

DIDACTIQUE DES MATHÉMATIQUES ET EVALUATION : PETITE HISTOIRE DE 40 ANS DE CHEMINS PARALLELES.

À TERME, L'EVALUATION EST-ELLE SOLUBLE DANS LA DIDACTIQUE ?

Antoine **Bodin**

Résumé

Le résumé n'a pas été communiqué par l'auteur.

Mots clés

INTRODUCTION

En complément d'une communication faite au séminaire national de didactique de mars 2015, je suis amené dans le présent texte à reprendre et à compléter mes notes et à essayer de développer ma pensée.

Rétrospectivement, je suis un peu effrayé par le titre que j'ai moi-même donné à mon intervention : ce titre exprime davantage ce qui me semble utile de faire que ce que je suis réellement capable de faire. J'ai de fait été témoin et modestement acteur de ces 40 années de recherches en didactique des mathématiques et de ces 40 années de travaux et d'opérations relatives à l'évaluation, en particulier à l'évaluation des acquis des élèves et il me semble aujourd'hui nécessaire de faire le point des rapports entre évaluation et didactique.

Au début des années 70, il était encore assez facile de faire le tour des recherches en didactique et des travaux sur l'évaluation, mais, en France comme dans nombre d'autres pays, ces deux domaines ont connus depuis un développement quasi exponentiel. À dire vrai, une partie de ces recherches et de ces travaux me sont restés plus ou moins étrangers et il est pour le moins présomptueux de prétendre faire le point.

L'EVALUATION... OU EN EST-ON ?

Avant le début des années 70, le terme évaluation était peu utilisé dans le domaine de l'éducation. Le souci de « mesurer » les acquis des élèves pour « mesurer » l'efficacité des systèmes éducatifs s'était cependant développé aux USA depuis le début du siècle. On parlait là-bas « d'*accountability* », terme que l'on traduit aujourd'hui en français par « redevabilité » : il s'agit de l'idée que les institutions ont des comptes à rendre à leurs mandants. On y utilisait aussi le vocable anglais d'*evaluation* et de nombreuses théories et dispositifs avaient été développés dans ce cadre.

En France, on parlait alors de contrôle, de notation, d'examens lorsqu'il s'agissait de rendre

compte des acquis individuels des élèves. Le souci de « mesurer » et de suivre globalement ces acquisitions (enquêtes à grande échelle) ne s'est imposé qu'à partir des années 70. Ensuite, peu à peu, l'évaluation a envahi l'ensemble des activités humaines et pas seulement le domaine de la formation. Il est dorénavant nécessaire de considérer l'évaluation comme un phénomène social et politique qui conditionne la plupart des aspects de notre vie.

Si l'on se limite à l'enseignement, un basculement s'est peu à peu effectué d'une évaluation largement placée entre les mains des enseignants et donc pilotée par l'amont à une évaluation pilotée par les institutions. Les études internationales comme les études nationales, de plus en plus invasives (ce qui ne préjuge ni de leur intérêt ni de leurs qualités éventuelles) tendent à réduire l'importance de l'enseignant dans le processus évaluatif des acquis des élèves.

Nous avons développé dans un autre article l'idée qu'il y a continuité entre les différents aspects de l'évaluation. De façon plus ou moins consciente, l'enseignant, dans sa pratique quotidienne, directement ou indirectement, est conditionné par les évaluations de masse et par les jugements qu'elles produisent, implicitement ou non, sur son action.

Dans les analyses que l'on peut faire comme dans les études que l'on peut lire sur l'évaluation, il est souvent difficile de séparer les aspects idéologiques et les aspects techniques de la question. Nombre d'essais publiés ces dernières années, tendent à vilipender l'usage qui est fait de l'évaluation ; ses aspects réducteurs et les conditionnements auxquels ils conduisent. La liste serait longue ; voici juste quelques titres récents pour qui mettent dans l'ambiance :

- La tyrannie de l'évaluation.
- L'évaluation, une menace ?
- La folie évaluation.
- Faut-il avoir peur de l'évaluation ?
- La fabrique des imposteurs.

Pour nous limiter à l'éducation, et à côté de ces ouvrages grand public, les articles des sociologues et des psychologues ainsi que ceux produits par les sciences de l'éducation sont nombreux qui dénoncent des effets négatifs de l'évaluation sur la qualité de l'enseignement et sur la motivation des élèves.

Le temps n'est donc plus où l'on pouvait considérer, avec une certaine naïveté, l'évaluation comme un moment de l'activité pédagogique plus ou moins éloigné de l'activité didactique proprement dite et ignorer les interdépendances existant entre les différents actes et dispositifs qui se prévalent de l'évaluation.

