

Évaluer des compétences

Didier Delignières
Université de Montpellier

10 octobre 2022

<https://didierdelignieresblog.wordpress.com/2022/10/10/evaluer-des-competences/>

L'évaluation reste une zone d'ombre dans le système éducatif. Si chacun s'accorde sur sa nécessité, il y a peu de consensus sur ses finalités. Est-elle au service de la hiérarchisation objective des élèves, pierre angulaire de leur sélection et de leur orientation ? Est-elle dédiée au suivi des apprentissages, à la régulation de l'enseignement, à la construction de sens pour les élèves ? Lorsque l'on questionne les étudiants à ce sujet, ils évoquent en premier lieu le bréviaire hérité des années didactiques, distinguant doctement évaluations diagnostique, formative, formatrice, sommative, certificative. Mais dans ce domaine comme dans d'autres, le fait d'être capable d'énumérer une liste à puces ne garantit guère que l'on ait compris la question.

L'objectif de ce texte est principalement d'analyser les représentations, les idéologies, les pesanteurs, qui formatent la réflexion et les pratiques, dans la mise en œuvre de l'évaluation en général, et de manière plus particulière de l'évaluation des compétences. Il s'agit, comme la plupart des textes publiés sur ce site, d'une invitation à la prise de recul, à la réflexion critique. On n'y trouvera pas de solutions « clé en main », de procédures directement exploitables sur le terrain, mais plutôt des principes et analyses permettant à chacun d'interroger ses propres pratiques.

Le poids de l'évaluation certificative

Toute réflexion sur l'évaluation est vite polluée par des représentations tenaces liées à l'activité de notation. Même si chacun est capable de clairement différencier une évaluation support de l'apprentissage et une évaluation destinée à rendre des comptes à l'institution, il reste difficile pour les enseignants de les concevoir comme deux processus complètement indépendants.

L'évaluation certificative reste dans le système actuel un élément central. Même si l'EPS n'est pas souvent considérée comme déterminante, par rapport à d'autres disciplines, il n'en demeure pas moins qu'elle participe aux procédures d'orientation, ou de passage dans l'année supérieure. On pense évidemment aux évaluations réalisées en vue des examens, mais au-delà, les notes attribuées, notamment au lycée, sont prises en compte dans l'examen des dossiers lors de la procédure Parcoursup. L'évaluation génère de fortes attentes, et aussi de fortes angoisses, chez les élèves et leurs parents, une réalité qu'il est difficile de ne pas prendre en compte ([Duru-Bellat, 2022](#) ; [Perrenoud, 2005](#)). Les enseignants sont évidemment sommés de justifier leurs notations, en termes d'équité (la note correspond-elle réellement au travail réalisé ?), de fiabilité (l'enseignant est-il vraiment certain de la note attribuée, un autre enseignant aurait-il donné la même note ?), et de crédibilité (les notes permettent-elles de hiérarchiser sans ambiguïté les élèves, l'enseignant n'a-t-il pas tendance à sur- ou à l'inverse à sous-noter ses élèves ?).

Au-delà des problèmes proprement docimologiques, l'évaluation certificative charrie d'autres dérives, finement analysées par André Antibi (2003) au travers du phénomène de « constante macabre » : la sacralisation de la « moyenne », et de la répartition symétrique des notes de part et d'autre de cette valeur de référence, condamne nécessairement une partie de la classe à l'échec, quel que soit par ailleurs le niveau des élèves. Cette rigueur dans la notation détermine aussi une hiérarchie entre disciplines : les disciplines « sérieuses » sont celles qui génèrent de l'échec, mettant en lumière les réussites d'une élite, alors qu'une discipline qui ferait réussir tous les élèves ne serait qu'une « école des fans », sans exigences significatives ([Observatoire des pratiques en éducation prioritaire, 2021](#)).

Il est évidemment tentant de clamer que l'on va s'y prendre autrement, que l'on va se livrer à une évaluation plus qualitative, que l'on va faire de l'évaluation plus un instrument de pilotage des apprentissages qu'un outil de sélection. Mais comme le note Perrenoud, « *une description plus clinique, qualitative, avec référence à des objectifs, des niveaux de maîtrise, des lignes de progression, n'a de sens que pour des destinataires qui veulent en savoir plus. Ce n'est pas le désir de la majorité des parents, qui n'ont pas les moyens d'interpréter de telles informations* ». Et il ajoute : « *Dès lors qu'ils vivent la scolarité de leurs enfants sous la menace d'un échec, d'une exclusion ou simplement d'une relégation dans les filières secondaires sans avenir, il est normal que les parents se soucient de ce qui se passe en classe, épluchent les carnets scolaires, scrutent les indices avant-coureurs d'un échec* » ([Perrenoud, 2005](#)).

