

FORMATION ET PRATIQUE DES ENSEIGNANTS

REGARD PORTÉ SUR L'ÉVALUATION EN EPS

Résumé : L'évaluation est une thématique au centre des préoccupations des enseignants. Les avancées théoriques diffusées largement en formation et faisant l'objet de nombreuses publications font pourtant l'objet de réajustements lors des pratiques en classe. Nous avons interrogé 179 enseignants d'EPS du secondaire au moyen d'un questionnaire dans le but d'identifier la manière dont les enseignants adaptent sur le terrain les théories sur l'évaluation dans cette discipline singulière. Nos résultats montrent des pratiques très variées et une persistance du côté traditionnel et sommatif de l'évaluation : l'évaluation formative est régulièrement détournée de sa fonction initiale en étant notée, créant une confusion des finalités de l'évaluation en EPS.

Mots-clés : évaluation sommative, évaluation formative, notation, pratique pédagogique, formation des enseignants.

Depuis la création des IUFM en 1989, la formation des enseignants est régulièrement réformée, tiraillée entre la forte demande de professionnalisation et l'apport de données théoriques issues de la recherche. Si la théorie et la pratique apparaissent interdépendantes pour assurer une formation cohérente, les liens entre ces deux entités sont complexes ; la mise en pratique des savoirs enseignés en formation fait émerger toute la complexité de la situation d'enseignement, nécessitant une adaptation au contexte qui est peu prévisible surtout pour les enseignants débutants (Rayou et Zanten, 2004). En dépit des recommandations officielles et des apports issus de la recherche, les enseignants disposent d'une liberté pédagogique et du choix d'une démarche didactique (Bressoux *et al.*, 1999). Effectivement, les textes officiels rendent compte des finalités de l'enseignement, et l'enseignant détermine les moyens d'y parvenir. Les méthodes utilisées sont largement personnalisées et reflètent pour une part l'identité de l'enseignant, rendant l'analyse compréhensive des pratiques à la fois très pertinente mais aussi complexe. Les travaux menés par le courant de l'École de Chicago (Becker, 1963) prennent en considération dans l'analyse des conduites humaines le poids du contexte et la signification que les individus se font des situations. Cette conception interactionniste permet de porter un regard sur la tradition qui perdure, sur les pratiques qui évoluent peu, même si les recommandations officielles (par exemple le rapport de l'Inspection Générale 2007-048) ou certaines recherches incitent à faire autrement (Marcel *et al.*, 2007). Dans ce cadre, il est nécessaire de distinguer, dans le travail enseignant, ce qui relève de la prescription et des réalisations effectives des enseignants, car comme l'indique Leplat (1986, p. 26) : « l'activité dépend bien de la tâche (ou si l'on préfère des conditions externes de travail), mais aussi des caractéristiques des individus qui l'exécutent ».

Plus spécifiquement nous nous centrerons ici sur l'activité évaluative des enseignants du secondaire. Étant obligatoire selon la loi d'Orientation sur l'éducation du 10 juillet 1989 où il est mentionné que les professeurs assurent le suivi individuel et l'évaluation des élèves, l'évaluation sous toutes ses formes implique un recueil d'informations, une confrontation entre un référé et un référent dont le résultat amène à une prise de décision (Hadjji, 1997). Aujourd'hui, tout un chacun s'accorde pour dire que l'évaluation est un passage obligé dans un parcours d'apprentissage ; elle fait partie des compétences à développer chez les enseignants qui, en formation, prennent connaissance de ressources théoriques et vivent parallèlement des ateliers d'analyse de pratiques.

Même si l'évaluation est un sujet partagé par les disciplines, elle est l'objet de publications transversales (Figari, 1994 ; Scallon, 1982) mais aussi plus spécifiques, comme faisant partie des didactiques disciplinaires. Nous avons choisi de réaliser une étude auprès des enseignants en éducation physique et sportive (EPS), discipline singulière puisque la formation des enseignants d'EPS est réalisée dans les UFR STAPS où les étudiants ont très tôt la possibilité de réaliser des stages, chaque semestre, en établissement scolaire. De plus, les enseignements scientifiques et la pratique des activités physiques sont la plupart du temps orientés vers l'analyse des pratiques, vers les difficultés rencontrées par les débutants, voire par les élèves. Cette discipline possède aussi des modalités d'évaluation spécifiques comme par exemple le contrôle en cours de formation lors des examens scolaires ; les enseignants évaluent ainsi directement la portée de leur enseignement. Une étude publiée par le Ministère de l'Éducation Nationale (Braxmeyer et al., 2004) montre que les enseignants d'EPS se distinguent largement des enseignants des autres disciplines en utilisant davantage les évaluations à des fins diagnostique et formative. Ces mises en œuvre sont utilisées pour créer des groupes de besoin et individualiser les apprentissages. Notre étude vise ainsi à interroger les pratiques évaluatives de ces enseignants au regard de deux hypothèses principales : tout d'abord, nous supposons que ces pratiques sont loin d'être standardisées, et qu'elles évoluent notamment avec le niveau d'expertise. Ensuite, même si les évaluations formatives et sommatives ne sont pas confondues par les enseignants, certaines pratiques – comme l'utilisation de la notation – peuvent paraître contradictoires et provoquer un décalage entre les intentions formatives de l'enseignant et les situations proposées, par exemple lorsque les effets secondaires de la notation sont peu pris en compte.

L'ACTIVITÉ ÉVALUATIVE DES ENSEIGNANTS : LA FONCTION SOMMATIVE DOMINANTE.

Au niveau diachronique, l'évaluation dont la finalité première fut le contrôle des apprentissages, a été redéfinie, élargie pour devenir un procédé au service des apprentissages. La réflexion sur l'évaluation est d'une importance sans équivoque et de nombreuses questions quasi quotidiennes témoignent de cet intérêt : qu'évalue-t-on ? À quel moment et pourquoi faire ? Le fort impact de l'évaluation sur les apprentissages et les productions d'élèves a notamment été démontré par Rayou (2002) à partir d'une étude sur les copies de philosophie de lycéens ; l'évaluation certificative amènerait davantage vers la standardisation des productions d'élèves que vers l'affirmation de leurs idées, de leur créativité dans cette discipline où justement la normalisation n'est pas recherchée.