Ces réserves faites, nous nous limiterons dans ce texte aux rapports que, de fait, l'évaluation entretient avec l'action didactique et à la place qu'elle occupe ou qu'elle pourrait occuper dans les recherches en didactique des mathématiques.

ÉVALUATION ET DIDACTIQUE DES MATHÉMATIQUES

Le paragraphe précédent insiste sur le fait que la question de l'évaluation a maintenant envahi le champ de la formation à un point tel qu'il n'est plus possible d'ignorer le rôle de l'évaluation et les effets retour de l'évaluation sur l'acte d'enseigner comme sur le processus d'apprentissage. Dit autrement, le contrat didactique se constitue sous le contrôle et l'emprise de l'évaluation.

Venons-en donc à la didactique ou plus précisément à la recherche sur la didactique des mathématiques.

La recherche en didactique des mathématiques est elle aussi née dans les années 70, c'est à dire en même que l'évaluation commençait à s'imposer dans le système éducatif français.

Les précurseurs de la didactique, en particulier Guy Brousseau ont de suite vu et anticipé les risques que l'évaluation faisait planer sur l'action didactique. Toutefois, Guy Brousseau ne parle essentiellement d'évaluation qu'en référence à une action de type quasi administrative qu'il

convient de tenir à distance sans y prêter grande attention. En ce qui concerne le suivi des processus d'acquisition des élèves, tout était semble-t-il contenu dans la théorie des situations didactiques en particulier avec la dialectique de la validation.

D'une façon générale, la recherche en didactique s'est largement démarquée de la question de l'évaluation qu'elle a largement abandonnée aux psychologues et aux sciences de l'éducation. Il est symptomatique que, parmi les quelque 400 thèses de didactique des mathématiques répertoriées depuis 1975, 3 seulement contiennent le terme évaluation dans le titre (thèses de Régnier (1983), de Gagatsis (1982) et de Pilet (2012)).

D'autres thèses cependant font, explicitement ou non référence, à l'évaluation. Il en est ainsi de la thèse de Claire Margolinas (1989) qui consacre tout un chapitre sur le sujet et qui en particulier explicite la différence entre validation et évaluation.

Bien que l'étude de l'évaluation ne fasse pas partie du projet de thèse de Claire Margolinas, elle y consacre un chapitre entier dans lequel elle fait une étude détaillée des places respectives de la validation et de l'évaluation dans la classe. Cela à partir d'une définition minimale de l'évaluation qu'elle situe dans la phase de conclusion d'une situation didactique :

« Une phase de conclusion est une phase d'évaluation quand, dans cette phase, la validité du travail de l'élève est évaluée par le maître sous la forme d'un jugement sans appel ».

Elle précise encore : « *pour nous, phase de validation et phase de conclusion ne sont pas synonymes* ».

De façon indirecte, beaucoup d'autres travaux sont de nature à intéresser l'évaluation, par les méthodes qu'ils utilisent, par les instruments qu'ils développent, par les questionnaires qu'ils produisent et les analyses qu'ils font des observations faites.

On le sait, l'évaluation est polymorphe. Nous faisons toutefois l'hypothèse que sous ses différentes formes elle influe sur l'activité didactique.

- On sait depuis longtemps que les examens et autres évaluations externes ont une influence directe sur l'action des enseignants et sur la qualité des apprentissages des élèves.
- On sait de même que les pratiques d'évaluation des enseignants au quotidien peuvent motiver les élèves et stimuler les apprentissages, mais qu'elles peuvent tout aussi bien démotiver les élèves et contrarier les apprentissages.
- Dans bien des cas, l'évaluation a pour premier effet de substituer une motivation exogène à la motivation endogène qui préserve davantage le sens des apprentissages.
- Les études de masse, qu'elles soient nationales ou internationales, influent fortement sur les programmes d'enseignement et sur la formation des enseignants. Le fait que, à tort ou à raison, ces études mettent la priorité sur tel ou tel aspect de la formation, génère de façon quasi directe de nouvelles formes d'enseignement et de nouveaux thèmes d'étude. Il en est ainsi de l'insistance de ces études sur les liens à favoriser entre les mathématiques et la vie dite réelle ou sur la place des mathématiques dans la formation du citoyen. Le fait que les études PISA par exemple ne font aucune place à la démonstration est, me semble-t-il, à mettre en relation avec l'effacement progressif de la place de la démonstration dans l'enseignement du collège.

Parmi les formes d'évaluation, une place particulière doit être faite à l'évaluation dite formative qui s'est imposée à partir de 1990 dans les discours sur l'évaluation. En effet, par définition, l'évaluation formative se veut intégrée au quotidien de l'action didactique.

