



HAL
open science

Chapitre 4. L'évaluation scolaire en France : la spécificité de l'EPS

Lucie Mougenot

► **To cite this version:**

Lucie Mougenot. Chapitre 4. L'évaluation scolaire en France : la spécificité de l'EPS. Évaluation et autoévaluation. Quels espaces de formation?, De Boeck Supérieur, pp.157-174, 2015, Pédagogies en développement, 10.3917/dbu.ceon.2015.02.0157 . hal-04244712

HAL Id: hal-04244712

<https://hal.science/hal-04244712>

Submitted on 16 Oct 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

L'évaluation scolaire en France : la spécificité de l'EPS

Lucie Mougnot

1. Introduction

L'évaluation occupe une place très importante dans les pratiques scolaires : au quotidien, selon Braxmeyer et *al.* (2004), l'activité évaluative représente 18 à 20 % du temps de travail des enseignants. Elle est une préoccupation majeure pour ces acteurs de l'enseignement, mais aussi pour les élèves, l'institution et les parents, puisque ses résultats aboutissent à une prise de décision (Hadjji, 1997). Toutefois, l'action d'évaluer est souvent problématique pour l'enseignant : elle donne lieu à un foisonnement de questions, de compromis à réaliser, de précautions à prendre, d'autant plus si l'on s'inscrit dans une conception éthique de l'évaluation. L'élève qui échoue est-il celui qui n'a pas appris ? Les évaluations proposées reflètent-elles réellement et uniquement les apprentissages réalisés en amont ? Peuvent-elles être appréhendées autrement que comme une sanction par les élèves ? L'activité évaluative des enseignants est complexe car elle se situe au carrefour de multiples dimensions que nous soulèverons par la suite. De fait, former les enseignants à évaluer relève d'un processus complexe et ne peut s'envisager uniquement par le biais de la formation initiale. Au regard d'une quarantaine de visites de terrain que nous avons réalisées auprès d'enseignants débutants, une des difficultés majeures semble être la concordance entre les attentes des jeunes enseignants et les résultats obtenus lors des évaluations. Certaines représentations et pratiques sont de véritables obstacles à une évaluation utile pour l'élève, construite de façon à l'aider à apprendre. L'évaluation chez les enseignants débutants paraît davantage orientée vers une fonction sociale (sommative voire certificative), plutôt que vers une fonction pédagogique : cela questionne d'emblée le rôle et les manques de la formation à l'évaluation. L'enjeu de ce chapitre est de pouvoir apprécier les conceptions des enseignants vis-à-vis de l'évaluation scolaire ; il s'agit de tenter de comprendre, par une démarche exploratoire, ce qui fonde leurs choix de modalités évaluatives. L'évaluation est un élément fondamental de l'enseignement et de l'apprentissage, un élément complexe du fait des formes différentes qu'il peut adopter et des enjeux qu'il sous-tend. Au-delà d'une pluralité de mises en œuvre observables, il convient de s'intéresser aux facteurs susceptibles d'orienter ces pratiques pédagogiques.

Les données présentées ici sont le résultat d'une enquête menée en Education Physique et Sportive (EPS) en France auprès de 179 professeurs. Nous avons choisi cette discipline

singulière car la formation est réalisée dans des établissements universitaires (les UFR STAPS) où les étudiants ont très tôt la possibilité de réaliser des stages en milieu scolaire. De plus, les enseignements scientifiques et la pratique des activités physiques sont la plupart du temps orientés vers l'analyse des pratiques, vers la compréhension des obstacles à l'apprentissage – que vivent les étudiants –, ce qui ouvre la voie assez tôt à une prise en compte de l'hétérogénéité des élèves. Cette discipline possède aussi des modalités d'évaluation spécifiques comme le contrôle en cours de formation lors des examens scolaires ; les enseignants évaluent ainsi directement la portée de leur enseignement et sont tenus d'être toujours informés des directives institutionnelles.

Cette recherche est une enquête sous forme de questionnaire dans laquelle les enseignants abordent les fonctions des évaluations et, de façon concrète, ce qu'ils mettent en place. Sans avoir pour ambition de comprendre la complexité de leurs pratiques ni de généraliser ou d'extrapoler les résultats, cette démarche exploratoire a pour objet d'appréhender les différences de mise en œuvre qui traduisent des interprétations multiformes d'une culture commune de l'évaluation. L'enquête peut ainsi être un support aux formations : elle témoigne des difficultés rencontrées pour mettre en pratique les connaissances *a priori* transmises en formation initiale. En effet, sous des intentions pertinentes au départ, les choix pédagogiques sont divergents et régulièrement en décalage avec les finalités poursuivies. Or, en se référant au philosophe James (1978, p. 124), « pour développer la signification d'une pensée, il nous faut seulement déterminer la conduite qu'elle est apte à produire ; cette conduite est pour nous sa seule signification et sa seule importance ». Sous de mêmes termes, les pratiques peuvent être différentes selon les individus, tout comme différents concepts peuvent aboutir à des propositions similaires.

2. Des conceptions protéiformes

Le choix des modalités évaluatives est lié aux priorités pédagogiques et didactiques des enseignants qui disposent d'une grande liberté (Bressoux et *al*, 1999). Les pratiques sont encadrées par des recommandations officielles mais l'espace de décision reste vaste et soumis à la complexité de la situation d'enseignement. L'apport de la formation initiale ne permet pas toujours de répondre aux nombreuses attentes des enseignants, puisque la singularité des situations vécues demande une adaptation constante (Baillat, 2009), et nécessite des allers-retours permanents entre théorie et pratique. Les enseignants choisissent au regard de leur vécu, de leurs connaissances, du contexte et des contraintes institutionnelles, ce qui semble le

plus approprié pour leurs élèves. Bien que la standardisation des pratiques ne soit pas recherchée, la variété des conceptions et des mises en œuvre mérite d'être analysée pour comprendre en quoi la culture de l'évaluation est partagée, et dans quelle mesure les ressources théoriques sont redéfinies, interprétées par chacun ; cela engendre une complexité et une pertinence de l'analyse des pratiques.