En réponse aux théories déclinant les différentes formes d'évaluations et leurs intérêts (Allal et al. 1991 ; Perrenoud, 1998) les dérives et effets parfois né-

fastes ou paradoxaux (Rayou *op. cit.* ; Merle, 2007 ; Vantourout et Goasdoue, 2010), il paraît pertinent d'analyser dans quelle(s) mesure(s) les pratiques évoluent dans la direction des approches issues des recherches. Par exemple pour Perrenoud (*op. cit.*), ce qui transparait nettement, c'est la primauté de la fonction traditionnelle et sommative de l'évaluation. Malgré les avancées qui proposent de nouvelles pistes d'enseignement visant à optimiser les processus d'apprentissage, Perrenoud relève que le côté traditionnel perdure, se manifestant par la persistance de la fonction sommative de l'évaluation et de ses rapports étroits entretenus avec la notation. Pourtant, la fonction sommative n'est plus la référence unique en termes d'évaluation enseignée en formation. Les recherches menées autour de l'évaluation formative puis formatrice s'enchaînent depuis une quarantaine d'années (Scallon, 1982 ; Delorme, 1987 ; Cardinet *et al.*, 2001). Au regard des propos de Perrenoud (*op. cit.*) – partagés par bon nombre de formateurs et conjugués à la vingtaine d'observations empiriques que nous avons réalisées durant des visites de terrain de futurs enseignants (en amont de notre enquête) –, l'évaluation pose un véritable problème d'adéquation entre ce qui est enseigné et les modalités d'évaluation utilisées. Dans ce contexte, nous nous sommes interrogés sur la mise en place de trois formes d'évaluation régulièrement utilisées en EPS : (i) l'évaluation diagnostique – dont l'objectif est de faire le point sur les connaissances antérieures et spontanées des élèves allant servir de point d'appui aux apprentissages nouveaux (De Ketele, 1988) – ; (ii) l'évaluation formative, qui a aussi une fonction pédagogique en participant à la régulation des apprentissages et au développement de l'élève dans le sens du projet éducatif (Perrenoud, *op. cit.*) ; (iii) et l'évaluation sommative qui clôture une phase d'apprentissage, sous la forme d'une vérification de l'acquis, en sanctionnant la performance (Talbot, 2012).

Sans avoir pour ambition de comprendre la complexité des pratiques ni de généraliser nos résultats, nous avons élaboré un questionnaire afin que les enseignants puissent décrire leurs pratiques et exprimer leur opinion à propos des évaluations qu'ils mettent en place. En EPS, à l'instar de David (2000), la mise en œuvre des évaluations pose des problèmes spécifiques. Il cite par exemple le biais d'enseigner au final que ce qui est évalué pour mettre en œuvre des procédures d'évaluation précises et objectives. Nous sommes partis du postulat selon lequel l'enseignant adapte sa mise en œuvre selon ses conceptions et la situation dans laquelle il évolue, tout en s'appuyant sur les approches actuelles orientées vers une approche formative. L'enquête présentée ici a pour objet d'appréhender les différences de mise en œuvre qui traduisent des interprétations différentes d'une culture commune de l'évaluation. Elle témoigne des difficultés rencontrées par les enseignants à mettre en pratique les connaissances qu'ils ont acquises et à ce détacher du versant sommatif. Les difficultés pointées par les théoriciens ou relevées lors de nos observations empiriques sur le terrain vont-elles apparaître dans les résultats de notre enquête ?

Dans une première partie nous présenterons la méthodologie de cette recherche. Ensuite, les résultats les plus saillants relatifs à l'utilisation des évaluations formative et sommative dans l'enseignement en EPS seront déclinés. Enfin, nous discuterons les résultats obtenus au regard des théories actuelles.

MÉTHODOLOGIE

Les enseignants interrogés

La réalisation d'un questionnaire rempli sous couvert de l'anonymat nous a permis de recueillir 179 réponses exploitables. L'objectif de cette étude a été présenté comme une recherche de type universitaire relative à l'activité évaluative des enseignants en EPS.

Pour obtenir une quantité de réponses suffisante, nous avons diffusé ce questionnaire par différents moyens. Tout d'abord, en contactant des Inspecteurs Pédagogiques Régionaux en EPS : deux d'entre eux ont accepté de diffuser notre questionnaire sur leur site académique, mais le nombre de répondants a été très faible (une seule réponse dans une académie de l'Île de France). Ensuite, toujours par voie hiérarchique, nous avons contacté des chefs d'établissements : par courrier électronique nous avons adressé le questionnaire à plus de 500 établissements du secondaire dans diverses académies, ce qui nous a permis de récolter une centaine de réponses. Enfin, nous l'avons diffusé par voie directe, lors d'une formation continue dans l'académie de Lyon, de rencontres au sein de l'Union Nationale du Sport Scolaire dans l'académie de Créteil, et plus simplement entre collègues.

Le profil des enseignants interrogés est proche du profil des enseignants d'EPS en 2009 au niveau national¹. Ainsi, 55% ($n = 99$) des enseignants interrogés sont des hommes, contre 54% au plan national. La part d'agrégés est plus importante, soit 16% ($n = 28$) dans notre enquête. 77% ($n = 138$) sont certifiés et 7% sont d'une autre catégorie (stagiaire, contractuel...). Enfin, la moyenne d'âge au plan national est de 42 ans et 55% de nos répondants ont entre 31 et 50 ans ; 21% ont moins de 30 ans et 25 % plus de 50 ans.

Une majorité enseigne en collège (56% ; $n = 100$ dont 27 en SEGPA²) et près d'un tiers enseigne dans deux types d'établissements, essentiellement des lycées, le plus courant étant les enseignants de lycée général et technologique ($n = 36$).

Le questionnaire

Pour concevoir le questionnaire, nous avons d'abord réalisé trois interviews d'enseignants en EPS d'une durée de 30 à 45 minutes, pour orienter les thèmes et les questions abordés. De là, nous avons défini trois champs de questionnement répartis en quinze questions :

- Une question ouverte, une échelle d'attitude à 4 niveaux et 13 questions fermées voire mixtes concernent les mises en œuvre et conceptions des trois formes d'évaluation : diagnostique, sommative et formative. Par exemple :

Question 1 : En une phrase, pour vous, qu'est-ce que l'évaluation ?