Dans la théorie des situations il n'est pas question d'évaluation dans le cadre du processus didactique proprement dit. La dialectique de la validation ferme le cycle des dialectiques sans mettre en doute que ce qui pouvait être compris avait été compris et ce qui pouvait être appris avait été appris. Cela sans s'opposer à des évaluations de type sommatif (contrôle) effectuées de façon extérieures aux situations didactiques proprement dit.

Provenant de domaines plus ou moins éloignés de la didactique ou même s'opposant à elle, la notion d'évaluation formative est apparue comme superfétatoire pour l'enseignement des mathématiques. Les objectifs et les vertus qu'on lui attribuait étant supposées être assurés par la dialectique de la validation.

Quelques jalons

Revenons un peu en arrière et cherchons à préciser de façon chronologique les premières intrusions des problématiques évaluatives dans les recherches en didactique des mathématiques.

La thèse de Régis Gras (1979) ne porte pas directement sur l'évaluation mais s'appuie sur un travail d'évaluation d'une expérimentation didactique (projet OPC). Bien que Régis Gras précise que ses « *intentions se limitent à expliciter une grille d'appropriation conceptuelle* », sa thèse est le point de départ des travaux sur l'analyse statistique implicite qui offre une alternative encore insuffisamment exploitée à l'approche « mesure » des acquisitions des élèves. La démarche taxonomique qui y est développée permet en effet de hiérarchiser les niveaux d'appropriation du savoir mathématique et permet donc, mieux que l'approche « mesure », de déterminer la valeur des observations faites, et ainsi permet d'évaluer dans le sens premier de ce terme.

La thèse de François Pluvinage (1977) porte directement sur l'évaluation avec sa référence à la classification NLSMA utilisée dans les évaluations internationales de l'IEA, mais surtout, de façon indirecte sur l'éclairage qu'elle porte sur la difficulté des questions posées aux élèves.

En 1977, Guy Brousseau fait un exposé remarqué lors du 29^{ème} congrès de la CIEAM dont le thème était « *Évaluation et enseignement des mathématiques* ». Sa communication : « *l'observation des activités didactiques* » a ensuite été reprise dans sa thèse. Le mot évaluation n'y est prononcé que très marginalement, ce qui est un signe que l'on peut décoder en nous référant en particulier un autre exposé fait par Guy Brousseau à ICME 5 en 1978 : « *Évaluation et théories de l'apprentissage en situations scolaires* ». On y lit en particulier :

« Beaucoup d'enfants et ensuite de professeurs n'acquièrent qu'une fausse pratique du savoir...

L'évaluation seule ne permet pas de corriger les phénomènes rapportés plus haut et parfois elle les accentue.... »

« 1 - L'évaluation permet à certains maîtres de pallier en partie aux insuffisances les plus criantes de son enseignement

2 - Elle le conduit ensuite à une diversification et à une démultiplication des objectifs intermédiaires, donc à recourir à des situations fermées visant des apprentissages à court terme, avec pour corollaire l'effet de rationalisation.

3 - Enfin, tôt ou tard l'évaluation se limite à certains objectifs convenus. Les maîtres les visent par des apprentissages d'algorithmes et de savoirs non fonctionnants. L'économie conduit à céder sur les exigences relatives à la compréhension...

Il n'est pas encore possible de distinguer, à l'aide des évaluations classiques, des connaissances acquises par une suite organisée d'assimilations ou même de conditionnement de celles qui sont acquises par une genèse authentique des concepts. »

Ailleurs, à la même époque, Guy Brousseau a bien anticipé, ce qui, au fil des ans, est devenu une réalité de plus en plus préoccupante, à savoir : « *l'écrasement des objectifs sur l'évaluation* ».

Dans les mêmes années, Georges Glaeser introduit l'évaluation dans son cours de didactique des mathématiques. Glaeser (1995) portait en effet un intérêt particulier à l'évaluation et a

suscité plusieurs thèses et autres travaux sur l'évaluation dans le cadre de la didactique des mathématiques.

Son cours porte en partie sur la docimologie et sur l'histoire du côté « mesure » de l'évaluation, mais il s'intéresse aussi à la validité du questionnement et à l'analyse des observations. Cela le conduit à repérer de profils de réponses à un questionnaire et à s'éloigner de la notation traditionnelle par addition de points qui n'ont à l'évidence pas la même valeur.