Il n'est pas non plus facile de déconnecter une évaluation formative au service de la progression des élèves d'une évaluation certificative qui donnerait in fine des classements différents. Les élèves eux-mêmes n'y comprendraient pas grand-chose. Philippe Perrenoud estime que l'expertise enseignante devrait essentiellement consister « *à savoir dresser un bilan analytique des acquis, à mesurer le trajet parcouru, à identifier les obstacles et les résistances, à apporter des régulations* ». Mais il reconnaît que « *la guerre des notes empêche de s'attaquer à ce problème. Elle masque les vrais enjeux. Elle fait perdre du temps et de l'énergie au détriment de la lutte contre l'échec scolaire* » ([Perrenoud, 2005](#)).

Ce qu'il faut retenir de ce propos liminaire, c'est que les enseignants restent poussés par l'institution, par les parents et les élèves eux-mêmes, à une évaluation qui hiérarchise et sanctionne, et qu'il ne faut pas minimiser les contraintes qu'ils peuvent rencontrer à penser autrement (voir aussi la [contribution récente de Charles Hadji, 2022](#)).

La « révolution » des compétences

Dans l'ensemble des disciplines scolaires, la mise en avant des compétences a évidemment profondément interrogé les pratiques évaluatives. Tant que l'on en restait sur une pédagogie traditionnelle, centrée sur la restitution des connaissances, des démarches basées sur des questions fermées ou des questionnaires à choix multiples pouvaient s'avérer pertinentes. Mais la pédagogie des compétences amène d'autres exigences. Dans ce cadre, il ne suffit plus à l'élève de restituer les connaissances, mais de les exploiter en synergie pour réagir de manière favorable à des situations complexes.

Toutes les disciplines n'ont pas été interpellées de la même manière par cette évolution. Certaines, assises sur un patrimoine séculaire de savoirs académiques, y ont vu une régression des exigences et un « *abandon des savoirs* » ([Hirtt, 2011](#)). D'autres, comme les sciences de la vie et de la terre, ou l'histoire-géographie, ont pu y voir une opportunité pour mettre leur discipline au service de l'éveil critique des élèves vis-à-vis de grandes problématiques sociétales. Les enseignements technologiques et professionnels ont plutôt accueilli cette évolution de manière favorable, comme reconnaissance des spécificités de leurs contenus.

En EPS, l'introduction du concept de compétence a également été plutôt bien accueillie. Même si l'on peut considérer que dans l'histoire de la discipline, certains courants de pensée ont tenté de faire entrer l'EPS dans le giron des disciplines académiques, par exemple lorsque Delaunay et Pineau ([1989](#)) affirmaient qu'en EPS « *on enseigne des structures, des relations, des principes, des règles...* », la sportivisation de la discipline, à partir des années 60, ouvrait sans conteste la porte à des conceptions de contenus d'enseignement proches de ce que l'on désigne aujourd'hui comme compétences. Dans un texte intitulé « *S'il te plaît écris-moi un programme* », le SNEP avait ainsi justifié son intérêt pour le concept de compétences : « *L'EPS n'est-elle pas plus une discipline de compétences qu'une discipline de savoirs ? et notre implication pour faire « entrer » les compétences reposait sur l'idée simple que l'EPS ne devait pas abandonner le champ des savoir-faire au profit de méta-savoirs (principes...)* » ([SNEP, 2005](#)).

Néanmoins, ravalier les compétences à de « simples » savoir-faire technico-tactiques constitue certainement une impasse. J'ai évoqué dans des contributions récentes la nécessité d'une définition

plus holistique des compétences, intégrant totalement ce qu'il est convenu de désigner sous le vocable de « compétences méthodologiques et sociales », si l'on souhaite réellement que les contenus de la discipline soient sans conteste alignés sur ses finalités, qui depuis une vingtaine d'années sont orientées vers la formation citoyenne ([Delignières, 2022](#)). Et la solution institutionnelle qui consiste à traduire dans le *livret scolaire unique* les acquisitions scolaires en termes de compétences transversales reste un exercice bien artificiel, qui correspond davantage à une justification *a posteriori* de la contribution des enseignements disciplinaires au Socle commun qu'à l'attestation avérée de compétences utiles.

Des principes généraux pour l'évaluation des compétences

Avant de se poser la question de la mise en œuvre pratique de cette évaluation des compétences, il est sans doute nécessaire d'en évoquer les principes généraux, et d'envisager ce qui peut la différencier des procédures traditionnelles. Lorsque l'on questionne les étudiants ou les enseignants à ce sujet, on a souvent l'impression que le fait d'abandonner l'évaluation chiffrée suffit à basculer dans l'évaluation par compétences. C'est sans doute un peu plus complexe, et avant de se demander comment on « évalue par compétences », il faut déjà savoir si ce sont bien des compétences que l'on tente d'évaluer...

Un certain nombre de préconisations ont été avancées pour construire des démarches d'évaluation des compétences (voir notamment [Perrenoud, 2004](#), [Gérard et Braibant, 2004](#), [Gérard, 2007](#) ou [Delignières, 2009](#)). Pour ceux qui n'ont pas trop de réticence à lire un peu d'anglais, on pourra aussi consulter le remarquable article de [Wiggins \(1989\)](#), sur la conception de « tests authentiques » pour l'évaluation des compétences. Ces préconisations concernent toutes les disciplines, même si comme on l'a vu précédemment, toutes ne sont pas impactées de la même manière par cette approche.

