

Comité scientifique des IREM

Extrait (« Evaluation ») du procès-verbal de la réunion du 11 avril 2014

Adopté lors de la réunion du 13 juin 2014

Membres du CS présents : Michèle ARTIGUE, Eric BARBAZO, Laurent BOUDIN, René CORI, Philippe DUTARTE, Bernard EGGER, Brigitte GRUGEON-ALLYS, Catherine HOUEMENT, Jean-Charles JACQUEMIN, Gérard KUNTZ, Yann LEFEUVRE, François MOUSSAVOU, Isabelle MOUTOUSSAMY, Jean-Claude ORIOL, Dominique POIRET-LOILIER, Jean-Pierre RAOULT, Fabrice VANDEBROUCK

Invités : André ANTIBI, Sylvie COPPE, Corinne CROC, Paul DRIJVERS

Ce procès-verbal, dont le projet a été rédigé par le secrétaire de séance, Jean-Pierre Raoult, avec l'aide des notes de Catherine Houdement, complète le document « relevé de conclusions », qui a été diffusé, ainsi que ses annexes, après la réunion par la présidente du comité scientifique, Michèle Artigue¹. Celui-ci est référencé dans le corps du procès-verbal par *relevé*.

Des notes en bas de page donnent des références ou des indications postérieures à la réunion.

.....

3. Réflexion et débat sur l'évaluation, perspectives pour le travail des IREM sur ce thème.

Nous retraçons ici successivement des échanges suivant chacun des exposés, puis la discussion globale qui les a suivis.

3.1. Echanges suivant les exposés

a. Après l'exposé de Sylvie Coppé (*Développer les pratiques d'évaluation formative dans les classes, annexe 2 au relevé de conclusions*)

Sylvie Coppé est interrogée sur la distinction entre évaluation formative planifiée et évaluation formative non planifiée (« on the fly ») ; Catherine Houdement craint le risque que cette dernière catégorie en vienne à inclure de manière indistincte de interactions très diverses entre enseignants et élèves et non nécessairement significatives. Sylvie Coppé répond que le visionnement des vidéos d'enregistrement des séances en classe favorise le tri des éléments les plus pertinents.

Jean-Claude Oriol revient sur la phrase, dans le diaporama de Sylvie Coppé : *fournir à l'élève de l'information sur l'écart entre ce qui est attendu et ce que l'élève sait*. Il s'interroge sur le sens de l'expression *ce que l'élève sait* et pointe la difficulté de cette notion, *dans laquelle signifiant et signifié se chevauchent*. Sylvie Coppé précise qu'il s'agit de cerner « là où en est l'élève » sur le point de mathématiques en jeu.

Jean-Claude Oriol s'interroge aussi sur le rôle des enseignants « experts » intervenant pour apprécier une telle action internationale sur un sujet assez pointu. Il souligne le risque de « connivence » sur le sujet entre les chercheurs et les experts, vu le recrutement de ces derniers parmi les enseignants travaillant déjà avec les chercheurs.

Michèle Artigue s'intéresse à la phrase, dans le diaporama de Sylvie Coppé : *s'assurer des effets positifs des actions* ; elle souligne la difficulté d'un tel constat. Sylvie Coppé indique l'importance, à cet égard, du fait qu'il s'agit d'un travail sur le long terme, où l'on dispose de l'enregistrement de 18 séances successives.

b. Après l'exposé de Paul Drijvers (*Evaluation en mathématiques : innovation et technologie, le bon mariage ? annexe 3 au relevé de conclusions*)

Les rôles respectifs des psychométriciens et des didacticiens suscitent la curiosité de membres du comité scientifique. Paul Drijvers précise que l'analyse psychométrique intervient à titre de contrôle de la pertinence des outils numérisés d'évaluation élaborés.

Paul Drijvers est invité à préciser quelques points du contexte néerlandais :

¹ Il a été placé le 14 mai 2014 dans la rubrique du CS sur le portail des IREM : <http://www.univ-irem.fr/spip.php?article1077>

- nature de l'organisme CITO auprès duquel s'effectuent les études présentées : il s'agit d'un institut indépendant qui reçoit du ministère de l'éducation la tâche de faire des évaluations (ce qui confère aussi au ministère une fonction de validation de ces études),

- organisation du diplôme de fin d'études secondaires : 50% attribué localement par les enseignants, 50% sous forme centralisée.