- Deux questions fermées se réfèrent à la notation, sa fréquence et relation à l'évaluation. Par exemple :

Question 6 : Vous arrive-t-il d'utiliser une évaluation formative pour noter vos élèves ? Jamais : ; Oui parfois : ; Oui souvent : ; Oui, systématiquement :

- Une question fermée et l'échelle d'attitude déjà mentionnée abordent l'effet de l'évaluation sur le stress et la motivation des élèves. Par exemple :

¹ Cf. « Être professeur d'éducation physique et sportive en 2009 » – *Les dossiers évaluations et statistiques* - DEPP - N°195 - février 2010 <http://www.education.gouv.fr/cid50749/etre-professeur-d-education-physique-et-sportive-en-2009.html>.

² Section d'Enseignement Général et Professionnel Adapté.

Question 7 : Pensez-vous que les élèves puissent adopter des conduites motrices (dans l'action, c'est-à-dire hors conduites sociales) différentes entre, d'une part, les séances d'apprentissage hors évaluation, et d'autre part :

Les situations d'évaluation formative : Oui non

Les situations d'évaluation sommative : Oui non

La passation dure de 15 à 20 minutes, un extrait du questionnaire est présenté en annexe. Les définitions des différentes formes d'évaluation abordées sont données en note de bas de page pour éviter les confusions liées aux nombreuses acceptions, points de vue ou représentations sur cet objet.

Le questionnaire a été testé à deux reprises avant sa forme définitive, une première fois au sein d'un séminaire de recherche, et la seconde auprès de huit enseignants d'EPS, différents de ceux ensuite interrogés.

Recueil des données

Le recueil et l'analyse des données ont été réalisés à l'aide du logiciel *question data 6* permettant de croiser les données recueillies pour tester statistiquement l'existence de relations entre différentes variables (calcul de khi2 total et khi2 par case pour chaque tableau de contingence). Les questions ouvertes ont été traitées manuellement par une analyse de contenu qui consiste à repérer des fréquences d'apparition de mots, et qui crée des catégories en fonction des réponses apportées (Bardin, 1977). Quelques tableaux de contingence sont présentés en annexe afin de rendre plus lisible les résultats.

RÉSULTATS GÉNÉRAUX

Les résultats seront abordés comme suit : ceux issus de la question ouverte relatifs à la définition de l'évaluation. Puis, nous présenterons les conceptions des enseignants relatives aux fonctions de l'évaluation, pour enfin approcher leurs pratiques de la notation.

Conceptions de l'évaluation dans sa globalité

Aux questions liées à l'évaluation sommative, nous constatons que tous les enseignants mettent en place des évaluations, notent leurs élèves et affirment les informer des critères d'évaluation : pour la majorité (61%), au cours du cycle³ et pour les autres, dès la première séance. La mise en place d'une évaluation diagnostique est aussi très répandue (98%) bien que non obligatoire, ceci confirmant les résultats de Braxmeyer (*op. cit.*) qui témoignent du fort recours à ce type d'évaluation en EPS.

Cependant, bien que ces pratiques générales semblent partagées, des divergences apparaissent déjà au regard de la définition de l'évaluation : d'abord, pour 85% des sujets interrogés, l'évaluation est un recueil de données, une mesure. Une distinction émerge cependant puisque 65% réfèrent cette évaluation à un niveau de départ, des objectifs, des exigences. Parmi eux, 38% comparent les données recueillies avec ce qui a été enseigné, 41% avec un barème ou des exigences spécifiques. De façon minoritaire 4,5% évaluent leurs élèves en fonction de leur niveau de départ et 16% les comparent entre eux. Notons que 11% des enseignants précisent qu'ils évaluent pour mettre une note.

³ Nous avons estimé dans l'enquête qu'un cycle d'apprentissage dure 8 séances.

Ensuite, 93% définissent ce qu'ils évaluent (cf. tableau 1). Même si nous pouvons mettre en relief des termes variés, rappelons que les acceptions autour d'un même mot peuvent être aussi différentes selon les personnes.

Ce qui est évalué : réponses spontanées données en question ouverte.	% (n = 166)
Compétences	33% (n = 54)
Acquisitions, apprentissages	23% (n = 39)
Niveau de pratique	23% (n = 39)
Progrès	8% (n = 14)
Savoirs	4% (n = 7)
Performance	3% (n = 5)
Conduites motrices	2% (n = 3)
Réussite	2% (n = 3)
Capacités	1% (n = 2)

Tableau 1 : ce que les enseignants disent évaluer, en pourcentages parmi ceux le précisant dans leur définition (n = 166).

Les divergences observées dans ces réponses reflètent la variété des termes employés notamment dans les discours officiels. Par exemple, le terme *compétence* est le plus représenté, faisant référence à la prédominance de cette notion dans les programmes scolaires, apparue avec la Charte des programmes du 20 février 1992, puis actuellement dans le socle commun de connaissances et de compétences (décret du 11 juillet 2006) et la charte des programmes de 2014⁴. La référence aux apprentissages est une donnée plus générale qui témoigne des modifications de la conduite du sujet dues à de nouvelles acquisitions. La définition de niveau de pratique ou d'habileté est aussi très utilisée en EPS pour se démarquer de la prise en compte de la seule performance motrice et ce, depuis l'arrêté du 17 juillet 1984, tout comme l'évaluation des progrès qui a longtemps figuré dans les évaluations au baccalauréat et au brevet des collèges. La définition apportée à l'évaluation par les enseignants est ainsi plutôt proche des concepts employés dans les textes officiels.

Fonctions de l'évaluation

En premier lieu, l'évaluation diagnostique est une pratique largement répandue chez les enseignants interrogés, comme l'ont montré Braxmeyer *et al.* (*op. cit.*) : 41% des enseignants l'utilisent systématiquement et 57% parfois. Cette fonction pronostique est d'utilité directe à la spécification des objectifs de travail et des contenus.