Des situations complexes

Tout d'abord, les compétences ne peuvent être évaluées que dans des situations complexes. Ceci signifie essentiellement que les situations d'évaluation ne doivent pas demander la restitution de connaissances localisées et prévisibles. Une situation complexe est avant tout mal définie, c'est à dire qu'il est difficile de la circonscrire de manière définitive. Chacun, sur la base du sens qu'il donne à la situation, peut en redéfinir les contours. Une situation complexe admet une multitude de solutions, plus ou moins pertinentes. C'est justement le degré de pertinence de la solution produite qu'il conviendra d'évaluer. On trouvera dans [Delignières \(2019a\)](#) une analyse des relations théoriques entre les concepts de complexité, de situation complexe et de compétence.

Des situations contextualisées

[Perrenoud \(2004\)](#) insiste sur le fait que les situations d'évaluation doivent être contextualisées, c'est à dire chargées de sens, proches des situations de la vie réelle. On retrouve dans cette préoccupation les réflexions qui ont présidé à l'élaboration du Socle commun de 2015, visant à permettre aux élèves de comprendre et d'agir dans la complexité du monde qui les entoure. Comme il a été dit précédemment, certaines disciplines scolaires ont pu éprouver quelques difficultés à inscrire leurs démarches dans le monde réel. Ce principe de contextualisation résonne dans notre discipline avec le concept de « tranche de vie », introduit par Maurice Portes ([1999](#)) pour désigner les expériences que devaient traverser les élèves dans l'étude des APSA. L'EPS a l'avantage de se référer à une culture vivante, largement diffusée dans la société, et pour peu qu'elle ne renie pas cet ancrage, elle a peu de chance de basculer dans le formalisme et l'abstraction. Il convient cependant de rester vigilant à ce niveau, à l'heure où l'on voit émerger des tendances récurrentes au retour à l'hygiénisme et à la défiance vis-à-vis de la culture sportive ([Delignières, 2020a](#)).

Des situations collaboratives

On ne peut envisager la construction des compétences et leur évaluation en faisant abstraction des relations entre pairs et de la dynamique des groupes. [Wiggins \(1989\)](#), dans la tradition curriculaire nord-américaine, affirme que « *les tests authentiques nécessitent une certaine collaboration avec*

d'autres. La plupart des défis professionnels auxquels sont confrontés les adultes impliquent la capacité d'équilibrer les réalisations individuelles et collectives ». Le Socle commun de 2015 retient, d'une manière plus générale, la nécessité d'être capable de vivre ensemble, de participer à des projets collectifs, comme ciment d'une citoyenneté progressiste. Coston, Testevuide et Ubaldi (2013) expriment clairement comment, dans notre discipline, le collectif est incontournable : « *la construction des compétences pour une personne, c'est toujours une affaire de rencontre entre « lui » et un « autre »* ». *L'autre : cet enseignant qui construit des milieux « consistants », « résistants » aux velléités de contournement des élèves. L'autre, ce pair, partie prenante des projets du pratiquant, l'autre comme « aide » pour faire mieux comprendre, apprendre et interagir dans des contextes porteurs de savoirs et de culture que sont les APSA. En EPS il est toujours question de s'inscrire dans un projet avec les autres* ». On peut évidemment ajouter que l'EPS, parce que ses pratiques de référence sont essentiellement des pratiques collectives (bien au-delà d'ailleurs des « sports collectifs »), intègre naturellement cette dimension, au contraire d'autres disciplines dans lesquelles elle peut apparaître comme une contrainte artificielle (Delignières, 2019b). On conçoit qu'en EPS, une tranche de vie authentique ne saurait se limiter aux aspects moteurs, technico-tactiques, mais intègre de manière indissociable les méthodes d'entraînement et de préparation, les responsabilités et le fonctionnement des groupes.

Une situation attendue, et coutumière

L'évaluation ne doit pas être la première confrontation à la complexité. La situation d'évaluation doit rappeler d'autres situations similaires, de complexité équivalente, travaillées lors de la séquence d'enseignement. En EPS, ce principe a été développé de manière consistante par les propositions du CEDREPS à propos des « formes de pratiques scolaires » (FPS, Ubaldi et al., 2006). Rappelons que cette démarche est basée sur le constat d'un éparpillement néfaste des objectifs en EPS, et de la nécessité de cibler l'activité des élèves, à un moment donné de leur cursus, vers un objet d'enseignement nettement circonscrit. Les auteurs définissent cet objet d'enseignement comme un « pas en avant », réaliste car atteignable par l'ensemble des élèves, et ancré dans la culture de l'APSA car déterminant une étape essentielle dans la formation du pratiquant. Dans ce cadre, une FPS est une situation dont le système de contraintes va favoriser l'émergence du « pas en avant ». Il ne s'agit pas d'une tâche localisée, qui comme par magie pourrait faire émerger une solution efficace en quelques essais, mais d'une situation qui doit être pratiquée sur la durée (à l'aune d'une séquence d'enseignement). La forme de pratique scolaire doit « permettre à l'élève de « triturer longtemps » l'objet d'enseignement et les contenus qui ont été choisis » (Coston et al., 2013). Et dans ce sens la situation d'évaluation terminale n'a pas à être autre chose que la forme de pratique scolaire, étudiée durant l'ensemble de la séquence.

