

INTRODUCTION

« La foi dans l'harmonie par le calcul a ressurgi avec la globalisation ultra libérale et les revendications libertariennes, qui sont les deux faces – économique et “sociétale” – d'une même médaille. Mais [...] la promesse moderne d'émancipation n'est pas faite pour les faibles : elle condamne à la misère ou à la solitude la masse de ceux qui ne parviennent pas à s'affirmer en maîtres dans un univers de compétition généralisé. » (Supiot, 2015, p. 292.)

Les procédures d'évaluation – qui envahissent tous les secteurs du réel – se présentent le plus souvent sous la forme de données fiables et objectives, qui ont pour vocation principale d'atténuer le parti pris de l'évaluateur. Quant à cette invasion, elle s'inscrit à l'intérieur d'une mutation idéologique plus vaste qui a institué la mise en concurrence comme le principal facteur de la performance. En pratique, ces procédures ne se contentent pas d'informer mais elles affectent la nature du sujet ou de l'objet évalué ; si bien que l'évaluation pose – quel que soit son champ d'application – un problème éthique. La nature d'un problème éthique consiste précisément dans le fait que la solution du problème n'est jamais donnée par avance et qu'elle doit faire l'objet d'une invention singulière qui n'est jamais assurée en son résultat. Dans la mesure où les procédures d'évaluation sont incontournables, une évaluation éthique essaie de tenir compte de la singularité des situations, tout en cherchant à atténuer les effets violents et discriminants inhérents à l'évaluation elle-même. Nous ne cherchons donc pas ici à condamner l'évaluation en elle-même, mais à promouvoir une approche éthique qui tienne compte de la complexité des situations. Tout au long de cet ouvrage, nous interrogerons les connaissances disponibles issues de la recherche en éducation pour appuyer et discuter nos propos.

L'OMNIPRÉSENCE DE L'ÉVALUATION

L'évaluation justifie son omniprésence en faisant croire à sa nécessité, à sa fiabilité. Elle se présente comme une procédure témoignant sans cesse à l'évalué qu'il peut accomplir des progrès, dépasser ce qu'il est pour être toujours plus productif et plus efficace. Officiellement, les évaluations sont censées apporter des données fiables et bénéfiques dans un contexte de transparence, mais dans la majorité des cas, elles consistent seulement à divulguer des données chiffrées susceptibles de hiérarchiser les individus ou les objets suivant une visée normative. Concrètement, cela signifie que seuls les résultats affichés – les productions – attirent l'attention, au détriment d'éléments qualitatifs nécessaires à la compréhension de ces résultats.

Le milieu scolaire n'échappe pas à cette logique. L'évaluation y est omniprésente et elle est nécessaire, pourvu qu'elle soit utile à l'élève pour l'aider à progresser. Dans ce cas, il est nécessaire de s'interroger sérieusement sur les procédures d'évaluation proposées ; cela concerne de façon non exhaustive la nécessaire cohérence entre les compétences visées et la situation d'évaluation prévue, mais aussi la pertinence et la transparence des référents, critères et indicateurs retenus, le choix du moment prévu pour évaluer, la fiabilité et la bienveillance de l'évaluateur, la gestion des erreurs, le type d'informations transmises aux élèves voire aux parents, l'implication des élèves dans leur propre évaluation ou celle de leurs pairs... Les caractéristiques d'une évaluation qui permettrait à l'élève de mieux apprendre ne peuvent faire l'objet d'une simplification ou de l'application de principes établis selon le « bon sens », car les enjeux qui en découlent ensuite sont décisifs pour les sujets évalués.

Les évaluations scolaires internationales (par exemple PISA ou le classement de Shanghai), nationales, les classements d'établissements scolaires, la comparaison des classes entre elles sont devenues des données *a priori* voire naturellement fiables ou objectives pour comparer l'efficacité de l'enseignement et le niveau de performance des élèves. Ils suscitent de vifs débats et provoquent des sentiments néfastes, surtout quand les évalués s'aperçoivent qu'ils sont plutôt dans le bas de l'échelle. Au plan national, ces classements affectent peut-être moins les individus scolarisés faisant l'objet de ces évaluations car ce ne sont jamais eux – individuellement – qui sont évalués et montrés, mais le groupe scolaire auquel ils appartiennent. Par contre, plus on se situe à l'échelle d'une classe ou d'un individu, plus les conséquences de ces comparaisons peuvent être pernicieuses et entraîner des effets inverses à ceux attendus. Pour Abelhauser *et al.* (2011, p. 17) : « On veut nous faire oublier ainsi que l'évaluation repose sur une logique gestionnaire et sur des techniques de management dont les conséquences sont pires que les bénéfiques attendus. » Échouer, c'est montrer son incompetence, c'est montrer qu'on n'a pas réussi comme les autres, qu'il nous manque quelque chose. Pour Miller (2008, p. 63) : « Il y a dans la culture de l'évaluation le désir d'imposer des normes, d'humilier l'autre, de le faire céder sur son être. » Évaluer en référence à une norme, c'est renvoyer chacun à sa place, lui rappeler ce qu'il vaut. Mais selon quels critères et suivant l'autorité de qui ?

