquée au Développement (TOULON)	medbijou.eco@gmail.com
BILGOT	Anne	ESPE Paris	anne.bilgot@espe-paris.fr
BLUM	Joëlle	Académie de Paris	Joelle.blum@ac-paris.fr
BODIN	Antoine	IREM de Marseille	antoinebodin@mac.com
BLANC	Christophe	EDA	cblanc@enfanticages.com
BOUSSARIE	Véronique	Académie Paris	veronique.boussarie@ac-paris.fr
BROUSSAN	Nicolas	Académie Créteil	nicolas.broussan@ac-créteil.fr
CABANE	Robert	IGEN	robert.cabane@education.gouv.fr
CASTEL-DOMPS	Sylvie	Académie Créteil	sylvie.castel@ac-créteil.fr
CASTEL	Frédéric	ESPE Reims	frederic.castel@univ-reims.fr
CELI	Valentina	Lab-E3D (ESPE Aquitaine)	valentina.celi@u-bordeaux.fr
CHAUVET	Catherine	Académie Bordeaux	catherine.chauvet@ac-bordeaux.fr
CHENEVOTOT	Françoise	LDAR (ESPE Lille)	chenevotot.francoise@neuf.fr
CHOQUET	Christine	CREN (ESPE Nantes)	christine.choquet@univ-nantes.fr
De CONINCK	Brigitte	SEGEC (Belgique)	brigitte.deconinck@segec.be
COLIN	Stéphanie	Académie Paris	stephanie.colin@ac-paris.fr
COPPÉ	Sylvie	Université de Genève	sylvie.coppe@unige.ch
CORNE	Robert	Académie Créteil	robert.corne@ac-créteil.fr
CORNET	Christine	Académie Créteil	christine.cornet@ac-créteil.fr
COULANGE	Lalina	LACES (ESPE Aquitaine)	lalina.coulangue@u-bordeaux.fr
DEMOULIN	Patricia	ESPE Aquitaine	patricia.demoulin@u-bordeaux.fr
DEPREZ	Michèle	ESPE Paris	michele.deprez@espe-paris.fr
DIKA	Sandrine	Académie Paris	sandrine.dika@ac-paris.fr
DIVOUX	Catherine	Académie Créteil	catherine.divoux@ac-creteil.fr
DRIJVERS	Paul	Institut Freudenthal	p.drijvers@uu.nl
DUFY	Cécile	ESPE Caen	cecile.dufy@unicaen.fr
DUQUESNE	Mathilde-Cécile	Académie Paris	mathilde-cecile.duquesne@ac-paris.fr
DUROISIN	Natacha	Université de Mons	natacha.duroisin@umons.ac.be
FARIAS	Luiz Marcio	Université Bahia-BRÉSIL	lmsfarias@ufba.br
FEYFANT	Annie	IFÉ	annie.feyfant@ens-lyon.fr
FERRARI	Anne-Catherine	Académie Versailles	aferrari1@ac-versailles.fr
FLOURET	Romain	Académie Créteil	romain.flouret@ac-creteil.fr
FOULQUIER	Laurianne	ESPE Aquitaine	laurianne.foulquier@u-bordeaux.fr
FRETIGNE	Patrick	IREM de Rouen	pf@univ-rouen.fr
GAONA	Jorge	LDAR	jorge.gaona.p@univ-paris-diderot.fr
GATEAU	Agnès	DSDEN 89	agnes.gateau@ac-dijon.fr
GENELOT	Sophie	IREDU (ESPE Bourgogne)	sophie.genelot@u-bourgogne.fr