En second lieu, suite à l'intérêt porté à la définition générale de l'évaluation, nous avons souhaité approfondir les conceptions des enseignants en leur proposant une échelle d'attitude (cf. tableau 2) relative aux fonctions de l'évaluation et ses effets sur les élèves.

Dans le cadre de la professionnalité enseignante, l'évaluation est protéiforme et bigarrée ; fait saillant, la majorité (70%) s'accorde à penser que « l'évaluation formative est utile pour aider mes élèves à progresser ». Cette intention est renforcée par les 75% qui se désignent comme étant assez ou tout à fait d'accord avec « j'évalue pour optimiser l'apprentissage ». Ces données montrent une volonté de se détacher de la fonction traditionnelle de l'évaluation perçue comme une étape de contrôle. D'ailleurs, interrogés sur la mise en place d'évaluations

⁴ Cette charte est consultable sur internet :

http://cache.media.education.gouv.fr/file/04_Avril/37/5/charte_programme_csp_312375.pdf

formatives, 97% des enseignants disent les utiliser, régulièrement ou non. Or, l'intention de placer l'évaluation au service de l'apprentissage est-elle suivie de mises en œuvre réellement en adéquation avec cet objectif affiché ?

<i>n</i> = 179	<i>Pas du tout d'accord</i>	<i>Peu d'accord</i>	<i>Assez d'accord</i>	<i>Tout à fait d'accord</i>
J'évalue pour noter.	9 %	22%	43%	26%
L'évaluation formative est utile pour aider mes élèves à progresser.	1%	1%	27%	70%
J'évalue pour maintenir l'engagement de mes élèves.	13%	26%	44%	17%
J'évalue pour optimiser l'apprentissage.	5%	21%	49%	26%
L'évaluation sommative motive mes élèves.	3%	22%	53%	22%
L'absence de notation entraîne un désengagement de mes élèves.	11%	34%	43%	12%
Évaluer mes élèves me permet de les classer.	31%	33%	28%	8%

Tableau 2 : choix relatifs aux différentes propositions selon une échelle à 4 niveaux. Résultats présentés en pourcentages.

Évaluation et notation

Au croisement des réponses apportées, nos résultats révèlent que la notation est souvent associée à l'évaluation. Déjà, à l'affirmation « j'évalue pour noter » 69% sont d'accord. Sans que nous ayons spécifié dans la proposition le type d'évaluation (diagnostique, formatif ou sommatif), une majorité y rattache spontanément la notation.

Précisément, la notation est régulièrement présente lors des évaluations formatives. En effet, seuls 37% disent ne jamais noter l'évaluation formative alors que 97% affirment la mettre en place régulièrement ou non. De fait, le nombre de notes – attribuées aux élèves par cycle d'apprentissage – étant en moyenne de 2,5, les 63% d'enseignants notant les évaluations formatives (parfois, souvent ou systématiquement) mettent en moyenne plus de quatre notes par cycle de huit séances. Dans ce cadre, l'évaluation formative est ainsi considérée comme une étape dans l'apprentissage pour certes, renseigner les élèves sur leur niveau, mais de façon quantitative à l'aide d'une note. Au final, en croisant nos données, seulement 11 % des enseignants ($n = 21$) proposent une évaluation formative à chaque cycle et ne la notent jamais.

Malgré la volonté des enseignants de faire progresser les élèves, les résultats témoignent de pratiques n'allant pas toujours dans le sens d'une aide à l'apprentissage, puisque noter une évaluation en change sa finalité.

RÉSULTATS PLUS SPÉCIFIQUES À CERTAINS PROFILS

Les effets de la notation

Les premiers résultats globaux montrent un attachement à l'évaluation formative, mais la présence forte de la notation entache cette volonté d'utiliser l'évaluation au service des apprentissages.

Concernant la fréquence de cette notation, nous constatons une dépendance très significative ($p = 0.01$) entre les enseignants qui notent beaucoup les

élèves, et ceux qui affirment ne pas les noter pour les hiérarchiser. À l'unanimité, ceux qui notent 6 à 8 fois par cycle⁵ disent ne pas évaluer pour classer ($n = 10$). Ici nous observons une représentation des notes qui ne prend pas en compte les effets néfastes qui peuvent en découler. Effectivement, huit de ces mêmes enseignants sont tout à fait d'accord avec le fait que l'évaluation permette de maintenir l'engagement des élèves et qu'elle n'est pas une source de stress pour eux. Plus précisément, concernant les représentations des enseignants liées au stress produit par l'évaluation chez les élèves, nous observons que plus les enseignants pensent qu'elle est source de stress, moins ils mettent de notes en moyenne par cycle (cf. tableau 3).

<i>L'évaluation est source de stress pour les élèves</i>				
	Pas du tout d'accord	Peu d'accord	Assez d'accord	Tout à fait d'accord
Nombre de notes en moyenne	4.09	3.02	2.15	1.94

Tableau 3 : fréquence de notation selon la perception du stress engendré chez les élèves.

Ici les avis sont très partagés concernant les effets que peut avoir l'évaluation sur le stress. Ceux étant les moins convaincus du stress engendré sont les enseignants ayant des classes de Section d'Enseignement Général et Professionnel Adapté (SEGPA) ($n = 27$) ainsi que les jeunes enseignants ($p = 0.01$).

De même, la tendance à penser qu'évaluer et surtout noter a un effet sur l'engagement des élèves révèle aussi des disparités : les certifiés assimilent davantage l'évaluation à une sollicitation de la motivation et donc de la participation des élèves ($p = 0,05$). Ceci s'accroît chez les enseignants interrogés intervenant par ailleurs en SEGPA ($p = 0.01$).

Nous pouvons établir un constat similaire à propos de la fonction de l'évaluation qui consisterait à classer les élèves entre eux. Les résultats de l'échelle d'attitude montrent que 64% des enseignants ne sont pas d'accord avec cette proposition « évaluer mes élèves me permet de les classer ». Or, en nous intéressant aux 36% restant, les données statistiques montrent que les enseignants ayant des classes SEGPA, tout comme les enseignants certifiés sont surreprésentés parmi ceux étant plutôt d'accord avec cette proposition ($p = 0,01$).