Si les pratiques sont disparates, cela peut être consécutif à des priorités individuelles spécifiques au regard des multiples possibilités qui sont offertes, ou alors, cela peut aussi être la résultante de choix par défaut. Dans ce deuxième cas, le manque de formation est fondamental mais pas suffisant, car la relation entre les théories et la pratique dépasse la seule connaissance des méthodes ou dispositifs révélés efficaces par la recherche. D'ailleurs – et notamment depuis la création des IUFM en 1989 –, la formation des enseignants est régulièrement réformée, tiraillée entre la forte demande de professionnalisation et l'apport indispensable de ressources théoriques. Si la théorie et la pratique apparaissent interdépendantes pour assurer une formation pertinente, les liens entre ces deux entités sont complexes (Crahay, 2007). En effet, les conceptions des enseignants sont la résultante d'influences multiples que la formation à elle seule ne peut bousculer. Par exemple, Morissette (2010) relève le fait que les enseignantes du primaire affirment employer fréquemment des méthodes traditionnelles car celles-ci sont valorisées par les différents acteurs qui gravitent autour de la sphère éducative. De plus, Coburn (2004) souligne que les enseignants sont plus enclins à tenir compte des nouvelles connaissances ou expériences si celles-ci sont compatibles avec leur propre système de croyances. Les résistances seraient en grande partie dues à une idéologie dominante dans laquelle les enseignants s'attachent au principe de justice méritocratique qui renforce l'adhésion à des conceptions traditionnelles de l'évaluation (Crahay, 2010). La question de l'évolution des pratiques et des conceptions touche à celle de l'identité professionnelle et plus largement à l'identité de la personne.

Aujourd'hui en France, les formations initiales en Master « Enseignement » et les formations continues académiques¹ en didactique disciplinaire abordent l'enseignement par compétences dans lequel une place est réservée aux différents modes d'évaluation (notamment formative²).

¹ Par exemple, dans l'académie d'Amiens en EPS, plusieurs formations ont pour intitulé : « Construire des formes de pratiques scolaires des compétences propres 2 et 3. Concevoir une évaluation progressive, formative (...) ».

² Talbot (2009) dans « L'évaluation formative : comment évaluer pour remédier aux difficultés d'apprentissage » définit l'évaluation formative comme un ensemble de procédures plus ou moins formalisées visant à adapter l'action pédagogique et didactique en fonction des progrès et des difficultés d'apprentissage.

Néanmoins, ces formations impliquent peu d'allers-retours réguliers entre la théorie et la pratique, et l'on peut se demander si elles sont suffisantes pour induire des mises en œuvre en rupture avec une évaluation-sanction, une évaluation discriminante.

À partir d'une analyse des propositions pédagogiques relatives aux évaluations, l'objectif est de cerner les différences fondamentales qui émergent en partie selon deux variables : l'ancienneté et le grade. Ce choix permettra de mettre en lumière l'importance de l'expérience et d'une formation plus poussée (l'agrégation), pour faire évoluer les pratiques selon une perspective éthique, qui rompt ainsi avec la logique normative. De ce fait, l'activité évaluative de l'enseignant sous-tend une certaine conception de l'apprentissage, une vision singulière de l'élève.

3. L'évaluation en EPS

Même si l'évaluation scolaire est un sujet transversal aux disciplines, chacune d'entre elles connaît des problèmes spécifiques de mise en œuvre. Comme le souligne David (2000), la mise en œuvre des évaluations en EPS est problématique : il cite par exemple le biais d'enseigner au final que ce que l'on peut évaluer, pour espérer mettre en œuvre des procédures d'évaluation précises et objectives. Par ailleurs, à la grande différence de ses collègues, l'enseignant d'EPS évalue des conduites motrices³. Cette spécificité implique des difficultés supplémentaires comme la nécessité d'observer chaque élève pendant l'action – bien souvent collective –, de saisir les intentions des élèves, sans prise de recul ni retour en arrière possible. Le résultat est immédiat, tant pour l'élève qui ne peut pas revenir sur son action, que pour l'enseignant qui est contraint de traduire dans l'immédiat ses observations, selon les critères d'évaluation définis. La prise de décision est donc très rapide et demande une grande maîtrise des critères d'observation.

Concernant les démarches évaluatives, Brau-Antony (2001) a réalisé une recherche auprès de 275 enseignants d'EPS à propos de leurs conceptions de l'enseignement des jeux collectifs, par le biais d'un questionnaire. Il montre que les enseignants ont tendance à faire cohabiter les fonctions – formative et certificative – de l'évaluation. Cela nous a incités à relever les différences que les enseignants établissent entre ces deux types d'évaluations qui ont des

³ Parlebas aborde la notion de conduites motrices dès 1968. Elle se définit comme l'organisation signifiante du comportement moteur.

finalités bien différenciées. En définitive, même s'ils affirment les distinguer, les mises en œuvre proposées parfois mettent en lumière une certaine confusion : par exemple, lors de visites de terrain, nous avons pu constater que les évaluations formatives sont parfois notées, ce qui relègue au second plan leur intérêt pédagogique. Le rapport à la notation est un élément central dans la recherche que nous présentons ici, car il modifie les buts pour l'élève (Butera *et al.*, 2006).

Dans l'enseignement, les décisions prises semble émaner d'habitudes voire d'intuitions. L'expérience et le vécu de chacun orientent les choix mais, malgré la tendance à penser que ceux-ci sont justifiés et pertinents, les conséquences des modalités évaluatives proposées ne reflètent pas toujours l'effet escompté. L'enjeu est de taille, puisque ces conséquences impactent directement l'apprentissage.

L'étude présentée dans ce chapitre vise à définir le sens que les enseignants accordent à l'évaluation, à mesurer les écarts interindividuels et à les confronter aux pratiques jugées efficaces par les experts. Les modalités évaluatives mises en œuvre seront discriminées selon l'ancienneté des enseignants, leur grade, et leurs conceptions à l'égard des effets de l'évaluation sur la motivation et le stress de leurs élèves. Cela permettra indirectement de saisir la portée de la formation à l'évaluation sur les choix pédagogiques.