Ce qui permet de relativiser l'idée couramment avancée selon laquelle une situation d'évaluation des compétences devrait nécessairement être inédite (voir par exemple Crahay, 2006). Ceci peut avoir du sens dans certaines disciplines, par exemple en mathématiques où l'agencement d'une démonstration aurait pu être entièrement dévoilé dans un exercice précédent. Mais dans le cadre des situations communément pratiquées en EPS, cette injonction est totalement hors de propos : une situation complexe ne se reproduit jamais à l'identique, et la compétence réside justement dans la capacité à s'adapter à ces évolutions perpétuelles.

Enfin la situation d'évaluation terminale n'a pas à être plus difficile que les situations d'apprentissage. Au contraire, on devrait pouvoir aller plus loin en situation d'apprentissage qu'en situation d'évaluation.

Un rapport nouveau au temps

Traditionnellement, l'évaluation est concentrée sur une épreuve ponctuelle, en fin de séquence d'enseignement, et strictement délimitée dans le temps. La définition même des compétences amène à relativiser ce dogme scolaire. Une évaluation pertinente des compétences ne peut se baser sur une seule session : parce que la situation est complexe et la compétence imprévisible, l'échec est toujours possible. Perrenoud (2004) retient ce principe comme essentiel : « *Il n'y a aucune contrainte de temps*

fixée arbitrairement lors de l'évaluation des compétences ». Dans un rapport de l'Observatoire des pratiques en éducation prioritaire (2021), on retrouve ce commentaire d'un enseignant d'EPS, critiqué par une enseignante d'une autre discipline, surprise que ses élèves obtiennent de si bonnes notes : « *Comment se fait-il que tes élèves aient tous 20/20 ? À croire que c'est un peu l'école des fans, quoi. Et je lui explique que j'avais mis en place un objectif. Ils avaient autant de passages qu'ils voulaient, mais l'objectif à la fin, il fallait l'avoir atteint. C'était en danse. Ils avaient une note. Ils avaient autant d'essais jusqu'à la note finale, 20* » (p. 56).

Au-delà de l'évaluation terminale, l'évaluation des compétences participe essentiellement au processus de leur construction progressive : « *L'évaluation doit contribuer à ce que les [élèves] développent davantage leurs compétences* » (Perrenoud, 2004). Wiggins (1989) affirme ainsi que ce qu'il appelle les *tests authentiques* « *mettent l'accent sur les progrès des élèves vers la maîtrise* », ils « *se reproduisent et valent la peine d'être pratiqués, répétés et repris* ». Cette logique est clairement mise en œuvre dans les productions du groupe EPIC, pour lequel il s'agit avant tout de « *renseigner conjointement l'enseignant et les élèves sur les progrès qui s'opèrent au fil des leçons* » (Röösl, 2017). L'auteur ajoute : « *L'évaluation est davantage envisagée comme un message et moins comme une simple mesure. Nos productions visent ainsi à permettre aux élèves de s'auto et se co-évaluer pour mieux apprendre [...] et déterminer leurs apprentissages futurs (fonction d'aiguillage)* » (Röösl, 2017).

On comprend, lorsque l'on progresse dans ces principes, que si l'on réfléchit en termes d'évaluation des compétences, on quitte nécessairement le schéma traditionnel du cycle, succession de situations d'apprentissage, se terminant par une situation d'évaluation, plus complexe, dans laquelle les élèves seraient supposés être capables de mobiliser de manière conjointe les savoir-faire disparates acquis lors des leçons qui ont précédé. L'évaluation des compétences ne peut être conçue comme un exercice ponctuel, sanctionnant au terme d'une séquence d'enseignement le niveau acquis par les élèves. Elle est nécessairement continue, distribuée, au fil de la séquence d'enseignement. C'est dans ce sens que les distinctions « classiques », entre évaluations diagnostique, formative, formatrice, sommative, certificative, me semblent clairement dépassées dans le cadre d'une pédagogie des compétences. L'évaluation, sans changer de nature, doit être en mesure de remplir toutes ces fonctions.

Les critères d'évaluation de la compétence

Les auteurs qui ont réfléchi de manière générique à l'évaluation des compétences proposent en général de recourir à une évaluation multi-critériée. Wiggins (1989) affirme que « *les tests authentiques utilisent des systèmes de notation à multiples facettes au lieu d'une seule note globale. Les nombreuses variables de performances complexes sont dissociées dans le jugement* ». Gérard et Braibant (2004) évoquent les critères génériques suivants : (1) *la pertinence*, c'est-à-dire l'adéquation de la production à la situation proposée, (2) *l'utilisation correcte des savoir-faire* : même si la production manque de pertinence, est-elle pour autant bien ou mal réalisée ? (3) *la cohérence de la démarche*, même si elle n'est pas pertinente, et (4) *La complétude*, c'est-à-dire le caractère complet de la production.