La mise en place d'évaluations formatives

Même si la quasi-totalité des enseignants dit mettre en place des évaluations formatives, plus ou moins régulièrement, des divergences apparaissent concernant les mises en œuvre. À l'inverse du précédent bilan relatif à l'engagement des élèves, nous constatons que les enseignants, non d'accord avec le fait que l'absence de notation puisse entraîner un désengagement, sont surreprésentés de façon significative ($p = 0.01$) parmi ceux qui mettent en place à chaque cycle des évaluations formatives et qui ne les notent jamais. Ce profil d'enseignant est aussi surreprésenté parmi ceux n'étant pas du tout d'accord avec le fait d'évaluer pour maintenir l'engagement ($p = 0.002$). Ils dissocient ainsi l'évaluation de la notation ; d'autres moyens sont donc utilisés pour favoriser les apprentissages et susciter l'intérêt, que l'évaluation notée.

L'expérience des enseignants paraît jouer aussi en faveur d'une prise en compte des effets néfastes de l'évaluation et de l'intérêt des évaluations forma-

⁵ Huit étant le nombre de séances de deux heures par cycle proposé dans notre questionnaire.

tives : 42% des enseignants en proposent à chaque cycle et, parmi eux, les plus de 50 ans sont surreprésentés ($p = 0,01$). Ces derniers se démarquent aussi en faisant partie des enseignants qui notent le moins (cf. tableau 4), avec les agrégés ($p = 0,05$).

Age des enseignants interrogés						
		20 à 30	31 à 50	plus de 50	Total	%
Nombre de notes mises par cycle	1	8	- 16	+ 16	40	23%
	2	13	42	16	71	40%
	3	7	25	7	39	22%
	4	2	3	4	9	5%
	5	0	(+) 7	0	7	4%
	6	0	0	0	0	0%
	7	(+) 3	1	0	4	2%
	8	3	3	0	6	3%
Total		36	97	43	176	.
%		20%	55%	24%	100	.

Tableau 4 : effectif des enseignants selon leur âge et leur fréquence de notation ($n = 176$). Tableau issu du logiciel *Question data 6*.

Légende : lorsque la case est statistiquement significative à $p = 0,05$ par rapport à l'effectif théorique, le résultat est précédé d'un « + » ou d'un « - » : « + » : effectif observé supérieur à l'effectif théorique, « - » : effectif observé inférieur à l'effectif théorique.

Le signe entre parenthèses signifie que l'effectif observé est faible pour tirer toute conclusion affirmée.

Les résultats montrent que la grande majorité des plus de 50 ans (32 sur les 43) notent une à deux fois par cycle ; ils sont surreprésentés parmi ceux qui notent le moins. La spécificité des pratiques des jeunes enseignants a déjà fait l'objet de recherches, par exemple Rayou et Van Zanten (*op. cit.*) qui soulèvent leur difficulté à gérer la complexité par manque de professionnalisation, ou encore par Boraita et Issaieva (2013) qui montrent que les conceptions en matière d'évaluation des jeunes enseignants en formation sont différentes de celles des enseignants en activité. De fait, selon l'expérience professionnelle voire le degré d'expertise, les priorités des enseignants ne sont pas les mêmes et engagent des choix pédagogiques différents.

L'utilité des évaluations est ainsi traduite de différentes façons, même si au départ les intentions paraissent semblables. Par ailleurs, si la majeure partie des enseignants est tout à fait d'accord avec le fait que l'évaluation est utile pour aider les élèves à progresser, cet avis ne concerne cependant que la moitié de ceux exerçant en lycée professionnel. L'évaluation formative semblerait moins utile aux élèves en lycée professionnel selon notre population étudiée.

DISCUSSION

Globalement, au vu de la fréquence des évaluations proposées, nous constatons qu'elles sont une préoccupation majeure des enseignants en EPS. Or, bien qu'ils distinguent dans leurs réponses les évaluations sommatives de formatives, différents points nous éclairent sur des mises en œuvre parfois contradictoires, notamment avec les effets attendus de l'évaluation formative : la forte relation entre la notation et l'évaluation soulève certaines questions. Noter une évaluation formative c'est en faire une échéance qui ne permet pas par exemple le droit à l'erreur car la sanction est immédiate. De plus, la validité de la note proposée n'est

pas garantie car celle-ci ne reflète pas le niveau de l'élève mais plutôt les modalités de notation ayant été définies par l'enseignant (Brau-Antony *et al.*, 2005). Or, la note centre l'attention de l'élève sur celle-ci au détriment des appréciations plus qualitatives destinées à l'aider et l'accompagner dans son parcours d'apprentissage. La finalité du dispositif est ainsi modifiée. Pour Perrenoud (1998, p. 14), « rares sont les enseignants qui s'opposent résolument et ouvertement à une pédagogie différenciée et à une évaluation formative. Ils n'y adhèrent toutefois qu'à condition qu'elle soit donnée par-dessus le marché sans compromettre aucune des fonctions traditionnelles de l'évaluation ». L'évaluation formative serait un plus destiné à l'apprentissage mais, malgré ses effets bénéfiques, elle ne serait que facultative. Selon Altet (2001, p. 78), « les enseignants ayant le souci de modifier leurs pratiques d'évaluation et de mettre en œuvre une démarche d'évaluation formative, appelée aussi évaluation-régulation, rencontrent des difficultés certaines dans la mise en application de cette évaluation formative ». Ils ne parviennent pas à intégrer les éléments nécessaires dans leur organisation et gestion pédagogique habituelle.