L'évaluation chiffrée, qui quantifie l'individu, n'est pourtant pas préconisée explicitement à l'école. Les textes officiels, les rapports de l'inspection générale ne cessent de plaider pour une évaluation positive qualitative. Mais, comme nous le verrons par la suite, les résistances sont nombreuses, tant du côté des parents d'élèves et des enseignants, que du côté des politiciens. Il semble opportun de parler de *résistances*, car la recherche universitaire a montré l'impact positif de certaines procédures d'évaluations dans une séquence d'apprentissage – par sa fonction de régulation (Scallon, 1982), ou par l'utilité des informations transmises aux élèves (Schmidt, 1992). Or, on peut s'apercevoir que dans les pratiques professorales, les évaluations proposées sont majoritairement normatives, faisant état d'une proximité quasi systématique avec la notation (OCDE, 2008) et qu'elles se fondent le plus souvent sur une comparaison des individus entre eux.

Ce culte du chiffre ne concerne pas exclusivement le milieu éducatif, mais la « gouvernance par les nombres » (Supiot, 2015) et l'imaginaire d'une harmonie globale fondée sur le calcul affectent l'ensemble des domaines du réel. Or, il peut être dangereux de « penser

qu'il puisse y avoir une évaluation automatique chiffrée, systématique, complètement objective, et qui vous appelle à participer à cette objectivité, alors que vous avez bien le droit de ne pas être comme les autres, d'avoir une pratique différente » (Miller, 2008, p. 63). La notation – qui est le moyen le plus utilisé à l'école pour rendre compte des résultats d'une évaluation – ne permet pas la prise en compte de la singularité des personnes. Elle permet seulement de comparer un résultat chiffré à une moyenne (de classe, d'école, régionale, etc.), elle-même largement dépendante du contexte et de l'évaluateur (Merle, 1998) ; elle sanctionne le sujet au regard d'une norme à atteindre et favorise en priorité la perception des manques, des erreurs et des problèmes, plutôt que celle des acquisitions ou des procédures utilisées par chacun. Elle induit la comparaison aux autres sans questionner de façon fine la validité des critères qui ont été utilisés pour parvenir à ces résultats. Elle met en avant la performance atteinte, la production finale, et minimise l'originalité des réponses, la personnalité et les réussites de chacun.

Deux difficultés majeures émergent et limitent considérablement les effets attendus des évaluations scolaires : d'une part, cette quantification des résultats qui entraîne une perte de sens et, d'autre part, la multiplication des évaluations comme moyen de contrôle. Pour Foucault (1994, p. 74) : « Nous sommes entrés dans un type de société où le pouvoir de la loi est en train non pas de régresser, mais de s'intégrer à un pouvoir beaucoup plus général : celui de la norme. Ce qui implique un système de surveillance, de contrôle tout autre. Une visibilité incessante, une classification permanente des individus, une hiérarchisation, une qualification, l'établissement de limites, une mise en diagnostic. » À de nombreuses reprises dans cet ouvrage, nous ferons référence aux travaux de recherche qui s'intéressent aux évalués, en mettant en avant les effets nocifs des évaluations chiffrées, que ce soit pour les élèves les moins compétents au regard des normes scolaires, mais aussi pour ceux qui se trouvent en haut de l'échelle. La multiplication des évaluations scolaires ne peut se faire au détriment des apprentissages et au détriment du bien-être des élèves. Nous aborderons cette problématique en nous appuyant sur les résultats d'une expérimentation menée à l'école.