Même si j'en conçois l'intérêt dans certaines disciplines scolaires, je ne suis pas certain que ces critères puissent fonctionner dans toutes les situations. Les compétences restent très spécifiques aux activités dans lesquelles elles s'expriment, et ces activités présentent elles-mêmes des différences marquées. En ce qui concerne l'EPS, et vu la diversité des APSA proposées à l'étude des élèves, il me semble que l'élaboration d'une grille critériée unique, transversale à l'ensemble des pratiques, serait une entreprise peu pertinente. On peut évoquer à ce niveau le cadre d'évaluation proposé en 2019 pour le baccalauréat, qui ressemble davantage à une procédure de normalisation nationale qu'à une avancée pédagogie majeure (Delignières, 2020b).

D'une manière générale, les propositions actuelles pour l'évaluation des compétences vont dans le sens de la construction d'outils multi-critériés, spécifiques aux APSA enseignées et aux objets d'enseignement sélectionnés. Nous y reviendrons dans la partie suivante, notamment à propos des « scores parlants ».

Compétence et performance

Les collègues du CEDREPS ont fait une proposition intéressante, qui se démarque de la multiplication des critères que l'on voit fleurir ici et là. J'ai évoqué précédemment la logique du ciblage d'un objet d'enseignement, et de son insertion dans une FPS, destinée à favoriser l'émergence de la compétence correspondante. Ce modèle amène les auteurs à une conception assez novatrice de l'évaluation des compétences : « *La réussite dans la FPS doit dépendre directement de l'acquisition de l'objet d'enseignement. L'évaluation qui en découle prend alors en compte la performance mais qui est alors « une performance saturée en compétence ».* La construction de la forme de pratique scolaire doit donc offrir aux élèves le sens de l'activité en puisant ses objets dans la culture de l'activité et ne pas permettre la réussite de l'élève sans l'acquisition de la compétence recherchée » ([Coston et al., 2013](#)).

On ne peut que saluer la rigueur théorique de cette proposition, mais aussi reconnaître qu'elle impose de fortes exigences sur la conception des FPS, pour qu'en effet on puisse affirmer que la performance produite est « saturée en compétence ». A une époque où il est de bon ton de refouler l'idée de « performance », considérée comme résidu d'une conception surannée de l'EPS, cette ligne de réflexion me paraît courageuse et plutôt stimulante.

Quelques réserves, cependant...

On peut néanmoins se demander si dans le cadre théorique des FPS, cette évaluation de la performance rend réellement compte d'une compétence, telle que définie précédemment. C'est essentiellement le rapport à la culture, tel que suggéré dans les textes du CEDREPS, qui mérite d'être questionné. Dans un des articles princeps de cette thématique, Dhelemmes ([2004](#)) invite les enseignants à cultiver une certaine « irrévérence » vis-à-vis des pratiques de référence dans la construction des formes de pratiques scolaires. En d'autres termes, il admet que ces formes de pratique puissent légitimement prendre quelques distances vis-à-vis des modalités réelles de pratique des APSA. Becker et Kraemer ([2021](#)) dénoncent cette mise à l'écart de la culture, et le déni que les APSA puissent constituer par elles-mêmes l'objet d'étude de l'EPS. Ils y voient une rupture avec les conceptions culturalistes de l'éducation, concevant l'appropriation des objets culturels comme une condition essentielle à l'émancipation des individus.

On peut s'interroger aussi sur le postulat selon lequel les formes de pratiques scolaires seraient essentiellement construites sur la base du ciblage d'objets d'enseignement, principalement technico-tactiques. On peut craindre que ce ciblage s'opère au prix d'un affadissement de la complexité des dimensions sociales des pratiques ([Delignières, 2021](#)). Pour reprendre le modèle que j'ai proposé à propos de la référence culturelle ([Delignières, 2019b](#)), l'approche du CEDREPS se situe clairement dans une approche *savante* de la culture : on sélectionne des savoirs localisés, et l'on construit des formes de pratiques scolaires visant à en assurer l'appropriation. Je proposais dans ce texte une autre conception, *praxique*, de la culture, reconnaissant qu'au-delà des savoir-faire techniques, les modes de sociabilité, le fonctionnement des collectifs, les méthodes d'entraînement et de préparation, devaient être considérés également comme constitutifs de la culture. Même si les collègues du CEDREPS évoquent explicitement l'intégration des « composantes méthodologiques et sociales » dans les FPS (voir notamment [Dhelemmes, 2004](#)), ces dimensions semblent davantage renvoyer à un formalisme scolaire qu'à une prise en compte de la réalité sociale des pratiques de référence.