La relative absence des travaux et des références à l'évaluation dans les travaux de didactique est sans doute due, en grande partie au fait que la didactique des Mathématiques s'est d'abord développée à partir d'observations relatives à l'école élémentaire. Il est facile de montrer que, au moins à l'époque, les faits d'évaluation ne s'y présentaient pas de façon aussi prégnante que dans l'enseignement secondaire. Il n'est pas certain, s'il n'avait pas déplacé le focus sur les classes de collège et de lycée, que Yves Chevallard ait écrit en 1986 les phrases suivantes :

« [...] Pourtant lorsque [...] le didacticien tente de pénétrer dans l'histoire d'une classe, il doit se rendre à l'évidence : les faits d'évaluation qu'il peut alors y observer ne sont pas simplement un existant contingent, un mal nécessaire que l'on pourrait ignorer, mais bien l'un des aspects déterminants du processus didactique - qui règle et régule tout à la fois les comportements de l'enseignant comme l'apprentissage des élèves. Bref, quiconque pénètre un peu longuement dans la vie d'une classe ne peut longtemps ignorer la "tyrannie" du processus d'évaluation. » Chevallard et Feldmann (1986).

Ainsi, Chevallard avait bien vu à quel point l'évaluation dans la classe influait sur le processus didactique ; cela sans pour autant ignorer les effets en retour des autres types d'évaluation sur ces processus.

Par la suite Chevallard a en quelque sorte intégré ses réflexions concernant l'évaluation dans la Théorie Anthropologique du Didactique (TAD) en considérant l'évaluation faite par l'enseignant comme le sixième et dernier moment de l'organisation didactique (Chevallard 1999). Il est souvent revenu ensuite sur la question de l'évaluation (i.e. Chevallard 2004). L'approche théorique qu'il propose oblige à prendre du recul par rapport aux pratiques évaluatives, à mieux en saisir les enjeux et à débusquer les fausses évidences. Cette approche pourrait certainement permettre de mieux problématiser la question de l'évaluation dans le cadre de la didactique.

L'évaluation, que ce soit celle menée par l'enseignant ou celle menée par les institutions plus ou moins spécialisées, se nourrit en effet de fausses évidences. Par exemple, elle croit ou fait comme si elle croyait que les connaissances mathématiques pouvaient être mesurées autrement que dans un sens purement métaphorique. Qui plus est, elle croit ou fait semblant de croire que cette mesure peut être rapportée à une échelle unique.

L'évaluation dite critériée, comme l'évaluation dite par compétence, n'échappe pas à la critique. Ces démarches conduisent en effet souvent à mettre le savoir et l'évaluation en grilles aux mailles de plus en plus fines ; de plus, elle finit le plus souvent par se traduire par être rapportée à une échelle unidimensionnelle.

En utilisant partiellement les catégorisations de la TAD, il y a là une cascade de fausses évidences qui nourrissent la plupart des évaluations :

- Il serait immédiat de faire correspondre une tâche précise à toute case de la grille.
- Cette tâche suffirait pour représenter le type de tâche auquel cette tâche peut être associée.
- De la réussite (vs l'échec) à cette tâche particulière on pourrait déduire la réussite (vs l'échec) à toute tâche du type de tâche associé.
- De la réussite (vs l'échec) à cette tâche particulière on pourrait déduire quelque chose concernant la maîtrise de la technique ou de la théorie à laquelle elle est liée.
- Pour l'évaluation, cette tâche aurait la même valeur (le même poids) que toute autre

tâche associée à un autre type de tâche.

Exemple :

- Une case de la grille propose un type de tâche :
 - [Savoir /être capable de] Encadrer une fraction par deux entiers consécutifs.
« *Compétence* » du projet de programme cycle 3- 2015) :
- Une tâche pour l'évaluation qui appartient au type de tâche donné :
 - encadrer la fraction $\frac{17}{18}$ par deux nombres entiers.

De la réussite (vs l'échec) à cette tâche particulière que peut-on déduire concernant la maîtrise du type de tâche associé (ici la compétence en question) ?

Supposons que l'élève ait pu utiliser sa calculatrice, et ait observé que $\frac{17}{18} = 0,944\dots$ et en ait déduit que $\frac{17}{18}$ était compris entre 0 et 1, il aura sans doute fait preuve d'une certaine compétence, mais est-ce bien celle que l'on prétendait observer ?

S'il a, ce qui est sans doute ce que l'on attend de lui, observé sans autre calcul que $\frac{17}{18}$ est positif et inférieur à 1 ($\frac{18}{18}$), a-t-il autant démontré sa maîtrise du type de tâche ?

Que se passerait-il s'il avait eu à traiter la question sans calculatrice avec $\frac{18}{17}$, avec $\frac{259}{72}$, ou plus tard avec $\frac{-259}{72}$ ou $\frac{3,252}{0,27}$?