ÉVALUER POUR MIEUX APPRENDRE : REPENSER L'ÉVALUATION À L'ÉCOLE

L'évaluation est très controversée à l'école. De nombreuses études tendent à proposer des outils ou des méthodologies d'évaluation pour optimiser les apprentissages, pour valoriser les élèves et les impliquer davantage (dernièrement, Figari *et al.*, 2014) ; d'autres montrent que dans les pratiques, les modalités mises en œuvre tendent toujours à privilégier le versant normatif et sanctionnant qui ne permet pas aux élèves de mieux apprendre (Butera *et al.*, 2006 ; Dubet, 2009). C'est précisément cet écart – entre les effets attendus et les effets observés – qui a conduit à réaliser une étude comparative pour mieux appréhender les effets de différentes modalités évaluatives sur l'apprentissage. Si l'on veut repenser l'évaluation à l'école, il est nécessaire d'interroger les modalités mises en place par les enseignants, pour mesurer la portée des évaluations et ainsi éviter de tomber dans le contrôle, c'est-à-dire dans l'évaluation qui normalise et qui rejette les individualités. Pour que l'évaluation soit positive – du point de vue de l'optimisation des apprentissages et de la valorisation des individus concernés –, il ne s'agit pas seulement de prétendre

poursuivre des finalités éducatives (telles que la réussite de tous, le développement de la solidarité ou de l'autonomie), il faut également que tout un travail complexe de réflexion soit effectué en aval et en amont.

L'observation des pratiques professorales montre une hétérogénéité de l'activité évaluative des enseignants ; la mise en œuvre de pratiques variées témoigne d'une adaptation de chaque enseignant à son contexte d'enseignement. Si les pratiques sont disparates, cela peut être aussi consécutif à des priorités individuelles divergentes : certes, l'enseignant s'adapte au contexte, mais les décisions prises semblent aussi émaner d'habitudes voire d'intuitions (Tochon, 1989). Il peut s'agir d'évaluer souvent pour, par exemple, augmenter l'engagement des élèves, ou encore de mettre systématiquement en avant les meilleurs élèves, de les récompenser pour espérer motiver les plus faibles et leur donner envie de réussir aussi bien. Seulement, les habitudes qui émanent souvent d'expériences empiriques et de justifications par le bon sens, peuvent aussi être la résultante de choix par défaut. Dans ce cas, le manque de formation pourrait expliquer l'absence de choix des enseignants ; mais il n'est pas suffisant, car la relation entre les théories et la pratique dépasse la seule connaissance des dispositifs révélés efficaces par la recherche (Gélin *et al.*, 2007). 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 de données théoriques. Si la théorie et la pratique apparaissent interdépendantes pour assurer une formation cohérente, les liens entre ces deux entités sont complexes (Crahay, 2007). La mise en pratique des savoirs enseignés en formation n'est pas un automatisme et elle fait émerger toute la complexité de la situation d'enseignement ; cela nécessite notamment une adaptation au contexte qui est peu prévisible, surtout pour les enseignants débutants (Rayou *et al.*, 2004). L'hétérogénéité des pratiques professorales n'est donc pas seulement le fruit d'une adaptation de l'enseignant au contexte, mais elle est au carrefour d'une multitude de paramètres qui vont influencer les choix. Ceux-ci sont donc à questionner, pour appréhender leur capacité à atteindre les finalités scolaires.

Même si l'évaluation scolaire est une pratique transversale aux disciplines, elle pose – pour chacune d'entre elles – des problèmes spécifiques de mise en œuvre. D'après David (2000), la recherche par les enseignants de procédures évaluatives précises et objectives conduit à enseigner au final uniquement ce que l'on peut évaluer. Les études que nous allons présenter dans cet ouvrage ont été réalisées en éducation physique et sportive (EPS). Dans cette discipline, le contrôle en cours de formation mis en place lors des examens scolaires incite les enseignants à définir les contenus d'enseignement en fonction des évaluations, ce qui a pour conséquence la mise à l'écart d'un certain nombre d'apprentissages et une centration unique sur un niveau de compétence à atteindre. Par ailleurs, l'enseignant d'EPS évalue des conduites motrices¹ à la grande différence de ses collègues des autres disciplines. Cette spécificité amène des difficultés supplémentaires, comme la nécessité d'observer chaque élève pendant l'action, 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 tout de suite ses observations

1. Parlebas aborde la notion de conduites motrices dès 1968. Elle se définit comme « l'organisation signifiante du comportement moteur » (1999, p. 74). L'emploi de ce terme permet de se démarquer de la notion de comportement en mettant en avant le sens des actions motrices réalisées par les pratiquants ou les élèves.

en niveau de compétence ou d'habileté selon les critères définis. La prise de décision est donc très rapide et demande une forte maîtrise des critères d'évaluation et des techniques d'observation.