On conçoit dès lors que le principe avancé par Maurice Portes, selon lequel l'EPS se devait « *d'offrir aux élèves les conditions d'une réelle tranche de vie d'athlète, de combattant, de handballeur, de grimpeur, de gymnaste, etc.* » ([Portes, 1999](#)), principe que j'estime essentiel à la construction de compétences authentiques, est peu compatible avec une mise à distance des APSA, en tant qu'objets culturels. Serge Testevuide ([2020](#)) reconnaît d'ailleurs explicitement que « *la « tranche de vie » est comprise non pas comme un moment d'une vie de handballeur avec toutes ses dimensions, mais comme une « tranche » de son activité au sens d'un aspect particulier* ». Kraemer et Becker ([2022](#)) critiquent sans équivoque cette évolution : « *la discipline, au lieu de faire vivre des « tranches de vie »*

aux élèves, culturellement et socialement significatives, ne leur offre qu'une vie en tranches, les privant de la richesse et du réel humain contenu dans les APSA ».

Le quantitatif et le qualitatif dans l'évaluation des compétences

L'évaluation en EPS a été très précocement marquée par un souci d'objectivité et de quantification. Dans l'entre-deux-guerres, les médecins ont ainsi proposé des barèmes quantitatifs, pour évaluer les capacités physiques et déterminer les normes de santé que l'Éducation Physique devait établir chez les élèves ([Delignières et Duret, 1989](#)). Les années 60 ont été marquées par l'exploitation des tables Letessier. Les outils proposés dans les années 90 pour dépasser l'évaluation brute de la performance (par exemple les nomogrammes), reposaient également essentiellement sur des bases quantitatives (voir par exemple [Gréhaigne et Roche, 1993](#)). Notons que ces outils d'évaluation étaient à l'époque avant tout pensés dans le cadre d'une visée de répartition des élèves dans des groupes de besoins (pour l'approche hygiéniste de l'entre-deux-guerres), soit dans une optique d'évaluation certificative, pour hiérarchiser les élèves de manière objective ([Gréhaigne et Roche, 1993](#)).

Si des propositions alternatives, plus qualitatives, ont également émergé, sous forme de grilles de niveaux d'habileté ou de niveau de jeu (voir par exemple [Durand, 1993](#), ou [Laroche et Pion, 1995](#)), on assiste à l'heure actuelle à un retour en force des indicateurs quantitatifs, notamment au travers des « scores parlants ». Pour avoir une idée des origines de cette démarche, on pourra consulter les contributions princeps d'Ubaldi et Philippon ([2003](#)), de Bellard ([2006](#)) et de Mascret ([2009](#)). Les propositions les plus abouties sont présentées par le groupe EPIC ([Fayaubost, Gibon & Rossi, 2020](#)).

Il existe de nombreuses versions de scores parlants, le plus connu étant sans doute le 100/10/1, dans lequel trois indicateurs de performance produisent respectivement les centaines, les dizaines et les unités du score final. Ce type d'indicateur, multidimensionnel, renseigne l'élève (et aussi l'enseignant) non seulement sur sa performance, mais aussi sur la manière dont elle a été produite. Si son usage est répété au cours de la séquence, dans le cadre d'une forme de pratique stable, il révèle et situe les progrès réalisés et permet d'identifier les déficits à combler.

Ces indicateurs permettent aussi une dévolution complète de l'évaluation aux élèves, pour peu qu'on les ait formés de manière adéquate à la nature des observations à réaliser. L'évaluation devient d'ailleurs parfois un lieu majeur d'opérationnalisation des « compétences méthodologiques et sociales », un phénomène qui peut être interrogé... ([Delignières, 2022](#)).

Si ces propositions ouvrent des pistes intéressantes, il est peut-être utile d'en évoquer quelques limites. La démarche du groupe EPIC a été explicitement reliée à la proposition du CEDREPS sur les formes de pratique scolaire : le score parlant est notamment censé cibler le « pas en avant » visé par la construction de la FPS ([Röösl, 2017](#)). Coston et Ubaldi ([2007](#)) évoquent cependant le danger d'une conception « magique » de la FPS, qui pourrait laisser penser qu'elle pourrait représenter une solution idéale pour telle APSA et tel niveau de classe, dispensant l'enseignant d'une maîtrise technologique de l'APSA. De même, la conception d'un score parlant ne doit pas céder à la facilité : il doit réellement rendre compte du « pas en avant » ciblé dans la FPS, ce qui reste un exercice délicat.

Maurice Portes ([1996](#)) avait alerté sur certains effets pervers de l'évaluation, dans le domaine de sa spécialité de prédilection, le handball. Ces réflexions méritent d'être analysées dans le cadre du développement actuel des scores parlants. L'auteur constate que toutes les tentatives pour construire des outils et dispositifs d'évaluation en handball, et d'une manière générale dans toutes les activités collectives d'opposition, semblent guidées par deux préoccupations, l'objectivité et la faisabilité. Préoccupations qu'il considère légitimes, mais qui amènent logiquement les enseignants à donner une base quantitative à l'évaluation, et donc à recourir à des critères qui permettent aisément cette quantification.