On pourra objecter qu'il suffirait de poser un nombre suffisamment grand de questions appartenant au même type de tâche, mais ce n'est pas ce que l'on fait, ni ce que l'on peut raisonnablement faire dans les évaluations classiques.

Il ne s'agit donc pas ici de condamner ces évaluations mais seulement d'attirer l'attention sur les interprétations hâtives que sont souvent faites. Il s'agira aussi de montrer comment la recherche en didactique peut aider à améliorer la validité de ces évaluations.

L'ÉVALUATION DOIT-ELLE INTERESSER LA RECHERCHE EN DIDACTIQUE ?

Selon une définition proposée par Guy Brousseau, la didactique des mathématiques est la « *Science des conditions spécifiques de l'acquisition provoquée des connaissances mathématiques* ».

Dans ces conditions, si l'on admet que l'évaluation est de nature à perturber ou à conditionner ces acquisitions, il est clair que l'évaluation sous ses différentes formes ne peut manquer d'intéresser la recherche en didactique.

Il en est de même si l'on considère que, dans la plupart des cas, l'évaluation prétend avoir un effet positif sur le fonctionnement du système et sur les apprentissages. Cela, qu'il s'agisse des évaluations conduites par l'enseignant dans sa classe, qu'elles soient présentées comme formatives ou sommatives, ou qu'il s'agisse des évaluations externes. Il suffit par exemple de consulter rapidement les rapports des études PISA pour voir que ces études visent d'abord à améliorer la qualité de l'enseignement. Cela, évidemment, dans un sens favorable aux analyses concernant la nécessité d'adapter les systèmes d'enseignement aux besoins du monde moderne tels que l'OCDE les définit.

Dans tous les cas, l'une des questions qui doit intéresser la didactique est celle de la validité des évaluations.

On souhaite évaluer quelque chose (un objet o !). On met alors en place une évaluation. Qu'est-ce que cette évaluation pourra nous dire sur cet objet ? Quel est son pouvoir de dire quelque chose de vrai sur cet objet (pouvoir véridictionnel !) ? On sait bien que ce pouvoir, au mieux n'est que partiel ; souvent l'évaluation manque tout ou partie de l'objet et peut même tromper l'évaluateur.

L'enseignant qui ne se contente pas de lire un résultat, de compter des points sur une copie, ou de remplir des cases, sait assez bien, selon sa formation, compenser partiellement les défauts d'une évaluation en faisant intervenir son propre jugement. Le jugement professoral est certainement le complément nécessaire de toute évaluation valide et l'antidote de toute évaluation défailante. La recherche en didactique est de nature à informer et à former un jugement professoral lui-même valide.

En ce qui concerne les évaluations de masse, on sait qu'elles ont hérité de la psychométrie, discipline qui se préoccupe peu de didactique. Certes les méthodes ont évolué et sont de plus en plus sophistiquées (construction d'échelles et utilisation de méthodes probabilistes telle que l'IRT (théorie des réponses aux items)). Ces méthodes, souvent opaques pour le non spécialiste, ne prennent que partiellement en compte la question de la validité. Il y a là, me semble-t-il, un champ de recherche qui mériterait d'être davantage exploité par la recherche en didactique.

Une évaluation peut cependant être valide ou assez valide (rendant bien compte de ce qu'elle prétend observer) et porter sur des objets de peu d'intérêt, du moins pour la formation mathématique. À côté de la validité, et avant elle, se pose alors la question de la pertinence des choix. Sachant à quel point les études de type PISA, ou CEDRE pour la France, ont une influence sur les choix de programmes, cette question de la pertinence, au-delà des analyses de nature idéologique qui sont légitimes mais qui doivent être faites ailleurs, intéresse aussi la cohérence épistémologique des enseignements. De ce point de vue l'étude de la pertinence a évidemment sa place dans les recherches en didactique des mathématiques.

Les résultats des études évaluatives, qu'elles soient nationales ou internationales, sont le plus souvent présentées et analysées directement par ceux-là même qui les produisent, ce qui est bien sûr légitime. Mais ensuite ces résultats et ces analyses sont utilisés tels quels par les divers commentateurs (les médias, mais aussi les politiques et les sociologues qui s'emparent de ces études). Comme si le thermomètre ne pouvait plus être interrogé.

Certes de nombreuses critiques plus ou moins justifiées sont formulées par des chercheurs de tous domaines, mais elles restent le plus souvent confidentielles et, en général sans rapport avec la didactique.

La recherche en didactique des mathématiques aurait certainement beaucoup à gagner à se saisir de ces questions de validité et de pertinence et à les étudier du seul point de vue de leur rapport avec les organisations didactiques.