# LISTE DES PARTICIPANTS

NOM	Prénom	Institution	Adresse e-mail
GERON	Christine	Haute Ecole de Liège (Belgique)	christine.geron@gmail.com
GILOT	Morgan	Académie Orléans-Tours	morgan.gilot@ac-orleans-tours.fr
GOUASDÉ	Rémi	EDA	remi.goasdoue@parisdescartes.fr
GONCALVES	Patricia	Académie Aix Marseille	patricia.goncalves@ac-aix-marseille.fr
GONZALEZ	Natacha	Académie Créteil	natacha.gonzalez@ac-creteil.fr
GRAPIN	Nadine	Académie Créteil	nadine.grapin@u-pec.fr
GRASSART	Dominique	Académie Créteil	dominique.grassart@ac-creteil.fr
GRAU	Sylvie	ESPE Nantes	sylvie.grau@univ-nantes.fr
GRUGEON-ALLYS	Brigitte	LDAR (ESPE Créteil)	brigitte.grugeon-allys@u-pec.fr
GUIGNARD	Marc	IREM de Lyon	marc.guignard@free.fr
GUILLAUD	Sandrine	Académie Versailles	sandrine.guillaud@ac-versailles.fr
GUYOT	Daniel	Académie Créteil	daniel.guyot@ac-creteil.fr
HAAS	Ben	Université du Luxembourg	ben.haas@men.lu
HASPEKIAN	Mariam	EDA	mariam.haspekian@parisdescartes.fr
HEBERT	Carole	Académie Versailles	chebert1@ac-versailles.fr
HERAULT	Françoise	ESPE Paris	francoise.herault@espe-paris.fr
HERRERO	Stéphane	Académie Créteil	stephane.herrero@ac-creteil.fr
HOROKS	Julie	LDAR (ESPE Créteil)	Julie.horoks@u-pec.fr
HOUSSARD	Marie-José	Académie Créteil	mariejose.houssard@gmail.com
HUET	Catherine	Académie Paris	catherine.huet@ac-paris.fr
HUGUET	Thomas	Académie Versailles	thomas.huguet@ac-versailles.fr
HUILLERY PERRIN	Aurélié	Académie Créteil	mathshuillery@gmail.com
JEZEQUEL	Françoise	Académie Créteil	francoise.jezequel@ac-creteil.fr
JOURAN	Maëlle	Académie Rouen	maelle.jouran@ac-rouen.fr
KEBOUCHI	Kadir	Académie Créteil	kkebouch@gmail.com
KIPFER	Nadine	IFFP (Suisse)	nadine.kipfer@iffp-suisse.ch
KIWAN ZACKA	Michella	Beyrouth - Liban	michella.kiwan@gmail.com
LAGANIER	Pierre	Académie Créteil	pierre.laganier@ac-creteil.fr
LAHANIER-REUTER	Dominique	CIREL (Lille 3)	dominique.reuter@numericable.fr
LALAUDE	Marc	Académie Bordeaux	marc.lalaude@gmail.com
LEFORESTIER	Céline	Académie Versailles	celine.leforestier@ac-versailles.fr
LECLERC	Nathalie	Académie Créteil	nathalie.leclerc@ac-creteil.fr
Le ROUX	Julien	Académie Paris	Julie.leroux@ac-paris.fr
LI	Linxia	EDA	li.linxiash@gmail.com
LOYE	Nathalie	Université de Montréal	nathalie.loye@umontreal.ca
MAGNE	Adam	Académie Paris	adam.magne@ac-paris.fr
MAILLOUX	François	Académie Créteil	francois.mailloux@ac-creteil.fr
MALOT	Sonia	Académie Créteil	sonia.malot@ac-creteil.fr
MASSELOT	Pascale	LDAR (ESPE Versailles)	pascale.masselot@u-cergy.fr
MASSON	Emmanuelle	Haute Ecole de Liège	emmanuelle.masson@hel.be
MATULIK	Nicole	Académie Paris	nicole.matulik@ac-paris.fr
MAYENSON	Jean-Baptiste	Académie Paris	Jean-baptiste.mayenson@ac-paris.fr
MENOUAR VERFAILLIE	Elisabeth	CIREL (Lille 3)	elimenouar@gmail.com
MESNIL	Zoé	LDAR (ESPE Créteil)	zoe.mesnil@u-pec.fr
MINASSIAN	Laure	Université de Caen	laure.minassian@unicaen.fr
MOULIN	Marianne	ICAR	marianne.moulin@cue-lillenorddefrance.fr
MONDIN	Christophe	ESPE Aquitaine	christophe.mondin@u-bordeaux.fr