4. Méthodologie

L'observation et la description de ce que les enseignants proposent auraient pu permettre d'analyser les pratiques *in vivo*, mais nous nous sommes centrés davantage sur leurs conceptions, leurs intentions à partir desquelles les choix apparaissent. La démarche ici est plutôt exploratoire, c'est à dire que nous avons opté pour le questionnaire qui offre la possibilité de recueillir un nombre important de réponses. Même s'il ne permet pas d'appréhender directement les pratiques, mais plutôt ce que les enseignants en disent, les résultats seront utiles pour cerner la spécificité de l'évaluation en EPS et appréhender les choix évaluatifs dans une certaine mesure.

4.1. Les enseignants interrogés

Trois moyens ont été utilisés pour diffuser les questionnaires : la distribution de main à main ; la voie hiérarchique, en sollicitant quelques Inspecteurs Pédagogiques Régionaux ; et de façon plus aléatoire la voie électronique, auprès plusieurs centaines d'établissements scolaires. Nous

avons récolté 179 questionnaires exploitables.

La plupart des enseignants interrogés intervient dans plusieurs types d'établissements : 56% ($n = 100$) exercent en collège et 24 d'entre eux ont au moins une classe de SEGPA⁴ ; 39% ($n = 69$) sont en lycée général, 23% ($n = 42$) en lycée technologique et 32% ($n = 57$) en lycée professionnel (LP). La grande majorité des enseignants de lycée exerce sur plusieurs types d'établissements, essentiellement en lycée polyvalent⁵.

La répartition selon le genre et l'âge des enseignants interrogés (voir figure 1) correspond globalement à celle de l'ensemble des enseignants d'EPS en France en 2009⁶.

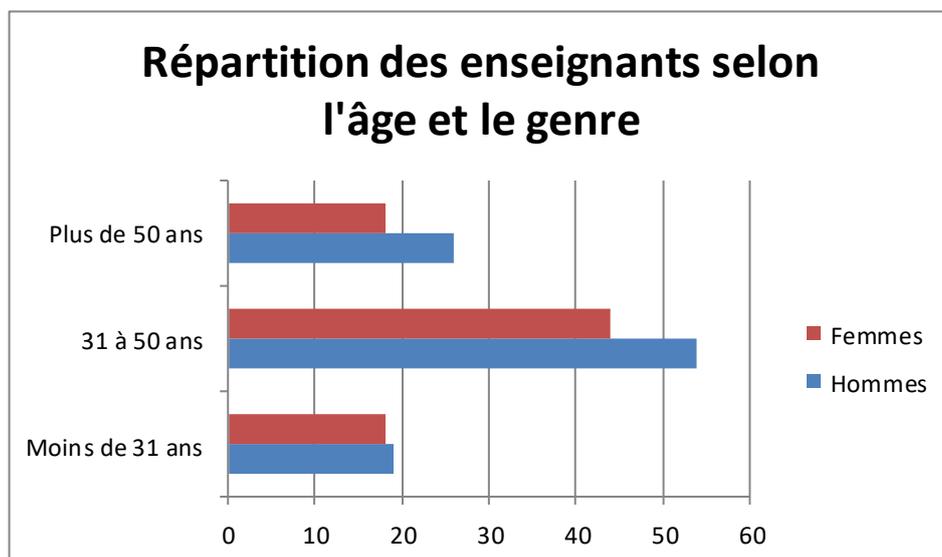


Figure 1 : répartition en nombre des enseignants, selon leur genre et leur âge.

Nous avons interrogé 55% d'hommes sans observer d'effet d'interaction entre le genre et le type d'établissement. Concernant le niveau de qualification, 78% de certifiés ont été interrogés ($n = 138$), 15% d'agrégés ($n = 28$) et 7% ayant un statut différent ($n = 13$) : stagiaires, vacataires, bi-admissibles ou PEGC⁷. Nous n'observons pas de dépendance significative entre le grade et le type d'établissement d'affectation.

⁴ Section d'Enseignement Général et Professionnel Adapté

⁵ Parmi les 69 enseignants intervenant en lycée général, trente-six ont aussi des classes technologiques. 34 enseignants ont des classes des trois sections en lycée.

⁶ DEPP, L'état de l'école. 29 indicateurs sur le système éducatif français. DEPP 2010, n°20.

⁷ Professeur d'Enseignement Général en Collège. Dans cette étude, nous ne relèverons pas de façon spécifique les résultats concernant cette troisième catégorie d'enseignants car ils ne sont pas assez nombreux.

Les données signalétiques de cette enquête dressent un portrait assez réaliste de la corporation en 2010 même si une proportion un peu plus grande d'enseignants de LP ou étant agrégés a été interrogée par rapport aux statistiques ministérielles.

42. Informations recueillies et choix des variables

A la suite d'entretiens réalisés auprès de trois enseignants en EPS, différents axes de questionnement ont été définis. Ces entretiens ont fait émerger des problématiques essentielles, notamment le rapport étroit entre la notation et l'évaluation. La définition des axes a abouti à celle de données à recueillir, comme présenté dans le tableau 1 ci-dessous :

Thèmes de recherche	Données à recueillir
Mises en œuvre des évaluations.	Définition de l'évaluation. Pratiques relatives aux différentes formes : diagnostique, formative, sommative.
La notation.	Fréquence de la notation. Rapport entre les formes d'évaluation et la notation.
Les fonctions de l'évaluation.	Utilité selon ses différentes formes pour l'enseignant et les élèves.
L'impact de l'évaluation sur les élèves.	Impact des évaluations sur les conduites motrices, le stress des élèves et leur engagement.

Tableau 1 : thèmes abordés dans le questionnaire.