Des travaux récents et en particulier ceux présentés dans le cadre de ce séminaire par Brigitte Grugeon, Nathalie Sayac et Nadine Grapin montrent que, sur ces questions, des avancées importantes ont été faites ces derniers temps

Des questions que l'évaluation pourrait poser à la didactique

De nombreux travaux de didactique des mathématiques utilisent des instruments d'exploration des comportements des élèves dans une situation donnée, instruments qui ressemblent souvent à s'y méprendre à ceux habituellement utilisés dans l'évaluation scolaire. L'évaluation qui se préoccuperait de validité didactique pourrait demander à la didactique de l'aider dans le choix de ses outils. Il pourrait s'établir, et semble-t-il, il y a déjà ici et là, des collaborations qui pourraient être fructueuses (cf. les communications faites dans le cadre de ce séminaire).

La didactique pourrait en particulier proposer des questions pour l'évaluation, questions qui auraient été analysées d'un point de vue didactique et du point de vue de la validité.

Par analyse didactique d'une question d'évaluation nous entendons :

- L'analyse de la tâche.
- Le repérage de la tâche dans une hiérarchie des niveaux cognitifs telle que celle de Régis Gras.
- La place que le type de tâche correspondant occupe dans l'édifice mathématique

- L'insertion probable ou possible de ce type de tâche dans un processus d'apprentissage.
- L'assurance que cette tâche particulière optimise l'information qui sera recueillie sur le niveau de maîtrise de ce type de tâche.
- Les précautions concernant l'utilisation de cette tâche en situation d'évaluation, de façon à ne pas perturber le processus didactique et, de préférence à l'améliorer (évaluation formative).

L'évaluation est soumise aujourd'hui au défi de l'évaluation des compétences. On peut à bon droit rejeter la notion compétence en qualifiant ce mot de « concept mou », mais l'évaluation pourrait demander à la didactique de l'aider à structurer ce concept comme elle pourrait lui demander de l'aider à rejeter les fausses évidences. Quelle que soit la définition que l'on donne à la notion de compétence, le savoir, l'articulation, et le sens des connaissances ne peuvent être qu'au centre des réflexions sur la question.

Les outils issus de la recherche en didactique tels que les outils d'analyse d'Aline Robert et l'analyse statistique implicative pourraient être mis à profit pour valider ou pour rejeter certaines constructions élaborées pour évaluer des compétences a priori mal définies. Ils pourraient aussi être utilisés pour définir de nouvelles compétences. L'analyse implicative en particulier permet de repérer des proximités cognitives qui ne sont pas évidentes.

L'évaluation pourrait aussi demander à la didactique de lui préciser les points clés des divers apprentissages et donc l'aider à distinguer parmi les savoir-faire (maîtrises de techniques) ceux qui sont des portes ouvertes sur des développements significatifs dans le cours de l'étude et ceux qui n'ont pas d'avenir et sont destinés à l'oubli.

Des questions que la didactique pourrait poser à l'évaluation

La didactique pourrait tout d'abord demander à l'évaluation de démontrer son utilité et son efficacité. L'emprise évaluative sur le système éducatif est devenue trop grande pour que l'exposé des intentions soit suffisant.

Un point qui concerne plus spécialement la didactique est celui de l'évaluation formative. Dans la plupart des pays, elle est devenue le leitmotiv de toute politique éducative. Les intentions sont connues et semblent parfaites : accompagner l'élève dans ses apprentissages et l'aider à apprendre en intervenant lorsqu'il rencontre un obstacle ; cela d'une façon bienveillante, non productrice de stress...

Oublions le stress qui, dans certains cas peut être bénéfique et voyons comment ces intentions sont traduites dans les faits.

D'une part l'évaluation formative (que j'ai largement défendue ailleurs) concerne a priori la relation enseignant-enseigné. Elle n'a pas à nourrir directement l'information de la chaîne qui va de l'enseignant au ministre en passant par le chef d'établissement, les parents, les différents corps de contrôle, etc....

Ce n'est malheureusement pas toujours le cas, et l'évaluation dite formative est souvent devenue l'alibi pour une évaluation permanente de nature plutôt sommative. Cela vaut pour l'enseignant dans sa classe qui, forcé ou non est amené à pervertir son évaluation formative pour rendre les comptes qui lui sont demandés ; cela vaut aussi pour les évaluations institutionnelles qui, elles aussi, se veulent formatives. La lecture du rapport de l'OCDE sur la question (2005) montre bien comment, dans certain cas l'évaluation formative dans la classe est utilisée de façon indue pour nourrir le processus de contrôle de l'activité des enseignants et des acquisitions des élèves ; cela sans que le rapport y trouve à redire.