Interrogé sur l'impact de ces travaux sur les pratiques locales, Paul Drijvers répond qu'il est encore trop tôt pour pouvoir l'apprécier.

c. Après l'exposé de Brigitte Grugeon-Allys (*Évaluation diagnostique et régulation, annexe 4 au relevé de conclusions*)

Interrogée par Jean-Pierre Raoult sur le sens du mot « régulation », Brigitte Grugeon-Allys répond qu'il renvoie à la différenciation des actions à mener auprès des élèves à la suite d'une évaluation diagnostique.

Se référant aux travaux de Gérard Vergnaud, Jean-Claude Oriol se demande si le travail présenté n'analyse pas plus le « signifiant » que le « signifié » pour l'élève. Il lui est répondu que, au travers des enregistrements effectués, l'analyse des interprétations des élèves est prise en compte.

François Moussavou s'interroge sur le lien entre les documents à disposition des professeurs et l'élaboration des évaluations. Brigitte Grugeon-Allys et Isabelle Moutousamy indiquent que les documents ressources, et même les commentaires dans les programmes, ne sont pas autant utilisés par les enseignants qu'il serait souhaitable.

d. Après la présentation par Michèle Artigue de l'exposé de Jean-François Chesné (*L'utilisation d'évaluations standardisées dans une formation d'enseignants de sixième : une expérimentation dans le domaine des nombres et du calcul, annexe 5 au relevé de conclusions*)

Cette présentation n'a pas été suivie de discussion.

e. Après l'exposé d'André Antibé et Corinne Croc (*L'évaluation par contrat de confiance, annexe 6 au relevé de conclusions*)

Sollicités par Bernard Egger de donner des précisions sur les effets de l'évaluation par contrat de confiance (sont-ils durables ? Ont-ils été quantifiés ?) André Antibé et Corinne Croc invoquent le cas d'élèves de l'académie de Créteil ayant bénéficié en école primaire de cette évaluation, dont les notes aux tests lors de l'entrée en sixième ont été particulièrement bons. Ils insistent sur le changement d'attitude des élèves face à l'apprentissage et sur le bénéfice lié au fait que ce fonctionnement de l'évaluation permet d'approfondir leur travail sur les items du programme.

Isabelle Moutousamy constate que ce changement présenté comme substantiel ne remet pas en cause la notion de note ; André Antibé répond que le problème n'est pas tant celui de la note que celui de la « culture d'évaluation » dans l'enseignement français ; pour obtenir l'adhésion des enseignants, il faut leur présenter une réforme qui leur apparaisse faisable,; la note leur apparaît comme un repère essentiel, surtout quand ils sont amenés à voir un très grand nombre d'élèves.

François Moussavou s'interroge sur l'adéquation de l'évaluation par contrat de confiance à l'apprentissage de « tâches complexes » qui est important dans le cadre de l'évaluation par compétences pratiquée en Lycée Professionnel. André Antibé rappelle que l'évaluation en temps limité n'occupe qu'une faible part du temps scolaire (il l'évalue à un douzième) et que les tâches complexes relèvent plutôt de situations d'apprentissage sans contrainte de temps trop forte.

Fabrice Vandebrouck s'interroge sur la manière dont des élèves ayant pratiqué l'évaluation par contrat de confiance se situent face aux conditions d'examen, en particulier au baccalauréat. Corinne Croc répond que les sujets proposés dans ces évaluations sont « classiques » et donc en particulier en terminale sont du type de ceux du baccalauréat et que ces candidats sont d'autant plus à l'aise qu'ils auront été ainsi incités à mieux travailler ces sujets. André Antibé fait remarquer que, s'agissant de l'entrée aux grandes écoles d'ingénieurs, la pratique de l'oral du « concours commun polytechnique », où un exercice d'oral sur deux est nécessairement extrait d'une banque connue, s'apparente au contrat de confiance.