Les différents axes présentés ci-dessus sont en étroite relation. Notons qu'au cours du questionnaire, la définition des types d'évaluation énoncés a été précisée en bas de page. Malgré ces précisions, des divergences apparaissent, ce qui renvoie directement à l'objet de cette enquête : montrer que, sous des termes *a priori* partagés, les enseignants y assimilent des définitions et mises en œuvre différentes.

Le questionnaire est anonyme ; les données signalétiques demandées à la fin concernent le sexe, l'âge (30 et moins ; 31 à 50 ; plus de 50 ans), l'ancienneté, le grade et le type d'établissement d'affectation.

43. Structure du questionnaire

Une fois la trame du questionnaire définie, celui-ci a été présenté au sein d'un séminaire du laboratoire du GEPECS⁸ où chaque question a été discutée, voire modifiée. Ensuite, il a été pré-testé auprès de huit enseignants, en présence du chercheur, pour effectuer les derniers ajustements.

La passation dure 15 à 20 minutes. Quinze questions sont posées : la plupart du temps, ce sont des questions fermées, parfois accompagnées d'une demande de complément en question ouverte. Pour cibler les conceptions des enseignants, nous avons aussi proposé une échelle d'attitude à quatre niveaux.

44. Traitement des données

Les données recueillies ont été traitées avec le logiciel *Question DATA 6*. Il permet de réaliser des tris croisés à l'aide de tableaux de contingence. Ceux-ci ont pour fonction de tester statistiquement l'existence ou non de relation entre les variables, par le test du *khi 2*.

5. Les résultats

Différents points méritent une attention particulière car ils font l'objet d'un consensus quasi exhaustif. Relever ces points d'accord, c'est aussi mettre en avant des divergences qui seront traitées ensuite.

51. Convergence de points de vue : les évaluations sommatives et diagnostiques

Tout d'abord, tous les enseignants interrogés mettent en place des évaluations pour noter leurs élèves, au moins une fois par cycle d'enseignement. Les enseignants utilisent la notation pour retranscrire les résultats des évaluations et informer les parents et l'institution du niveau des élèves. Ceci fait référence à une fonction sociale.

Ensuite, tous les enseignants affirment informer les élèves des critères d'évaluation : pour la plupart (61%) au cours du cycle et, pour les autres, dès la première séance. Pourtant, au regard des résultats d'une enquête de Lentillon (2008), on apprend que les élèves sont nombreux à ne pas savoir comment ils sont évalués, malgré les informations transmises par les enseignants. Il apparaît ainsi un écart entre les conséquences attendues et réelles des actions entreprises.

⁸ Groupe d'Etude Pour l'Europe de la Culture et de la Solidarité.

Enfin, la mise en place d'une évaluation diagnostique est une pratique aussi très répandue : 41% des enseignants l'utilisent systématiquement et 57% parfois. Ces données rejoignent celles de Braxmeyer et *al.* (2004) qui ont montré la forte mise en œuvre de ce dispositif en EPS, comparé aux autres disciplines.

En dehors de ces quelques points d'accord, nous allons à présent insister sur les différences de mises en œuvre et de conceptions.

52. Une centration sur la finalité sociale de l'évaluation

La finalité sociale de l'évaluation est majoritairement mise en avant. Ainsi, 124 enseignants (70%) sont d'accord pour dire qu'ils évaluent pour noter (voir figure 2 ci-dessous). La notation et le caractère sommatif de l'évaluation sont très présents dans les conceptions. Sans spécifier dans la question le type d'évaluation (qui aurait pu être diagnostique, formatif ou sommatif), plus des deux tiers des enseignants y rattachent spontanément la notation.

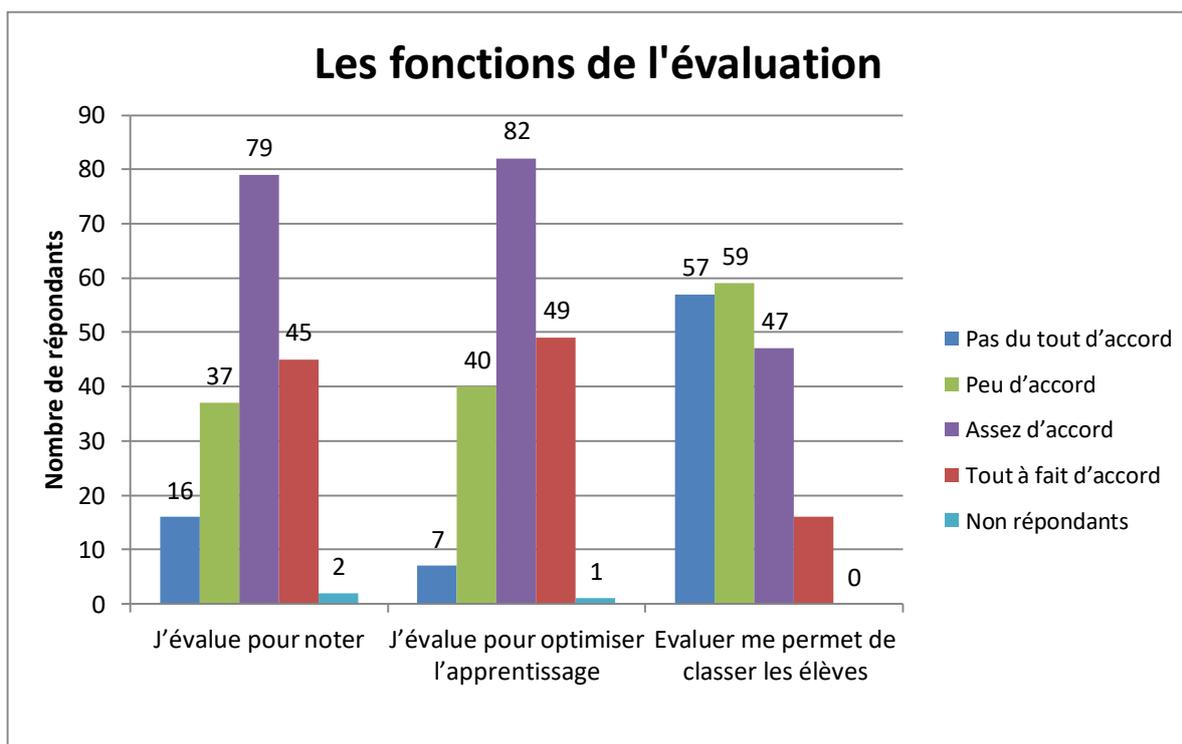


Figure 2 : répartition des enseignants approuvant ou non les propositions relatives aux fonctions de l'évaluation. Résultats présentés en nombre ($n = 179$).