Avoir pour ambition de proposer aux élèves des procédures d'évaluation au service des apprentissages – selon une logique qui rompt avec l'évaluation normative –, a conduit à questionner d'abord les enseignants sur leurs conceptions de l'évaluation, pour ensuite comparer *in vivo* des modalités variées d'évaluation afin d'en apprécier les effets réels. Il ne s'agira pas dans cet ouvrage de donner des pistes globales pour proposer des évaluations dites positives ou bienveillantes selon un mode prescriptif, mais nous montrerons précisément les effets que peuvent avoir différentes modalités d'évaluation sur l'apprentissage des élèves ainsi que sur leur plaisir d'agir.

POUR UNE APPROCHE ÉTHIQUE DE L'ÉVALUATION

Dans la mesure où l'évaluation interfère avec les conduites des élèves, elle pose un problème éthique. Autrement dit, il s'agit d'examiner le bien-fondé d'une pratique et son impact sur le devenir de ceux qui en sont affectés. En effet, selon Jaffro (1995, p. 222), l'éthique est un « point de vue que tout un chacun peut prendre sur une situation difficile ». L'évaluation pose un problème éthique car il n'y a pas de règles préconçues qui puissent être appliquées et qui seraient applicables à tout contexte ; et toute évaluation n'est jamais neutre dans le sens où elle a des effets (parfois difficilement perceptibles de l'extérieur) sur les personnes évaluées. La question de l'objectivité par exemple a toute sa place : nous aborderons dans cet ouvrage un ensemble conséquent de biais, démontrés depuis fort longtemps, qui remet en question la notation et la fiabilité du correcteur. La rigueur et l'objectivité apparentes des résultats chiffrés qui sont donnés par un enseignant – parfois au quart de point près sur une échelle de 0 à 20 – sont finalement largement contestables d'un point de vue scientifique. Mais l'enjeu est important, car les résultats d'une évaluation sont davantage perçus comme un verdict final qui touche à la valeur de soi, alors qu'ils ne devraient estimer que la valeur d'une production au regard de certains critères définis préalablement.

Dans la mesure où la solution du problème n'est jamais assurée, les problèmes éthiques ne relèvent pas de l'ordre du savoir. Celui qui revendique d'évaluer avec bienveillance et de façon éthique se questionne sur l'ensemble des effets et sur l'utilité de l'évaluation proposée. À l'inverse, celui qui n'examine pas sa pratique, qui ne se questionne pas sur le bien-fondé de ce qu'il propose, n'éprouve aucune inquiétude éthique. Le problème ne se pose pas, il n'est pas posé. Pour que l'aspect problématique d'une situation apparaisse, il faut que le témoin de la situation soit suffisamment heurté pour entrer dans une démarche réflexive.

Ce choix de confronter les conceptions des enseignants avec l'activité qu'ils développent et les apports des recherches est issu d'une volonté de notre part d'analyser l'efficacité de ce qui est proposé quotidiennement aux élèves. À travers l'efficacité, nous pensons aux effets des évaluations sur la régulation de l'apprentissage et sur le plaisir des élèves à apprendre. Nous pensons aussi à la cohérence entre les indicateurs ou critères d'évaluation et la situation proposée. Il paraît évident que nombre de décisions prises sont le fait