# LISTE DES PARTICIPANTS

NOM	Prénom	Institution	Adresse e-mail
MOTTIER LOPEZ	Lucie	Université Genève	lucie.Mottier@unige.ch
MOUNIER	Eric	LDAR (ESPE Créteil)	eric.mounier@u-pec.fr
NOLÉO	Chrystelle	ESPE de Martinique	chrystel.Noleo-guitteaud@espe-martinique.fr
ORRE	Line	Académie Versailles	line.orre@ac-versailles.fr
PAILLET	Vincent	Académie Orléans Tours	vincent.paillet@ac-orleans-tours.fr
PANERO	Monica	IFÉ	monica.panero@ens-lyon.fr
PASQUALINI	Frédéric	Académie Créteil	fpasqualini1@ac-creteil.fr
PASTOR	Jean-Marc	DEPP	jean-marc.pastor@education.gouv.fr
PENNISI	Mario	Université de Catania	pennisi@dmi.unict.it
PERRIN	Gaëtan	Académie Clermont-Ferrand	gaetan.perrin@ac-clermont.fr
PETIOT	Nicolas	Académie Orléans-Tours	nicolas.petiot@ac-orleans-tours.fr
PETITJEAN	Stephan	Académie Paris	Stephan.petitjean@ac-paris.fr
PFUFF	Nathalie	ESPE Créteil	nathalie.pfuff@u-pec.fr
PILET	Julia	LDAR (ESPE Créteil)	julia.pilet@u-pec.fr
POUSSET	Arnaud	Académie Créteil	arnaud.pousset@ac-creteil.fr
PRADO	Julie	Académie Paris	julie.prado@ac-paris.fr
QUIQUEMPOIS	Grégory	Académie Créteil	quiquempois.g@gmail.com
RAULIN	Dominique	EDA	raulin-pro@orange.fr
REYDY	Carine	Académie Bordeaux	carine.Reydy@u-bordeaux.fr
RIARD	Sophie	Académie Créteil	sophie.riard@ac-creteil.fr
RIVA	Françoise	Académie Bordeaux	francoise.riva@neuf.fr
RIVIERE	Christophe	Académie Versailles	christophe.riviere@ac-versailles.fr
ROBIN	Agnès	Académie Paris	agnes.robin@ac-paris.fr
RODITI	Eric	EDA	eric.roditi@parisdescartes.fr
ROUZE	Pascale	Académie Paris	pascale.rouze@ac-paris.fr
RUF	Fabrice	Académie Paris	fabrice.ruf@ac-paris.fr
SALLES	Franck	DEPP	franck.salles@education.gouv.fr
SAYAC	Nathalie	LDAR (ESPE Créteil)	nathalie.sayac@u-pec.fr
SCARON	Evelyne	LDAR	evelyne.scaron@univ-paris-diderot.fr
SCHEIBLING-SÈVE	Calliste	Académie Versailles	calliste.scheibling-seve@hec.edu
SERRA	Christelle	Académie Créteil	chrystelle.serra@ac-creteil.fr
SERVAT	Emmanuelle	ESPE Paris	emmanuelle.servat@espe-paris.fr
SIMONNEAU	Pascale	Académie Créteil	pascale.simonneau@ac-creteil.fr
SIREJACOB	Stéphane	LDAR	stephanesirejacob@hotmail.fr
SOMMIER	Fabien	Académie Créteil	fabien.sommier@gmail.com
SOUMAN	Denis	ESPE de Lorraine	denis.souman@univ-lorraine.fr
TEMPIER	Frédéric	LDAR (ESPE Versailles)	frederick.tempier@u-ergy.fr
THOMAS	Catherine	ESPE Strasbourg	catherine.thomas@espe.unistra.fr
UNN-TOC	Isabelle	Académie Créteil	isabelle.unn-toc1@ac-creteil.fr
VANDEBROUCK	Fabrice	LDAR	vandebro@univ-paris-diderot.fr
VANTOUROUT	Marc	EDA	marc.vantourout@paris5.sorbonne.fr
VELDHUIS	Michiel	Institut Freudenthal	m.veldhuis@uu.nl
VIEGAS	Odile	Académie Créteil	odile.viegas@ac-creteil.fr
VILLANYI	Denise	LUCET (Luxembourg)	denise.villanyi@uni.lu
VIRALDE	Chrystelle	Académie Créteil	chrystelleviralde@hotmail.fr
VOISIN	Samuel	CERSE (ESPE Caen)	samuel.voisin@unicaen.fr
WEISS	Laura	Université de Genève	laura.weiss@unige.ch

<http://evaluationmaths.sciencesconf.org>

**Avec le soutien de :**

- Laboratoire de Didactique André Revuz (LDAR), *Universités Paris Diderot, Cergy Pontoise, Paris-Est Créteil, d'Artois, de Rouen*
- Laboratoire Éducation Discours Apprentissages (EDA), *Université Paris Descartes*
- Laboratoire d'informatique LIP6, *Université Pierre et Marie Curie*
  - Agence Nationale pour la Recherche (ANR)
  - Association des Directeurs d'IEM (ADIREM)
    - MAIF
    - CASDEN