D'autre part, la fonction formative – au service de l'apprentissage – de l'évaluation est aussi largement approuvée, par 131 enseignants. En croisant les données, on observe une surreprésentation des enseignants tout à fait d'accord avec les deux propositions ($p = 0.001$) :

ceci pourrait paraître contradictoire, puisque optimiser l'apprentissage requiert des dispositifs plutôt formatifs, alors que la notation est davantage dénoncée car située à l'encontre des finalités d'apprentissage. Seulement, ces enseignants n'envisagent pas les effets néfastes de la notation : une analyse factorielle des données par ANOVA met en avant le fait qu'il existe un effet d'interaction entre « j'évalue pour noter » et « l'évaluation sommative motive mes élèves » ($p = 0,0001$). Le recours à la notation est fréquent, même en cas d'évaluation dite « formative », car la note est régulièrement associée à une source de motivation, comme nous le verrons plus loin. L'objectif annoncé est bien d'aider les élèves au travers de dispositifs formatifs, mais le fait que l'ajout d'une note modifie l'effet escompté semble minoré.

Une autre fonction apparaît dans l'échelle d'attitude : 35% des enseignants disent utiliser l'évaluation pour comparer les élèves entre eux, en les classant. En ce qui concerne les enseignants n'étant pas du tout d'accord avec le fait d'évaluer pour classer ($n = 57$ soit 32%) une première distinction apparaît selon le grade : les certifiés sont sous-représentés ($p = 0,01$) et les agrégés surreprésentés ($p = 0,05$).

Les différences selon le grade à propos de la fonction sociale de l'évaluation apparaissent aussi dans l'étude de la fréquence de notation : les agrégés notent moins souvent que les certifiés, respectivement et en moyenne 2.1 fois contre 2,7 fois par séquence, et les agrégés sont surreprésentés parmi ceux qui ne notent qu'une seule fois par séquence ($p = 0.05$). À l'inverse, les certifiés et non-titulaires sont sous-représentés parmi ceux qui ne notent qu'une fois ($p = 0.05$). Par ailleurs, selon l'ancienneté, 74% des plus de 50 ans notent une ou deux fois maximum leurs élèves. L'expérience et la formation paraissent ainsi influencer les choix pédagogiques et notamment l'usage de la notation.

53. L'évaluation formative en EPS

La quasi-totalité des enseignants affirme mettre en place des évaluations formatives de façon récurrente ou ponctuelle (97%). Selon la fréquence de son utilisation, notons que 43% en proposent à chaque cycle d'apprentissage et, parmi eux, les plus de 50 ans sont surreprésentés ($p = 0,01$). Les enseignants ayant le plus d'expérience notent moins souvent leurs élèves, et mettent en place de façon plus systématique des évaluations formatives dans un but d'optimisation de l'apprentissage.

Les résultats issus de tris croisés apportent un contraste à ce fort engouement apparent pour les évaluations formatives. Déjà, 38% affirment ne jamais les noter ; bon nombre d'entre eux les notent parfois et 15% souvent ou systématiquement (voire même, pour une minorité,

jusqu'à une note par séance). Ensuite, on s'aperçoit que les enseignants qui notent souvent ou systématiquement les évaluations formatives font partie de ceux qui disent en mettre en place à chaque cycle ($p = 0.00011$ et $n = 20$) ; et ce sont les enseignants en début de carrière qui notent le plus souvent. Finalement, seulement 12 % des enseignants interrogés proposent au moins une évaluation formative à chaque cycle sans jamais la noter. Sous des intentions semblables au départ, nous observons de grandes divergences dans les pratiques : au sortir de la formation initiale, le recours à la notation est fort et induit une reconfiguration de l'évaluation dite formative. Nous allons à présent confronter ces résultats aux représentations des enseignants vis-à-vis de la motivation et du stress de leurs élèves, ce qui permet de comprendre en partie la présence forte la notation.

54. Évaluation, motivation et stress des élèves

Quand on demande aux enseignants s'ils évaluent pour maintenir l'engagement des élèves, les avis sont partagés : 42% sont assez d'accord et 26% peu d'accord. Certaines précisions peuvent toutefois être apportées : tout d'abord les certifiés sont sous-représentés parmi ceux étant peu d'accord avec le fait que l'évaluation maintienne l'engagement ($p = 0,05$) ; ils auraient ainsi plus tendance à associer l'évaluation à une sollicitation de la motivation et de l'engagement des élèves : d'ailleurs, 21% d'entre eux ($p = 0.1$) sont tout à fait d'accord pour dire qu'ils évaluent dans le but de maintenir l'engagement, contre aucun agrégé.

Les enseignants sont plutôt partagés vis-à-vis de l'effet des évaluations sur le stress des élèves. Cependant, on remarque quelques dépendances très significatives : parmi ceux étant peu d'accord, les enseignants de 30 ans et moins sont surreprésentés ($p = 0.01$). En revanche, nous retrouvons, parmi ceux étant plutôt d'accord, les enseignants qui ont plus de 11 ans de carrière ($p = 0.01$). De plus, le stress engendré paraît pris en compte par une diminution du nombre de notes : les résultats témoignent d'une augmentation du nombre de notes allant de pair avec une non-prise en compte du stress (voir tableau 3 ci-dessous).

L'évaluation est source de stress pour les élèves				
	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 ($n =$

179).