Des démarches alternatives sont régulièrement proposées et utilisées pour l'évaluation des élèves. C'était le cas de l'évaluation critériée et maintenant de l'évaluation par compétences ou encore de l'évaluation dite « par contrat de confiance ». La recherche en didactique pourrait se pencher sur ces démarches pour étudier la façon dont elles modifient le contrat didactique et

sur la façon dont elles influent sur la qualité du savoir construit.

Jusqu'à présent, la recherche en didactique n'a pas pris très au sérieux la question de l'évaluation. Elle est pourtant bien placée pour lui poser les questions qui pourraient l'aider à sortir d'un certain simplisme ...

SOLUBILITE DE L'EVALUATION DANS LA DIDACTIQUE ?

Nous l'avons dit, évaluation et didactique des mathématiques sont apparues à peu près en même temps dans le paysage. Depuis les années 70 les recherches se sont multipliées et diversifiées, les théorisations se sont consolidées. Évaluation et didactique ont suivi des chemins parallèles, ne se rencontrant qu'à de rares occasions, comme cela fut le cas lors de l'intrusion de l'évaluation formative.

En fin de compte, en dépit d'exceptions que l'on peut souhaiter de plus en plus nombreuses, les enseignants continuent à avoir des pratiques d'évaluation dirigées davantage vers la notation que vers l'analyse de l'activité des élèves, à moins qu'elles ne soient principalement conditionnées par des demandes institutionnelles formelles telles que le remplissage de grilles dont la validité et l'utilité reste à démontrer. Sauf exceptions, la formation des enseignants de mathématiques en matière d'évaluation reste anecdotique ; elle intègre rarement la question des contenus et encore moins la didactique.

De leur côté, les institutions en charge des évaluations de masse ont le plus souvent ignoré non seulement la recherche en didactique, mais aussi les processus didactiques et même... l'enseignement.

D'une façon générale, les questions didactiques ont peu intéressées les spécialistes de l'évaluation, tandis que les chercheurs en didactique des mathématiques se détournaient des questions relatives à l'évaluation.

À un certain moment on aurait pu penser que la théorie didactique avait vocation soit à absorber la didactique soit qu'elle devait la tenir résolument à l'écart. D'où la question de la solubilité de l'évaluation dans la didactique.

Après ce retour en arrière, la question apparaîtra sans doute assez oiseuse. Il est aujourd'hui évident que l'évaluation n'est pas soluble dans la didactique et même que la question est sans intérêt. Le terme recouvre trop de pratiques et trop d'objectifs différents, chacune d'entre elles et chacun d'entre eux étant plus ou moins éloigné des préoccupations de la didactique.

L'écriture de ce texte m'a amené à relire certains textes anciens et à découvrir de nouveaux travaux. Ce faisant j'ai un peu l'impression d'avoir enfoncé ici des portes largement ouvertes. Les travaux de Marc Vantourout et Rémi Goasdoué, par exemple, présentés dans ce séminaire, sont la preuve que certaines de mes interrogations sont aujourd'hui largement dépassées. Il en est de même des autres travaux présentés dans ce séminaire. La question de l'évaluation est bien devenue un sujet qui intéresse la recherche en didactique.

En particulier, Marc Vantourout et Rémi Goasdoué (2014) attaquent résolument la question en articulant des entrées didactiques et psychologiques sur la question. Ils montrent en particulier comment une certaine « *conjonction de travaux en didactique des mathématiques et en psychologie permet de renouveler profondément les débats sur la validité* ».

CONCLUSION

L'évaluation n'est donc pas un concept de la didactique. Toutefois, l'omniprésence des faits d'évaluation (même si ceux-ci ne sont pas homogènes), la demande institutionnelle et sociale, et l'évidente participation d'une partie de ces faits à la genèse et à l'actualisation du contrat

didactique et au processus de transposition didactique, rend inévitable une plus grande attention des didacticiens à leur égard.

Après quarante années de chemins plus ou moins parallèles on voit mieux ce que chaque discipline peut demander ou apporter à l'autre. On voit mieux aussi les complémentarités et la nécessité d'une collaboration renforcée de chercheurs d'origines différentes, cela tant d'un point de vue théorique que pour la construction de dispositifs pratiques (évaluations de masse par exemple, mais aussi examens, domaine qui est trop souvent oublié).

L'évaluation et la recherche en évaluation peuvent en effet demander à la didactique des outils lui permettant de mieux assurer la validité de ses entreprises. Elle pourrait accepter ou même demander que la didactique assure un contrôle minimum sur ses entreprises.

La recherche en didactique des mathématiques peut demander à l'évaluation de mieux se poser la question des conséquences de ses actions sur la genèse du contrat didactique et donc sur la qualité des apprentissages.