Nos résultats semblent aller dans ce sens, car la quasi-exclusivité des enseignants affirment utiliser l'évaluation formative et pense qu'elle est utile à l'apprentissage. Or, en croisant les résultats, nous avons constaté que seuls 11% d'entre eux la proposent à chaque cycle d'apprentissage sans jamais la noter ce qui dévoile la grande variété de modalités mises en œuvre. Quand on s'aperçoit aussi, au regard de nos résultats, que l'expérience de l'enseignant (les plus de 50 ans et plus de 11 ans de carrière) est significativement dépendante d'une réflexion autour de la notation, nous pouvons de surcroît nous interroger sur l'impact de la formation initiale. L'expérience dévoile une reconnaissance du stress engendré par les notes, entraînant une réduction de leur fréquence, une plus grande mise en place d'évaluations formatives (non notées). Elle paraît être un critère important marquant l'évolution des représentations. La formation initiale aurait-elle moins de poids que la réflexivité issue de l'expérience pour faire évoluer les pratiques ? Selon Fabre (2006, p. 134), « l'analyse des pratiques – quel qu'en soit le modèle – commence lorsqu'on renonce à opposer d'autres solutions à la solution choisie par l'enseignant et qu'on tente de remonter avec lui au problème qui la fonde. » Dans un but de construction des actions de formation, l'objectif serait de renforcer les pratiques réflexives et les analyses de situations pour déceler ce qui fonde l'interprétation et la redéfinition pour chacun de l'évaluation.

Effectivement, la note qui sanctionne une évaluation formative contribue ainsi à la détourner de son utilité vis-à-vis des apprentissages et à l'assimiler à une étape intermédiaire à l'évaluation sommative. Cette pratique fréquente de la notation soulève deux principaux problèmes (Butera *et al.*, 2011) : le premier est que la note mesure une performance et non un apprentissage ; le second est relatif aux biais qui mettent largement en cause sa fiabilité et qui suscitent des sentiments néfastes chez les élèves (Merle, 2007 ; Vantourout et Goasdoue, 2010 ; Monteil et Huguet, 1999).

Dans ce cas, pourquoi nombre d'enseignants notent-ils les évaluations formatives ? Morissette (2010) avance l'idée selon laquelle les pratiques d'évaluation formative prennent la forme de processus de co-construction de sens dans l'interaction avec les élèves ; elles relèvent d'un système large impliquant la culture professionnelle des enseignants ainsi que leur milieu social ambiant. L'apparente non-prise en compte des pratiques jugées efficaces et validées scientifiquement doit se comprendre dans leur complexité pour permettre une démarche compréhensive, plutôt qu'une approche uniquement descendante visant à critiquer et

prescrire. Notre étude révèle que la note est largement perçue comme une source de motivation et d'engagement pour les élèves. Or, Butera *et al.* (*op. cit.*) a montré que si les élèves sont peut-être motivés par la note, elle ne suscite pas une envie de réussir mieux que les autres. Au contraire, elle induirait plutôt des buts connus pour limiter l'investissement et les performances scolaires. L'utilisation de la note comme source de motivation serait donc paradoxale. Les théories de la motivation vont dans ce sens (Deci et Ryan, 1985) en montrant que les sources de motivation externes agissent comme des contraintes posées au sujet et trouvent rapidement leurs limites. Pourtant, elles sont largement présentes à l'école mais cette abondance n'est évidemment pas à mettre sous la seule responsabilité des enseignants car l'institution aussi préconise ce recours comme moyen de communication avec les parents. De fait, il semble que les enseignants cherchant à pratiquer davantage l'évaluation formative soient contrariés par cette attente. Coburn (2004) considère que les enseignants sont plus enclins à tenir compte des nouvelles connaissances ou expériences si celles-ci sont compatibles avec leur système de croyances et de connaissances. Le recours à l'évaluation formative représenterait une rupture franche avec des conceptions bien ancrées et difficilement modulables.

Enfin, les résultats de l'enquête ont aussi révélé des relations entre des conceptions de l'évaluation et des caractéristiques contextuelles : par exemple les enseignants de lycée professionnel paraissent moins convaincus de l'utilité de l'évaluation formative que les autres enseignants. Nos résultats peuvent être mis en relation avec des recherches effectuées sur ce type d'établissements. En effet, généralement issus de milieux populaires, ces élèves sont souvent orientés par l'échec (Charlot, 1999 ; Jellab, 2005). La dévalorisation de l'enseignement professionnel et les politiques d'orientation successives ont favorisé la rupture des élèves avec l'image de leur formation : avant les élèves y entraient pour devenir ouvriers, et maintenant ils y entrent car ils sont en échec au collège. La perte de sens pour ces élèves peut en partie expliquer dans notre étude les différences constatées. Ces constats d'ailleurs se rapprochent de ceux pouvant être établis dans les classes de SEGPA qui regroupent des élèves en marge du système scolaire dès la classe de 6^{ème}. Les enseignants intervenant dans ces classes pensent l'évaluation davantage comme une source de motivation non stressante, n'hésitant pas ainsi à classer leurs élèves par leurs résultats. Mais ces pratiques mises en place sous prétexte de servir l'intérêt des élèves ne peuvent-elles pas mener, à long terme, à une plus forte rupture de ceux-ci avec le système scolaire ?

CONCLUSION

La présente enquête spécifique à l'EPS semble corroborer les observations empiriques effectuées dans la réalité ainsi que les différents constats établis par certains théoriciens à propos de l'activité évaluative des enseignants. Effectivement, les données témoignent d'interprétations différentes des formes d'évaluation proposées à l'école ainsi que leur adaptation aux contextes singuliers d'enseignement.

Cette étude a tenté, sur une population ciblée, de rendre compte de l'écart existant entre les discours et les conceptions des enseignants propres à l'évaluation des élèves, tout en précisant que ces résultats restent spécifiques à l'EPS, ce qui empêche toute généralisation. Nous avons constaté une interprétation de la fonction formative de l'évaluation par une partie des enseignants interrogés. Son intérêt pour l'apprentissage est reconnu mais le souci notamment d'optimiser l'investissement des élèves se traduit régulièrement par une notation de l'évaluation for-

mative, changeant ainsi sa finalité. À travers la définition que les enseignants interrogés donnent à l'évaluation et l'utilité attribuée à l'évaluation formative, les apports des formations et de l'expertise semblent réels. Cependant, ces éléments demanderaient à être approfondis et exploités différemment notamment selon le degré d'expertise de l'enseignant. Enfin, cette enquête a permis aussi de s'interroger sur les multiples effets de l'évaluation auprès des élèves (engagement, stress...), car ces effets ne sont pas perçus de la même manière par tous⁶.