Ce tableau 3 est explicite : plus les enseignants pensent que l'évaluation est source de stress, moins ils mettent de notes en moyenne. L'écart est progressif entre « tout à fait d'accord » et « peu d'accord » mais accélère brutalement pour les enseignants « pas du tout d'accord ». À ce stade, les avis paraissent bien plus tranchés et aboutissent à des pratiques encore plus accentuées vers la notation.

Ainsi, l'ensemble des données recueillies met en lumière de fortes divergences de conceptions qui se traduisent dans les pratiques évaluatives, selon l'expérience de l'enseignant. Former à l'évaluation est un véritable défi pour les formations initiales ; l'usage de la notation est perçu bien souvent par les jeunes enseignants comme un moyen de motivation, sans prise en compte de ses limites.

6. Discussion

6.1. La spécificité de l'évaluation en EPS

Sans avoir l'ambition de produire des savoirs, nous avons souhaité appréhender les interprétations faites des discours théoriques propres à l'évaluation en EPS. L'axe d'analyse a notamment permis de mettre en évidence des différences importantes entre les enseignants en début de carrière et ceux plus formés ou expérimentés.

L'évaluation est au service de l'apprentissage pour certains enseignants d'EPS, mais elle reste un moyen – d'autant plus chez les enseignants débutants – de gérer la classe, d'inciter les élèves à s'investir. Les conceptions de l'évaluation sont protéiformes : évaluer peut satisfaire à différentes fonctions d'ordre pédagogique et/ou du didactique (Allal et *al.*, 2007), mais cela peut servir aussi à maintenir la tranquillité dans la classe et à préserver un climat plutôt calme (Bressoux et *al.*, 2003), ou encore à hiérarchiser les élèves (Grisay, 1988). Bien que notre étude ait mis en évidence ces différentes fonctions – en particulier par le recours à la notation –, un des résultats importants réside dans l'attention portée par les enseignants à l'apprentissage. D'ailleurs, l'étude de Braxemeyer et *al.* (*op. cit.*) a montré que, même si l'évaluation sommative domine encore dans les pratiques, des évolutions vers des pratiques formatives sont constatées (particulièrement en EPS).

Les difficultés rencontrées par les enseignants en EPS les incitent à se questionner pour trouver des solutions pertinentes. Effectivement, ces derniers sont confrontés au fait de devoir

évaluer en temps réel toute une classe dans l'urgence, de ne pas pouvoir revenir en arrière, de prendre en compte les compétences individuelles dans des contextes souvent interactionnels (jeux collectifs, acrosport, etc), d'évaluer aussi les compétences sociales pendant l'action, etc. De plus, la forte référence des programmes scolaires aux pratiques sportives s'accompagne d'une remise en cause – paradoxale – de la compétition⁹, et de la nécessité de différencier parfois difficilement la performance atteinte du niveau de maîtrise de la tâche. Cela aboutit, par exemple en sport collectif, à ne pas vouloir prendre en compte le résultat d'un match dans l'évaluation, alors que celui-ci rend compte de l'atteinte du but du jeu. Faut-il « bien » jouer ou gagner ? Peut-on être plus compétent que ses adversaires en ayant perdu ses matches ? Et dans ce cas, comment gérer les groupes pour créer de façon la plus juste possible un rapport de force équitable et apprécier au mieux les compétences de chacun ?

L'évaluation en EPS renvoie à de multiples dilemmes. De fait, les réflexions engendrées aboutissent à des propositions concrètes, comme l'obligation de rédiger un projet pédagogique dès 1989, pour proposer un système docimologique applicable pour tous les élèves. Les évaluations mises en place doivent être rédigées dans leurs modalités de mises en œuvre et proposer un cadre commun pour un même niveau de classe, tout en laissant une marge de manœuvre à chaque enseignant pour s'adapter à ses élèves. Cette recherche d'harmonisation est renforcée par le fait que l'épreuve d'EPS aux examens est l'une des seules à se dérouler sous forme de contrôle en cours de formation, où l'enseignant de la classe évalue ses propres élèves, accompagné par un de ses collègues. De fait, les référentiels d'évaluation sont très cadrés, tant par le choix limité d'activités proposées, que par les critères et barèmes de notation spécifiés. Cette particularité a l'avantage d'impliquer les enseignants dans l'évaluation des compétences de leurs élèves ; en revanche, le choix des activités en terminale – et plus largement au lycée – s'en trouve réduit, limitant ainsi l'ouverture culturelle de l'enseignement ; une liste officielle spécifie les activités possibles, mais surtout, les enseignants choisissent et éliminent d'emblée certaines activités au regard des statistiques relatives aux notes obtenues qui paraissent chaque année. La pression exercée par l'évaluation externe au baccalauréat se répercute sur les niveaux de classe antérieurs.

Une autre problématique majeure est spécifique à l'EPS ; c'est la forte différence de résultats à l'avantage des garçons. À travers l'étude historique de l'EPS, Liotard (1996) montre que cette discipline favorise la reproduction et le renforcement des inégalités de genre, puisque

⁹ La compétition est une caractéristique inhérente aux sports et la supprimer reviendrait à dénaturer le sens de ces activités. Or, en EPS, l'hétérogénéité du public et les enjeux éducatifs sont peu compatibles avec l'usage récurrent de la compétition.

l'EPS est basée avant tout sur un univers sportif et masculin. Les enseignants conçoivent rarement des évaluations adaptées aux caractéristiques morphologiques et physiologiques de leurs élèves (qu'ils soient filles ou garçons d'ailleurs), alors que de fortes différences interindividuelles définissent leur groupe (Vigneron, 2004). La visibilité aux yeux de tous du niveau de compétence de chacun, des erreurs commises, et la mise en avant du corps renforcent cette nécessité de trouver les moyens de sortir des évaluations normatives pour s'orienter vers une approche plus juste – donc différenciée et formative – de l'évaluation.