d'habitudes (Tochon, *op. cit.*). L'expérience et le vécu de chacun incitent à agir de telle ou telle manière mais, malgré la tendance à penser que ces choix sont justifiés et pertinents, nous souhaitons apprécier plus finement leur portée, pour ne pas rester prisonnier du bon sens. « Le bon sens est la chose du monde la plus partagée : car chacun pense en être si bien pourvu que ceux même qui sont les plus difficiles à contenter en tout autre chose, n'ont point coutume d'en désirer plus qu'ils en ont. » (Descartes, 1962, p. 9.) L'enseignant travaille seul avec sa classe, si bien qu'il est difficile pour lui de percevoir d'autres façons de procéder et par conséquent les répercussions que cela entraîne sur les conduites des élèves. Car, au final, qui évalue l'activité des enseignants au quotidien ? Quels retours formatifs peuvent-ils avoir au sujet de leur propre pratique ? On remarque empiriquement par exemple que les enfants peuvent se comporter de façon très différente selon la personnalité des enseignants² ; selon le contexte, le rapport d'autorité, la bienveillance de l'adulte, les conduites diffèrent. Ce phénomène est observable notamment lors des conseils de classe. Les représentations que se font les enseignants d'un élève sont parfois très éloignées entre elles et soulèvent des débats houleux. L'expérience nous a montré qu'il n'est pas rare de constater que le professeur d'EPS est souvent là pour contredire, relativiser les avis négatifs ou au contraire émettre des bémols à propos d'élèves scolairement performants ; il apporte souvent un point de vue différent sur l'élève car le contexte et les situations d'enseignement sont très spécifiques à la discipline. Chacun agit comme bon lui semble, mais dans quelle mesure est-il possible d'affirmer que les pratiques sont efficaces ? Quelles erreurs sont commises de la part de l'enseignant et comment peut-il en prendre conscience ? L'enjeu est de taille puisque les conséquences impactent directement sur l'apprentissage des élèves qui est – lui-même – évalué. En tant qu'enseignant, il est difficile d'estimer la portée des choix ; le rapport d'autorité entre professeur et élève(s) et la portée symbolique des évaluations et notations (Merle, 2004) semblent entraver la prise de conscience de certaines conséquences néfastes des démarches entreprises. De fait, l'évaluation est une procédure qui reflète les différences de conceptions de l'enseignement et de l'élève, qui caractérise les représentations du professeur à l'égard de l'apprentissage. Les situations d'évaluation, les critères choisis, la communication faite aux élèves et les interprétations qui en découlent sont déterminants pour les élèves ; et pourtant ils paraissent très disparates.

Dans ce livre, nous proposerons de réfléchir aux actions entreprises et d'apporter des éléments concrets et testés de façon quasi expérimentale (Matalon, 1988). Comme le souligne Jaffro (1995), l'enseignant ne peut pas décider consciemment du moindre de ses faits et gestes ; en revanche, il peut prendre des décisions éthiques et réfléchir aux actions mises en œuvre et à leurs conséquences sur les autres. L'approche éthique vise ainsi à sortir d'une conception moraliste qui juge l'activité des élèves et leur réussite par rapport à ce qui est bien ou mal, c'est-à-dire par rapport à des normes préétablies, identiques pour tous et qu'ils n'ont pas choisies. Il s'agira au contraire d'essayer de partir des capacités des élèves pour les faire progresser dans une situation donnée, de se concentrer sur ce qu'ils font et non sur ce qu'ils devraient faire.

Pour mieux comprendre la conception contemporaine de l'évaluation, il nous a semblé indispensable dans une première partie de revenir sur l'histoire de l'évaluation et de la

2. Des recherches ont montré par exemple l'impact des attentes sur les conduites comme l'effet Pygmalion (ROSENTHAL et JACOBSON, 1968).

notation à l'école et en EPS plus spécifiquement, pour aborder l'activité évaluative des enseignants d'aujourd'hui ; différents résultats de recherches seront apportés et confrontés à une enquête que nous avons réalisée auprès des enseignants en EPS. Cela nous permettra de comprendre les conceptions des enseignants actuellement et les nombreuses résistances qui freinent le passage à une évaluation positive et bienveillante. Ensuite, dans une deuxième partie, nous nous centrerons sur l'impact des évaluations sur les conduites des élèves. Nous nous intéresserons à l'effet de modalités évaluatives différentes sur l'apprentissage et le ressenti des élèves en EPS. Nous analyserons de façon très précise leurs conduites motrices pour cerner les axes de progrès et nous compléterons ces données par une analyse des conduites verbales de quelques élèves. Ces données issues d'une recherche menée en milieu scolaire seront accompagnées de nombreuses références faites à d'autres études, dans le but de mieux comprendre l'intérêt de certaines procédures d'évaluations sur l'apprentissage et le ressenti des élèves. Nous nous attarderons aussi dans cette partie à mettre en relation la progression et le niveau des élèves avec certaines de leurs caractéristiques individuelles et socioaffectives pour mieux cerner les inégalités engendrées par l'enseignement des sports collectifs à l'école. Enfin, la dernière partie de cet ouvrage sera un aboutissement des précédentes parties qui mettra l'accent sur la prédominance de la compétition scolaire et ses effets néfastes sur l'apprentissage et le développement de la personne.