Dans ce cadre, il note que le volume de jeu est un indicateur souvent privilégié : « *déterminé par le nombre d'entrées en possession du ballon, [il] offre toutes les garanties de ce point de vue : il est omniprésent dans les propositions de nomogrammes, associé le plus souvent à un "indice d'efficacité" »*

(Portes, 1996). L'exploitation exclusive de ce critère quantitatif entraîne cependant pour l'auteur deux conséquences importantes. La première est « *le renoncement total ou partiel à la prise en compte de la pertinence des choix tactiques du joueur, celle-ci devant impliquer une appréciation subjective de l'évaluateur et exigeant de lui une bonne compréhension des logiques des rapports d'opposition... mais constituant le facteur fondamental de la performance en jeu* » (Portes, 1996). La seconde est « *la survalorisation, sans justification, de l'activité offensive, et plus précisément des actions sur le ballon ou à proximité de celui-ci, au détriment de tous les autres volets de l'activité du handballeur* ».

Le verdict de l'auteur est sans équivoque : « *on assiste à l'émergence d'un "handball de l'évaluation" qui induira des contenus d'enseignement largement amputés du fond culturel de cette spécialité* » (Portes, 1996). Cet argument est évidemment à mettre en rapport avec les constats que l'auteur fait sur la « méconnaissance technologique » des enseignants vis-à-vis des APSA (Portes, 2006). La quantification constitue aussi une béquille, pour des enseignants ne maîtrisant que modérément les subtilités de l'activité, pour mettre en place des dispositifs d'évaluation, et réguler le processus d'apprentissage. Mais elle peut entraîner aussi un ciblage sur certains objets d'enseignement dont la pertinence peut être questionnée. Maurice Portes défend une approche plus qualitative de l'évaluation, basée sur l'observation des intentions tactiques des joueurs et des équipes. Il est clair qu'il ne situe dès lors plus le rôle de l'enseignant comme celui d'un didacticien centré sur l'acquisition de savoirs prédéterminés, mais plutôt comme un accompagnateur, capable de guider les élèves dans la construction de leurs compétences (sur cette question des postures enseignantes, voir Delignières, 2016). Philippe Perrenoud avait clairement évoqué ce problème dans un article consacré à l'évaluation des compétences : « *Plus un enseignant se sent et fonctionne comme un formateur, un coach, un accompagnateur, plus il développe une expertise d'observateur. Cette dernière n'a rien à voir avec l'expertise d'un constructeur d'épreuves, de correcteur de travaux écrits, de compteurs de points ou d'erreurs et de fabricant de barèmes...* » (Perrenoud, 2004).

Je me suis interrogé par exemple sur le score parlant proposé par Stéphane Bellard (2006), en volley-ball. L'objectif de l'enseignant est clair, il s'agit de favoriser la construction collective de l'attaque. L'outil d'évaluation, de forme 1000/100/10/1, hiérarchise logiquement les actions réalisées en fonction du nombre de touches de balles qui a précédé le marquage du point.

Tableau 1. Exemple en volley-ball. Objectif poursuivi : construction collective de l'attaque.

Critères d'observation	Point marqué en 3 touches	Point marqué en 2 touches	Point marqué en 1 touche	Point marqué sur faute adverse
Unité de comptage des pt	1000	100	10	1
Score de l'équipe A	2	1	4	3
Score de l'équipe B		2	2	2

Ce score parlant présente une telle évidence que l'on pourrait être tenté de l'exploiter à tous niveaux de pratique. Il est clair cependant qu'il tend à promouvoir une forme tactique idéale (défenseur-passeur-attaquant), au détriment peut-être d'intentions tactiques plus pertinentes, à certains moments ou face à certaines équipes. Rappelons que la compétence réside dans une exploitation stratégique, située, des savoir-faire, quitte à produire des solutions inattendues, mais certainement pas dans la production systématique de savoir-faire idéaux. Les élèves pourraient être davantage pilotés par le score que par les intentions tactiques et l'intelligence du jeu.

Et encore ces arguments ne concernent que des activités qui se prêtent aisément à la quantification. Dans d'autres activités, la mesure a beaucoup moins de sens (les pratiques artistiques notamment). La nécessaire scolarisation de ses pratiques et de leurs procédures d'évaluation ne risquent-elles pas

d'introduire des ruptures complètes avec leur fonds culturel ? Amélie Broudisou évoque cette question dans un article consacré à la scolarisation du yoga : « *est-il seulement évaluable selon des critères scolaires sans remettre en cause ses présupposés culturels ?* » ([Broudisou, 2022](#)). Rappelons que le yoga est dorénavant dans la liste nationale des activités du champ d'apprentissage n°5 au lycée, et supposé supporter l'attribution de notes classantes au baccalauréat...

Ce texte suggère que l'évaluation évolue, tant dans ses fonctions que dans ses modalités. J'ai tenté de montrer les contraintes auxquelles elle était toujours soumise, dans une École tiraillée entre tradition méritocratique et velléités démocratiques. J'ai aussi évoqué les avancées majeures suscitées par la pédagogie des compétences, qui me semblent en effet susceptibles de modifier singulièrement le regard porté par les enseignants sur l'évaluation.