D'une façon générale la collaboration entre chercheurs et praticiens d'origines différentes est indispensable ; elle se met en place et cela laisse espérer de nouveaux progrès.

REFERENCES BIBLIOGRAPHIQUES

BROUSSEAU, G. (1978) Évaluation et théories de l'apprentissage en situations scolaires. Exposé fait à ICME 5 (Polycopié)

BROUSSEAU, G. (1978) l'observation des activités didactiques. Actes CIEAM 1977 (repris dans sa thèse et dans un article de la revue *Française de Pédagogie* n°45/1978)

CHEVALLARD, Y. (1999) L'analyse des pratiques enseignantes en théorie anthropologique du didactique. *Recherches en Didactique des Mathématiques* Vol 19/2

CHEVALLARD, Y. (2004) Le moment de l'évaluation, ses objets, ses fonctions : déplacements, ruptures, refondation. Site Web - <http://yves.chevallard.online.fr/>

CHEVALLARD, Y. FELDMANN, S. (1986) Pour une analyse didactique de l'évaluation - Irem d'Aix Marseille

GAGATSI, A. (1982) *Discrimination des scores au test de closure et évaluation de la compréhension des textes mathématiques*. Thèse, Université Louis Pasteur, Strasbourg.

GLAESER, G. (1995) : Fondements de l'évaluation en mathématiques. Cours de DEA Publication de l'APMEP N°96.

GRAS, R. (1979) *Contribution à l'étude expérimentale et à l'analyse de certaines acquisitions scientifiques et de certains objectifs didactiques en mathématiques*. Thèse, université de Rennes.

MARGOLINAS, C (1989) *Le point de vue de la validation : essai de synthèse et d'analyse en didactique des mathématiques*. Thèse – Grenoble 1.

MONS, N. (1999) *Les effets théoriques et réels de l'évaluation standardisée*. Eurydice

OCDE (2005) L'évaluation formative. Pour un meilleur apprentissage dans les classes secondaires (CERI – Centre de Recherche pour l'Innovation dans l'Enseignement).

PILET, J. (2012) *Parcours d'enseignement différencié appuyés sur un diagnostic en algèbre élémentaire à la fin de la scolarité obligatoire : modélisation, implémentation dans une plateforme en ligne et évaluation*. Thèse Paris Diderot.

PLUVINAGE, F. (1977) *Difficultés des exercices scolaires en mathématiques (étude des comportements de réponse par enquêtes à plusieurs modalités)*. Thèse de doctorat. Strasbourg.

REGNIER, J. C. (1983) *Auto-évaluation et autocorrection dans l'enseignement des mathématiques et de la statistique. Entre praxéologie et épistémologie scolaire*. Thèse, Université Louis Pasteur, Strasbourg. Repris et complété (1993) sous le titre : De l'importance du vrai et du faux dans la classe de mathématiques. La pensée Sauvage

SAYAC, N. GRAPIN, N. (2015) Évaluation externe et didactique et didactique des mathématiques : un regard croisé nécessaire et constructif. *Recherches en Didactique des Mathématiques* Vol 35/1.

VANTOUROUT, M. & GOASDOUE, R. (2014) Approches et validités psycho-didactique des évaluations. *Education et Formation* n°302.

REFERENCES NON CITEES DANS LE TEXTE

ARTIGUE, M. & WINSLØW, C (2010) Comparaisons internationales sur l'enseignement des mathématiques : un point de vue porté par la théorie anthropologique du didactique. *Recherches en Didactique des Mathématiques* vol. 31 n°1, pp. 47-82

BODIN, A. (1997) L'évaluation du savoir mathématique - Questions et méthodes. *Recherches en Didactique des Mathématiques* 17(1)

BODIN, A. (2007) Dissonances et convergences évaluatives. De l'évaluation dans la classe aux évaluations internationales : quelle cohérence ? *Bulletin de l'APMEP* n° 474 pp 47-79

CANTAT, A. (2009) Historique de l'évaluation des apprentissages : de l'enseignement des jésuites à l'approche par compétences. Université de Laval.

CHEVALLARD, Y. (1989) Evaluation, véridiction, objectivation - la relation didactique comme caprice et miniature. In *L'évaluateur en révolution - actes du colloque ADMEE 89 – INRP*

GUIMARD, P. (2010) *L'évaluation des compétences scolaires*. Presses universitaires de Rennes

RODITI, É. (2012) *Un point de vue didactique sur les questions d'évaluation en éducation*. In : Aline Robert ; Jacqueline Penninckx ; Marie Lattuati. Une caméra au fond de la classe de mathématiques, Presses universitaires de Franche-Comté, pp.275-289, 2012, Pratiques & Techniques