3.2. Débat général

Nous organisons la relation de ce débat suivant les thèmes abordés.

a. Rôle de la psychométrie

Jean-Pierre Raoult exprime qu'il est frappé par la montée de la psychométrie dans l'évaluation des élèves (référence obligée dans les indications sur les exercices de PISA, exposés que l'on vient d'entendre...). Il s'interroge sur l'importance de ce phénomène et sur ses retombées sur la recherche didactique.

Paul Drijvers répond que les mathématiciens peuvent avoir de « belles idées » sur l'enseignement mais que celles-ci peuvent être vaines si leur impact n'est pas mesurable ; ceci dit, l'élaboration de ce qui est mesurable en matière d'effet de l'enseignement n'est pas évidente et nécessite un dialogue entre les parties prenantes.

Brigitte Grugeon-Allys indique que, dans son travail, les psychométriciens présents à la DEPP² interviennent dès la conception du dispositif ; ils aident à évaluer la pertinence des exercices proposés, puis dans l'analyse des résultats d'un premier test sur échantillon, conduisant à la validation de la plupart des exercices, mais aussi au retrait de certains jugés trop difficiles ou se prêtant mal à une discrimination. Un problème est d'éviter que l'analyse des résultats des élèves ne conduise à les regrouper plus sur un critère « de réussite » chiffrée que sur les modes de raisonnement mis en pratique ; des informations complémentaires sur les travaux des élèves doivent donc aussi être recherchées.

b. Hétérogénéité des classes

Bernard Egger rappelle que, lors de la table ronde tenue dans la session finale de la semaine des mathématiques à la bibliothèque François Mitterrand le 22 mars 2014, tous les intervenants avaient fait état des difficultés liées à l'hétérogénéité des classes ; il demande si les études sur l'évaluation fournissent des éléments pour y remédier.

Sylvie Coppé montre qu'il s'agit d'un problème « politique », lié à des traditions néfastes de l'école française qui, à l'image de notre société très élitiste, tend à « fabriquer » de l'hétérogénéité. Notre conception de l'évaluation sommative doit donc être fortement révisée.

Brigitte Grugeon-Allys ajoute que cette hétérogénéité est accrue par le phénomène d'angoisse chez les élèves qui cherchent d'abord, en situation d'évaluation, à se situer par rapport à l'attente supposée du professeur. Faire évoluer ce climat est difficile car les enseignants tendent à reproduire ce qu'ils ont eux-mêmes vécu.

Bernard Egger revient sur sa question en demandant si cette hétérogénéité n'est pas aussi favorisée par des programmes au contenu trop ambitieux pour une part de la population scolaire ; Brigitte Grugeon-Allys pense plutôt que la question de la « culture » française dans l'exigence des résultats de l'élève est primordiale ; elle est appuyée par Michèle Artigue qui indique que la situation à cet égard tend maintenant à être différente en Lycée Professionnel, comme l'a montré la réunion précédente du comité scientifique.

c. Généralisation des résultats des recherches présentées

Jean-Claude Oriol pointe différentes difficultés relatives au profit à tirer des recherches qui viennent d'être présentées : le temps de la recherche n'est pas celui de la pratique courante en classe, les situations de tests tendent à créer des distinctions artificielles, la complexité de la tension entre évaluation et contrôle n'est pas la même en classe que lors de ces expérimentations où existe une forme de complicité entre maître et élève.

Brigitte Grugeon-Allys répond que ces travaux se déroulant sur une assez longue période (2 à 3 ans), il est possible d'observer une stabilisation des pratiques et de véritables processus de création de la part des enseignants.