La variété des pratiques fait appel à des représentations différentes de la définition, des fonctions et des effets de l'évaluation. Les représentations fonctionnent comme un filtre interprétatif où l'objet et le sujet sont interdépendants (Abric, 1987) ; elles permettent ainsi au sujet de donner du sens à ses conduites. De fait, les différences constatées sont le témoin de la liberté du sujet d'interpréter les connaissances qu'il a pu acquérir, de construire ses représentations au fil de son expérience et selon l'environnement dans lequel il agit. Rayou et Zanten (2004) ont d'ailleurs montré que, pour les jeunes enseignants, faire classe s'avère être un travail d'adaptation complexe au contexte dont le public est un élément central. Le manque de professionnalisation de leur formation, soulevée par ces jeunes professionnels, est un véritable problème pour apprendre à gérer la diversité. Notre recherche emboîte le pas à un tel constat, en témoignant d'une adaptation des enseignants à leur contexte d'enseignement. L'analyse réflexive de sa propre pratique devrait bénéficier d'une place de choix dans les formations initiale et continue pour favoriser cette adaptation. Car, tel que souligné par Fabre (2006), l'activité professionnelle n'est, de toute façon, jamais le reflet de la prescription.

Les contenus de formation en Master « Enseignement » s'appuient directement sur le référentiel de compétences des enseignants (Bulletin Officiel du 25 juillet 2013) à l'intérieur duquel figure clairement l'évaluation des progrès et des acquisitions des élèves mais l'articulation entre les stages et les apports théoriques n'est pas évidente à mettre en place. Le fonctionnement par alternance dans ces masters est un atout, en confrontant les futurs enseignants à la complexité du terrain tout en les accompagnant parallèlement pour répondre aux nombreux problèmes émergents de leur pratique. Les problématiques abordées en formation prennent tout leur sens pour les étudiants, encore faut-il que les formateurs qui accueillent les étudiants soient eux-mêmes formés à ce rôle complexe. C'est l'un des défis futurs à relever pour les Écoles Supérieures du Professorat et de l'éducation.

Lucie MOUGENOT
Université de Picardie
ÉSPÉ d'Amiens
CAREF EA 4697

Éric DUGAS
Université Bordeaux
ÉSPÉ d'Aquitaine
LACES EA 4140

⁶ Cette prise en compte de l'impact de l'évaluation sur les conduites des élèves *in vivo* fait en ce moment l'objet d'une recherche. Elle s'articule autour d'une analyse par observation différée par vidéo des conduites des élèves dans des situations avec ou sans évaluation.

Abstract : Evaluation is a central theme disseminated in teachers' everyday. Theoretical advances widely disseminated and being subject of many publications are not transposed as such at school. We asked 179 high school teachers in order to identify their conceptions and evaluation practices using a questionnaire. Finally, our results show a wide variety of practices, and especially the persistence of traditional and summative evaluation: formative evaluation is regularly deflected from its original purpose by being associated with a note, creating confusion in evaluation objectives.

Keywords : summative evaluation, formative evaluation, marking, teaching practice, teacher formation.

Bibliographie

- Abric J.-C. (1987) *Coopération, compétition et représentations sociales*. Cousset : Delval.
- Allal L., Cardinet J. & Perrenoud P. (dir.) (1979) *L'évaluation formative dans un enseignement différencié*. Berne : Lang [rééd. 1999].
- Altet M. (2001) « Pratiques d'évaluation et communication en classe » – in : G. Figari et M. Achouche *L'activité évaluative réinterrogée* (78-83). Bruxelles : De Boeck.
- Bardin L. (1977) *L'analyse de contenu*. Paris : PUF.
- Becker H.S. (1963) *Outsiders. Études de sociologie de la déviance* (trad. J.-P. Briand et J.-M. Chapoulié 1985). Paris : Éditions A.-M. Métailié.
- Boraita F. & Issaieva E. (2013) « Les conceptions de l'évaluation chez les futurs enseignants et enseignants en exercice : une étude dans le canton de Genève » – *Actes du colloque Admee 2013*. Disponible sur internet : http://www.admee2013.ch/ADMEE-2013/7_files/Boraita-Issaieva-ADMEE-2013.pdf
- Brau-Antony S. & Cleziou J.-P. (2005) *L'évaluation en EPS: Concepts et contributions actuelles*. Paris: Actio.
- Braxmeyer N., Guillaume J.-C. & Levy J.-F. (2004) « Les pratiques d'évaluation des enseignants en collège » – *Note de la DEP* 04.13.
- Bressoux P., Bru M., Altet M. & Lambert C. (1999) « Diversité des pratiques d'enseignement à l'école élémentaire » – *RFP* 126 (97-110).
- Butera F., Buchs C. & Darnon C. (2011) *L'évaluation, une menace ?* Paris : PUF.
- Cardinet J. & Laveault D. (2001) « L'activité évaluative en éducation: évolutions des préoccupations des deux côtés de l'atlantique » – in : G. Figari et M. Achouche (éds.) *L'activité évaluative réinterrogée* (15-29). Bruxelles : De Boeck.
- Charlot B. (1999) *Le rapport au Savoir en milieu populaire. Une recherche dans les lycées professionnels de banlieue*. Paris : Anthropos.
- Coburn C. (2004) « Beyond Decoupling : Rethinking the Relationship between the Institutional Environment and the Classroom » – *Sociology of Education* 77 (211-244).
- David B. (dir.) (2000) *Équité et arrangements évaluatifs. Certifier en éducation physique et sportive*. Paris : INRP.
- Deci E.-L. & Ryan R.-M. (1985) *Intrinsic motivation and self-determination in human behavior*. New York: Plenum Press.
- De Ketele J.-M. (1988) *Guide du formateur*. Bruxelles : De Boeck.
- Delorme C. (dir.) (1987) *L'évaluation en question*. Paris : ESF.
- Fabre M. (2006) « Autour des mots. Analyse des pratiques et problématisation » – *Recherche et Formation* 51 (133-145).

