

Site Paris Rive Gauche de l'Université de Paris Cité
Amphi Turing, Bâtiment Sophie Germain, 8 place Aurélie de Nemours, 75 013 Paris

Évaluations à grande échelle en mathématiques : regards croisés

L'évaluation dans l'enseignement, et notamment dans l'enseignement des mathématiques, recouvre beaucoup d'aspects. Dépassant le cadre d'une classe ou d'un groupe de travaux dirigés à l'université, ce colloquium s'intéresse aux évaluations à grande échelle. Dans le contexte actuel de l'enseignement, nous pensons par exemple aux évaluations nationales systématiques à chaque niveau scolaire mises en place depuis 2025, aux évaluations internationales (PISA, TIMSS), ou encore aux évaluations de grandes cohortes en licence de mathématiques. Ces évaluations à grande échelle amènent des questionnements sur l'enseignement et l'apprentissage des mathématiques : quel est l'impact de ces évaluations sur l'enseignement et l'apprentissage des élèves selon les usages qui en sont faits en classe (par exemple utilisation par l'enseignant pour gérer l'hétérogénéité des élèves), par les décideurs politiques (par exemple pour modifier les contenus des programmes ou les préconisations pour les modalités d'enseignement).

Les conférences viseront à éclairer la question des évaluations à grande échelle avec des points de vue variés. Nous cherchons à aborder les questions suivantes :

- Quelles sont les conditions et les contraintes qui sous-tendent les évaluations nationales et internationales, du point de vue institutionnel ?
- Qu'évaluent les évaluations à grande échelle, et notamment les évaluations nationales et internationales ? Quelles tâches mathématiques ? Quels sont les apports potentiels de la didactique des mathématiques pour outiller la conception d'évaluations à grande échelle, pour contribuer à l'élaboration et à l'analyse a priori des items, pour interroger les conditions de passation et pour interpréter les résultats ?
- Quel rôle peut jouer l'Intelligence Artificielle dans les évaluations à grande échelle pour élaborer, analyser, interpréter les résultats des évaluations ?

Programme

13h45 – 14h. Introduction du thème

14h – 14h50. **Méthodologie des évaluations standardisées en mathématiques : de la conception à l'interprétation des résultats dans PISA, TIMSS et les évaluations nationales.** Vincent Paillet & Franck Salles

14h55 – 15h45. **Interrogations didactiques à propos des évaluations à grande échelle en mathématiques.** Eric Roditi

Pause

16h05 – 16h55. **Concevoir une évaluation à grande échelle sur les fractions sans perdre de vue les savoirs.** Marie-Caroline Croset

17h – 17h50. **Évaluer les étudiants en mathématiques à l'ère de l'IA : entre efficacité pédagogique, équité et inclusion.** Lucie Jacquet-Malo & Joachim Lebovits

17h50 – 18h00. Clôture

Méthodologie des évaluations standardisées en mathématiques : de la conception à l'interprétation des résultats dans PISA, TIMSS et les évaluations nationales

Vincent Paillet, Direction de l'évaluation de la prospective et de la performance, ministère de l'Éducation nationale
Franck Salles, Direction de l'évaluation de la prospective et de la performance, ministère de l'Éducation nationale

Les évaluations standardisées à grande échelle en mathématiques, qu'elles soient internationales (PISA, TIMSS) ou nationales au secondaire, reposent sur des cadres de référence et des choix méthodologiques structurants. Cette présentation propose un éclairage sur les principales étapes de leur élaboration, depuis la définition des cadres et la construction des instruments jusqu'à l'analyse et l'interprétation des résultats. L'accent sera mis sur les questions de validité des contenus, de comparabilité des mesures et de limites d'interprétation, ainsi que sur la complémentarité de ces dispositifs pour appréhender le niveau des acquis des élèves. Il s'agira de montrer en quoi la lecture croisée des résultats issus de différentes évaluations contribue à une compréhension plus fine et plus nuancée des performances en mathématiques, dans toute la complexité du système éducatif.

Interrogations didactiques à propos des évaluations à grande échelle en mathématiques

Eric Roditi, EDA, Université Paris Cité

En nous appuyant sur des analyses didactiques d'items issus de différents questionnaires d'évaluation à grande échelle en mathématiques, nous nous interrogeons sur le concept de validité, c'est-à-dire la capacité à évaluer effectivement ce qu'on souhaite évaluer. Toujours en nous appuyant sur des exemples, nous prolongeons la réflexion sur la validité en fonction des modalités de passation : production libre ou question à choix multiple, temps libre ou temps limité, sur papier ou sur écran, etc. Nous nous interrogeons également à ce sujet sur les résultats produits et leurs interprétations. Dans un deuxième temps, nous présentons une étude visant à interroger les performances des élèves à une évaluation selon leur familiarité avec les questions auxquelles ils ont à répondre lors de cette évaluation. Enfin, nous développons un exemple où l'institution scolaire prend en compte les résultats des élèves à une évaluation internationale pour initier une modification des programmes scolaires.

Concevoir une évaluation à grande échelle sur les fractions sans perdre de vue les savoirs

Marie-Caroline Croset, LIG, Université Grenoble Alpes

Dans le cadre de l'évaluation à grande échelle d'un dispositif d'enseignement des fractions en cycle 3 dans l'académie de Grenoble, déployé dans près de 150 classes, nous avons été amenés à concevoir un test de connaissances portant sur le concept de fraction. Cette communication revient sur les choix méthodologiques effectués lors de la conception de ce test, en particulier sur la manière dont des outils issus de la didactique des mathématiques ont été mobilisés pour structurer l'objet de savoir, définir les tâches proposées et contrôler les connaissances effectivement évaluées. La conception d'une évaluation à grande échelle pose en effet la difficulté de concilier des contraintes de standardisation et de faisabilité avec le maintien d'une attention aux savoirs mathématiques effectivement en jeu, difficulté à laquelle la didactique peut peut-être apporter des réponses.

Évaluer les étudiants en mathématiques à l'ère de l'IA : entre efficacité pédagogique, équité et inclusion

Lucie Jacquet-Malo, coordinatrice du programme IA, Institut Mines-Télécom

Joachim Lebovits, Maître de conférences, Université Sorbonne Paris Nord, senior lecturer à NYU Paris, CEO PyxiScience

Les résultats des enquêtes internationales (PISA, TIMSS) révèlent un changement dans les compétences mathématiques chez les élèves, interrogeant la formation des enseignants, muée par de multiples réformes, et les méthodes d'évaluation. Parallèlement, l'afflux d'étudiants en L1/L2, couplé à la charge administrative et scientifique des enseignants-chercheurs, rend l'évaluation individuelle de qualité difficile. L'IA pourrait-elle offrir des solutions, sans sacrifier l'équité ou la dimension humaine de l'enseignement ? Comment concilier l'impératif d'évaluer efficacement des cohortes toujours plus nombreuses, tout en garantissant une pédagogie inclusive et en préservant le rôle central de l'enseignant ? Quels risques (biais, exclusion, déshumanisation) l'automatisation de l'évaluation fait-elle peser, et comment les atténuer ? Cette conférence explorera les pistes offertes par l'IA pour soutenir l'évaluation en mathématiques, sans déléguer aveuglément.