62. Les difficultés spécifiques chez les enseignants débutants

Loin de créer une standardisation des pratiques, les théories enseignées en formation provoquent de nombreuses interprétations et créent des conflits de représentations. Seule, la formation ne peut assurer une évolution sensible des pratiques. Les enseignants n'accordent pas la même importance aux différents types d'évaluation : Issaieva et *al.* (2011) montrent que les priorités s'organisent autour de deux axes fondamentaux, formatifs ou sommatifs, selon une logique normative. Dans l'étude présentée ici, nous retrouvons ces deux tendances, corrélées à l'ancienneté et au grade.

Au sortir de la formation initiale, les priorités des enseignants sont régulièrement en décalage avec les attentes des formateurs. Araújo-Oliveira et *al.* (2011) ont montré que les enseignantes débutantes au primaire recherchent davantage l'intérêt suscité chez les élèves que les apprentissages à construire. La gestion de classe par exemple est primordiale ; pour Rayou et *al.* (2004), faire classe s'avère être pour la plupart un travail d'adaptation complexe au contexte dont le public est un élément central. Le manque de professionnalisation de leur formation – soulevé par les jeunes enseignants –, est un véritable problème pour apprendre à gérer la diversité et orienter sa pratique vers l'apprentissage. Différents moyens sont utilisés pour gérer cette complexité, dont la notation comme nous l'avons vu, qui fait partie de l'arsenal pédagogique au service du contrôle de la classe, surtout en début de carrière (Merle, 1996).

L'étude présentée ici a aussi mis en évidence que les enseignants ont des intentions favorables à l'apprentissage mais qui entrent parfois en contradiction avec les pratiques. La complexité des situations d'enseignement, renforcée en début de carrière, rend difficile l'appropriation et la mise en application de principes pédagogiques qui sont non mécaniques et qui nécessitent un accompagnement et une réflexion de longue durée. Pour palier à ces écarts, l'accompagnement est primordial mais aussi l'apport de connaissances : par exemple,

comment évaluer, noter et justifier de son activité sans avoir connaissance des biais de la notation ? Les connaître serait susceptible d'améliorer la fiabilité et validité des évaluations et systèmes de notation mis en place, même si l'objectivité ne sera jamais possible. De plus, cela pourrait entraîner un jugement professoral plus flexible à propos des compétences des élèves, et induire peu à peu des pratiques davantage orientées vers une finalité pédagogique.

63. La notation : un moyen à l'encontre des finalités éducatives

L'étude présentée ici a fait ressortir la place prépondérante de la notation dans les pratiques évaluatives. Selon Merle (1998), si l'on entend souvent les professeurs regretter que les élèves ne travaillent que pour les notes, il est nécessaire de comprendre que, pour les élèves, les notes équivalent au salaire. Elles sont un puissant argument psychologique, source de stress et de pression. Comme l'ont montré entre autres Butera et *al.* (2006) ou Pulfrey (2010), trop de notation, de compétition et de hiérarchisation entre les élèves génère une décentration des buts d'apprentissage et favorise les buts de performance, voire d'évitement. Cela n'est pas sans conséquence sur la réussite scolaire. Cette pratique fréquente de la notation ne permet d'appréhender que la performance, au détriment de l'apprentissage (Butera et *al.*, 2011).

Or, si l'on se réfère aux caractéristiques de l'évaluation formative, la menace de l'échec n'existe pas *a priori*. Les effets d'un dispositif formatif ne peuvent être néfastes si celui-ci n'est pas utilisé à d'autres fins que pédagogiques (De Landsheere, 1992 ; Talbot, 2009). Seulement, Talbot (*ibid.*) regrette que certaines dimensions de l'évaluation soient bien souvent occultées dans les pratiques. Cela a pour conséquence une centration sur les finalités sociales. D'après les résultats de notre étude, les enseignants débutants sont donc au centre d'un paradoxe : ils pensent que les notes motivent les élèves, minimisent le stress engendré et, de ce fait, ils habituent les élèves à travailler dans une logique de performance, ce qui rend difficile ensuite le retour à une évaluation spécifiquement formative.

7. Conclusion

Même si les théories foisonnent autour des différents types d'évaluation, de leur utilité voire de leurs dérives, les pratiques ne reflètent pas ces apports théoriques (Fabre, 2006) ou du moins, elles mettent beaucoup plus de temps à évoluer pour prendre en compte les avancées issues de la recherche. De la même manière, les enseignants ont besoin de temps pour assimiler les données qu'ils reçoivent en formation et pour pouvoir les expérimenter sur le

terrain et apprécier à long terme leurs effets positifs. Les conditions dans lesquelles se déroulent les premières expériences d'enseignement n'aident pas les plus jeunes à centrer leur activité vers l'apprentissage des élèves. Le plus urgent est ailleurs. Toutefois, ce constat n'est pas nouveau : Huberman (1987), à travers une étude de quatre ans menée auprès d'enseignants du secondaire de la Suisse Romande, montrait déjà les difficultés des jeunes enseignants à concilier leur formation initiale et leurs premiers temps d'enseignement. L'étude présentée a pu mettre en avant des différences significatives qui existent notamment entre les jeunes enseignants d'EPS et les plus expérimentés ou formés, en matière d'évaluation. Ces différences se concrétisent essentiellement par une confusion des fonctions de l'évaluation et une redéfinition des évaluations formatives qui, de fait, sont détournées de leur finalité et utilisées comme des évaluations sommatives intermédiaires. Pourtant, l'étude de Braxmeyer et al. (2004) concernant les pratiques évaluatives au collège, souligne que les enseignants en EPS semblent mieux formés à l'évaluation que leurs collègues et motivés davantage vers d'autres finalités que sommatives. La spécificité de cette discipline a permis de faire évoluer plus rapidement les conceptions des enseignants en les plaçant face à des problématiques concrètes. L'approche proposée ici mérite d'être étendue à l'ensemble des disciplines et approfondie. Cela serait aussi l'occasion de mesurer les conséquences provoquées par les évaluations externes, en termes de priorités accordées par les enseignants, de dérives des pratiques et d'effets sur les ressources affectives et relationnelles des élèves. À ce propos, Abelhauser et al. (2011, p.9) mentionnent les paradoxes du culte du chiffre, largement prôné actuellement dans de nombreux champs, et mettent en lumière le fait que « l'évaluation se constitue comme un remarquable outil d'asservissement social ». Par conséquent, la formation des enseignants à l'évaluation constitue une étape décisive pour garantir des conditions optimales d'apprentissage.