- Figari G. (1994) *Évaluer : quel référentiel ?* Bruxelles: De Boeck-Wesmael.
- Hadji C. (1997) *L'évaluation démystifiée*. Paris : ESF.
- Jellab A. (2005) « Les enseignants de lycée professionnel et leurs pratiques pédagogiques : entre lutte contre l'échec scolaire et mobilisation des élèves » – *Revue Française de Sociologie* 46 (295-323).
- Leplat, J. (1986) « L'analyse psychologique du travail » – *Revue de Psychologie Appliquée* 31 (9-27).
- Marcel J.-F., Dupriez V., Perisset D. & Tardif M. (éds.) (2007) *Coordonner, collaborer, coopérer : de nouvelles pratiques enseignantes*. Bruxelles : De Boeck.
- Merle P. (2007) *Les notes : secrets de fabrication*. Paris : PUF.
- Monteil J.-M. & Huguët P. (1999) *Social context and cognitive performance: Towards a social psychology of cognition*. Hove : Psychology Press.
- Morissette J. (2010) « Une perspective interactionniste. Un autre point de vue sur l'évaluation des apprentissages » – *Revue de l'Association Internationale des Sociologues de Langue Française*. Disponible sur internet : <http://sociologies.revues.org/3028>
- Perrenoud P. (1998) *L'évaluation des élèves. De la fabrication de l'excellence à la régulation des apprentissages*. Bruxelles : De Boeck.
- Rayou P. (2002) *La « dissert de philo ». Sociologie d'une épreuve scolaire*. Rennes : PUR.
- Rayou P. & Van Zanten A. (2004) *Enquête sur les nouveaux enseignants. Changeront-ils l'école ?* Paris : Bayard.
- Scallon G. (1982) *L'évaluation formative*. Bruxelles : De Boeck.
- Talbot L. (2012 [2009]) *L'évaluation formative. Comment évaluer pour remédier aux difficultés d'apprentissage*. Paris : Armand Colin.
- Vantourout M. & Goasdoué R. (2010) « Correction de dissertations : analyse de l'activité de professeurs engagés dans une approche par compétences » – *Actes du congrès de l'Actualité de la Recherche en Éducation et Formation*, Université de Genève.

Annexe
Extraits du questionnaire proposé

En une phrase, pour vous, qu'est-ce que l'évaluation ?

Cochez un choix par ligne, reflétant votre avis relatif à l'affirmation donnée :

	Pas du tout d'accord	Peu d'accord	Assez d'accord	Tout à fait d'accord
J'évalue pour noter.				
L'évaluation formative est utile pour aider mes élèves à progresser.				
J'évalue pour maintenir l'engagement de mes élèves.				
J'évalue pour optimiser l'apprentissage.				
Evaluer mes élèves me permet de les classer.				
L'évaluation sommative motive mes élèves.				
Mes élèves ne différencient pas l'évaluation formative et sommative.				
L'absence de notation entraîne un désengagement de mes élèves.				
Les évaluations sommative et formative sont 2 phases du même travail, fondées sur les mêmes données mais utilisées à des fins différentes				
L'évaluation en général est source de stress pour les élèves, quels qu'ils soient.				

Tableau 5 : échelle d'attitude proposée aux enseignants.

Cette année, avez-vous utilisé la co-évaluation élèves ?

Oui non

Si oui, à quel moment utilisez-vous le plus la co-évaluation ?

Lors de l'évaluation diagnostique :

Lors de l'évaluation formative :

Lors de l'évaluation sommative :

Cette année, avez-vous réellement mis en place des évaluations diagnostiques ?

A chaque cycle

Parfois

Jamais

Cette année, avez-vous mis en place des évaluations formatives ?

Oui, à chaque cycle

Oui, parfois

Non jamais

Vous arrive-t-il d'utiliser une évaluation formative pour noter vos élèves ?

Jamais :

Oui parfois :

Oui souvent :

Oui, systématiquement

Si oui, pouvez-vous nous expliquer en quelques mots pourquoi ?

Pensez-vous que les élèves puissent adopter des conduites motrices (dans l'action, c'est-à-dire hors conduites sociales) différentes entre, d'une part, les séances d'apprentissage hors évaluation, et d'autre part :

les situations d'évaluation formative: Oui non
 les situations d'évaluation sommative : Oui non

Vos élèves ont connaissance des critères d'évaluation sommative :

Dès la première séance
 Au dernier moment
 Au cours du cycle
 Je ne fais pas d'évaluation sommative

Pensez-vous que les critères d'évaluation sommative soient vraiment assimilés par les élèves ?

Oui Non Parfois

Si vous avez répondu parfois, pouvez vous nous expliquer pourquoi, en quelques mots :

Exemples de tableaux de contingence :

<i>Avez-vous mis en place des évaluations formatives ?</i>		
<i>n = 179</i>	Effectif	Pourcentage
A chaque cycle	76	42%
Parfois	96	54%
Jamais	5	3%
Non-répondants	2	1%
Ensemble	179	100%

Tableau 6 : la mise en place des évaluations formatives.

<i>Vous arrive-t-il d'utiliser une évaluation formative pour noter vos élèves ?</i>		
<i>n = 179</i>	Effectif	Pourcentage
Jamais	67	37%
Parfois	85	47%
Souvent	17	9%
Systématiquement	9	5%
Non-répondants	1	1%
Ensemble	179	100%

Tableau 7 : la notation des évaluations formatives.

<i>Vos élèves ont connaissance des critères d'évaluation sommative :</i>		
<i>n = 179</i>	Effectif	Pourcentage
Dès la première séance	65	36%
Au dernier moment	3	2%
Au cours du cycle	111	62%
Total	179	100%

Tableau 8 : connaissance des critères d'évaluation par les élèves, selon leurs enseignants.

<i>Pensez-vous que les critères d'évaluation soient vraiment assimilés par les élèves ?</i>		
<i>n = 179</i>	Effectif	Pourcentage
Oui	65	36%

*FORMATION ET PRATIQUE DES ENSEIGNANTS
REGARD PORTÉ SUR L'ÉVALUATION EN EPS*

Non	17	9%
Parfois	95	56%
Non-répondants	2	1%
Ensemble	179	100%

Tableau 9 : assimilation des critères par les élèves, selon leurs enseignants.