Sylvie Coppé indique que ces travaux peuvent déboucher sur un changement global de la place de l'évaluation ; elle rappelle que, dans l'introduction des programmes de collège, il est précisé que l'évaluation ne doit pas être « à côté » de l'enseignement mais lui être intégrée. Elle se réjouit de constater que les esprits des enseignants évoluent dans ce sens et elle donne l'exemple parfois pratiqué de l'annulation d'une mauvaise note si l'élève s'est amélioré lors d'une évaluation ultérieure. Elle exprime sa conviction que la réflexion sur l'évaluation peut avoir un impact salutaire sur tout l'enseignement.

d. Instabilité des « contrats » passés avec les élèves

Gérard Kuntz prend l'exemple de la Finlande, vantée pour le succès d'un système d'enseignement dans lequel l'évaluation bénéficie d'un climat apaisé fort différent de celui connu en France, mais où les étudiants se trouvent souvent en grande difficulté quand ils sont jugés dans l'enseignement supérieur (il cite des échecs importants en écoles d'ingénieurs). Il relève aussi des cas de variations dans l'usage des calculatrices lors de contrôles ou examens, plutôt favorisé maintenant dans les lycées français mais parfois interdit ensuite (il cite l'exemple de l'INSA de Lyon), ce qui déstabilise les étudiants.

Yann Lefeuvre indique que les pratiques évoluent vers un usage raisonné des calculatrices, en équilibrant les épreuves où elles sont utilisées et celles où elles ne le sont pas (exemple du concours d'entrée aux écoles d'agronomie).

²Direction de l'Évaluation, de la Prospective et de la Performance au ministère de l'Éducation Nationale, qui est l'un des partenaires du projet de l'Agence Nationale de la Recherche NéoPræval dont relève cette étude.

Mais René Cori considère que l'exemple de difficultés d'adaptation des étudiants cité par Gérard Kuntz manifeste un échec de l'enseignement secondaire.

Pour André Antibi, ces problèmes sont aussi dus au fait que l'évaluation diagnostique n'est pas assez en vigueur en France ; si elle l'était, les élèves seraient mieux aidés face à leurs difficultés d'adaptation quand leurs conditions d'enseignement changent.

Pour Sylvie Coppé, il est essentiel d'armer les élèves en explicitant bien auprès d'eux, à la fin des séquences d'enseignement, ce qui était attendu ; l'évaluation formative prend là tout son sens. En effet, le drame des élèves en difficulté est justement qu'ils ne comprennent pas ce qui était attendu d'eux.

e. Travail des professeurs pour mettre en place l'évaluation et la remédiation

Bernard Egger relève l'ampleur du travail nécessaire pour que les professeurs fassent évoluer leurs pratiques d'évaluation, et en particulier élaborent les processus de remédiation qui les accompagnent. Pour les y aider, est nécessaire tout un travail de confection de documentation qui intéresse l'APMEP mais qu'elle ne peut mener seule.

Corinne Croc affirme que le travail de remédiation guidé par l'évaluation est facilité si l'évaluation elle-même n'a pas été un facteur d'échec ; elle évalue à 10% le pourcentage d'élèves apparaissant en difficulté après les évaluations par contrat de confiance.

Brigitte Grugeon-Allys confirme que ceci suppose de favoriser le travail en commun des enseignants. Elle indique que, dans les groupes de formateurs de professeurs des écoles, on mène ce type d'élaboration. Il faut avoir le souci d'en faire essaimer les résultats.

Paul Drijvers, après avoir constaté la multiplicité des points abordés dans ce débat, avec un éclairage très axé sur la situation spécifique de la France (ou jugée telle par les intervenants), met en évidence le lien étroit entre les différents aspects de l'enseignement, dont l'évaluation et la remédiation ne sont que des composantes ; mais, dit-il, « s'il faut changer tout, il faut bien commencer quelque part » ; et l'évaluation lui semble une bonne porte d'entrée, à condition de ne pas trop en rester à des généralités et de bâtir des propositions concrètes et accessibles.

3.3. Conclusion

Voir *relevé* pour la lecture complète de la conclusion à ce débat, qui porte sur l'importance de travailler, en particulier dans le réseau des IREM, sur tous les aspects de l'évaluation, sur la nécessité de nouer à cet effet des contacts diversifiés (APMEP, Institut Français de l'Education, Inspections générale et régionales, ESPE ..) et sur le caractère incontournable de la formation continue des enseignants pour transmettre la culture et les outils permettant de mettre en pratique les évolutions reconnues indispensables, afin d'améliorer, sur ce point, les conditions de l'enseignement des mathématiques en France.