Abelhauser A., Gori R. et Sauret J.-M. (2011). *La folie évaluation. Les nouvelles fabriques de la servitude*. Paris : Mille et une nuits.

Allal L. et Mottier Lopez L. (Ed.) (2007). *Régulation des apprentissages en situation scolaire et en formation*. Bruxelles : De Boeck.

Araújo-Oliveira A., Lisée V., Lenoir Y. et Maubant P. (2011). « Planification des savoirs d'enseignement chez des futures enseignantes du primaire au Québec : résultats d'une enquête par entrevues ». In P. Maubant (dir.), *Enjeux de la place des savoirs dans les pratiques*

éducatives en contexte scolaire. Compréhension de l'acte d'enseignement et défis pour la formation professionnelle des enseignants. Québec : Presses de l'Université du Québec, p. 129-155.

Baillat G. (2009). « Formation initiale : bilan contrasté, avenir incertain ». *Entretien du CRAP, L'actualité éducative* n°471 (mars 2009). Consultable sur internet :

<http://www.cahiers-pedagogiques.com/Formation-initiale%E2%80%89bilan>

Brau-Antony S. (2001). « Les conceptions des enseignants d'éducation physique et sportive sur l'enseignement et l'évaluation des jeux sportifs collectifs : résultats d'une enquête ». *STAPS*, 56, p. 93-108.

Braxmeyer N., Guillaume J.-C. et Levy J.-F. (2004). « Les pratiques d'évaluation des enseignants au collège ». *Note évaluation.* MEN.

Bressoux P., Bru M., Altet M., et Lambert C. (1999). « Diversité des pratiques d'enseignement à l'école élémentaire ». *Revue Française de Pédagogie*, 126, p. 97-110.

Bressoux P. et Pansu P. (2003). *Quand les enseignants jugent leurs élèves.* Paris : PUF.

Butera F., Darnon C., Buchs C., et Muller D. (2006). « Les méfaits de la compétition : comparaison sociale et focalisation dans l'apprentissage ». In R.-V. Joule et P. Huguet (Éd.), *Bilans et perspectives en psychologie sociale.* Grenoble : PUG, p. 14-44.

Butera F., Buchs C. et Darnon C. (2011). *L'évaluation, une menace ?* Paris : PUF.

Crahay M. (2010). « Les reformes pédagogiques échouent et pourtant l'école change ! » In Gilliéron Giroud P. et Ntamakiliro L. (dir), *Réformer l'évaluation scolaire : mission impossible?* Berne : Peter Lang, p. 231-264

Crahay M. (2007). *Peut-on lutter contre l'échec scolaire ?* Bruxelles : De Boeck.

Coburn C. (2004). « Beyond Decoupling: Rethinking the Relationship between the Institutional Environment and the Classroom ». *Sociology of Education*, 77 (3), p. 211-244.

David B. (dir.) (2000). *L'éducation physique et sportive : certification au Baccalauréat.* Paris : INRP, Documents et travaux de recherche en éducation n°39.

De Landsheere G. (1992). *L'éducation et la formation.* Paris : PUF.

Fabre M. (2006). « Autour des mots. Analyse des pratiques et problématisation. » *Recherche et formation*, 51, p. 133-145.

Grisay A. (1988). « La pédagogie de maîtrise face aux rationalités inégalitaires des systèmes d'enseignement ». In Huberman, M. (dir.), *Maîtriser les processus d'apprentissage. Fondements et perspectives de la pédagogie de maîtrise*. Paris : Delachaux & Niestlé, pp. 235-265.

Hadji C. (1997). *L'évaluation démystifiée*. Paris: E.S.F. Coll. Pratiques et enjeux pédagogiques.

Huberman M. (1987). « La recherche sur la formation : quelques hypothèses provocatrices ». *Recherche et formation*, 1, p.11-25.

Issaieva E., Pini G. et Crahay M. (2011). « Positionnement des enseignants et des élèves du primaire face à l'évaluation : une convergence existe-elle ? » *Revue française de pédagogie*, 176, p. 5-25.

James W. (1978 [1898]). « The pragmatic method ». In F. H. Burkhardt (General Ed.), *The works of William James, Essays in Philosophy in the works of William James*. Cambridge: Harvard University Press.

Lentillon V. (2008). « Les élèves de second degré face à l'évaluation en éducation physique et sportive ». *STAPS*, 79, p. 49 à 66.

Liotard P. (1996). « L'impossible spécificité de l'éducation physique féminine ». In P. Arnaud et T. Terret (Coord.), *Histoire du sport féminin. Sport masculin-sport féminin*. Tome 2. Paris : L'Harmattan.

Merle P. (1996). *L'évaluation des élèves. Enquête sur le jugement professoral*. Paris : PUF.

Merle P. (1998). *Sociologie de l'évaluation scolaire*. Paris : PUF.

Morissette J. (2010) « Une perspective interactionniste. Un autre point de vue sur l'évaluation des apprentissages », *SociologieS*. Disponible sur internet :

<http://sociologies.revues.org/3028>

Pulfrey C.-J. (2010). « Capitalism in the classroom : the impact of self-enhancement values and normative assessment on student motivation, goals and attitudes towards cheating ». Thèse de Doctorat. Université de Lausanne, Faculté des sciences sociales et politiques.

Rayou P. et Van Zanten A. (2004). *Enquête sur les nouveaux enseignants. Changeront-ils l'école ?* Paris : Bayard.

Talbot L. (2012 [2009]). *L'évaluation formative. Comment évaluer pour remédier aux*

difficultés d'apprentissage. Paris : Armand Colin, collection sciences humaines et sociales.

Vigneron C. (2004). « La construction des inégalités de réussite en EPS au baccalauréat entre filles et garçons. Thèse de doctorat, Univ. de Bourgogne (Dijon).