J'attire l'attention du lecteur sur les arguments développés dans la dernière partie, à propos des dérives potentielles d'une évaluation essentiellement basée sur des indicateurs quantitatifs. Même si l'on peut en comprendre l'intérêt, notamment en termes de suivi et de guidage de l'activité des élèves, on peut craindre que parfois les contenus d'enseignement soient davantage pilotés par les procédures d'évaluation que par des préoccupations éducatives. ***Il me semble que parfois on ne se pose plus guère la question d'évaluer ce que l'on a enseigné, mais plutôt d'enseigner ce que l'on est en mesure d'évaluer.***

Références

- Antibi, A. (2003). *La constante macabre ou comment a-t-on découragé des générations d'élèves*. Éditions Math'adore.
- Becker, A. & Kraemer, D. (2021). [A propos de la « forme de pratique scolaire » - De l'irrévérence à l'absence de révérence](#). *ContrePied, Hors-série n° 29*, 48-50.
- Bellard, S. (2006). [Évaluation : quand le quantitatif révèle le qualitatif](#). *Revue EPS*, 321, 31-35.
- Broudisou, A. (2022). [Yoga, intérêts et limites](#). Site EPS & Société, 3 juin 2022.
- Coston, A. & Ubaldi, J.L. (2007). [Une EPS malade de ses non-choix](#). *Les cahiers du CEDRE/CEDREPS*, 7, 20-31
- Coston, A., Testevuide, S., et Ubaldi, J.L. (2013). [Compétence et objet d'enseignement : les deux faces d'une même pièce](#). *Les cahiers du CEDRE/CEDREPS*, 13, 173-187.
- Crahay, M. (2006). [Dangers, incertitudes et incomplétude de la logique de la compétence en éducation](#). *Revue Française de Pédagogie*, 154, 97-110.
- Delaunay, M., & Pineau, C. (1989). [Un programme, la leçon, le cycle en EPS](#). *Revue EPS*, 217, 49-52.
- Delignières, D. (2009). [Enseigner des compétences](#). In *Complexité et compétences. Un itinéraire théorique en Éducation Physique* (chapitre 7). Paris : Éditions revue EPS.
- Delignières, D. (2016). [Les postures professionnelles de l'enseignant d'EPS: le scientifique, le didacticien, le coach](#). Site personnel, 17 avril 2016.
- Delignières, D. (2019a). [L'EPS, une éducation à la complexité](#). Site personnel, 10 novembre 2019
- Delignières, D. (2019b). [Culture, enseignement, et Éducation Physique et Sportive](#). Site personnel, 5 février 2019.
- Delignières, D. (2020a). [EPS et Sport : Faut-il enlever le « S » de « EPS » ?](#) Site personnel, le 27 novembre 2020.

- Delignières, D. (2020b). [L'évaluation de l'EPS au baccalauréat : trois pas en arrière...](#). Site personnel, 28 février 2020.
- Delignières, D. (2021). [Formes de pratique scolaire, formes scolaires de pratique Usages et mésusages des concepts](#). Site personnel, 23 octobre 2021.
- Delignières, D. (2022). [« Compétences méthodologiques et sociales » et « rôles sociaux ». Quelques arguments critiques](#). Site personnel, 12 septembre 2022.
- Delignières, D. & Duret, P. (1989). [Valeur physique et grandeur morale](#). *Échanges et Controverses*, 2, 155-200.
- Dhellemmes, R. (2004). [Des situations de « Ré\(v\)férence » aux formes de pratique scolaire](#). *Les cahiers du CEDRE/CEDREPS*, 4, 67-78.
- Durand, P. (1994). [Le saut de cheval](#). *Revue EPS*, 245, 13-15.
- Duru-Bellat, M. (2022). [Le vertige des notes](#). Alternatives économiques, 24 juin 2022.
- Fayaubost, R., Gibon, J. & Rossi, D. (2020). [Une démarche EPIC pour apprendre et enseigner en EPS](#). Éditions AEEPS.
- Gérard, F.-M. (2007). [La complexité d'une évaluation des compétences à travers des situations complexes : nécessités théoriques et exigences du terrain](#). *Actes du Colloque international « Logique de compétences et développement curriculaire : débats, perspectives et alternative pour les systèmes éducatifs »*. Montréal, 26 et 27 avril 2007.
- Gérard, F.-M. & Braibant, J.-M. (2004). [Activités de structuration et activités fonctionnelles, même combat ? Le cas de l'apprentissage de la compétence en lecture à l'école primaire](#). *Français 2000*, 190-191, 24-38.
- Gréhaigne, J.F., et Roche, J. (1993). [Évaluation au baccalauréat](#). *Revue EPS*, 240, 80-83.
- Hadji, C. (2022). [Faut-il continuer à noter les élèves ?](#) The conversation, 19 septembre 2022.
- Hirtt, N. (2011). [L'approche par compétences : une mystification pédagogique](#). Site Démocratisation scolaire, 25 janvier 2011.
- Kraemer, D. & Becker, A. (2022). [À propos de Forme de pratique scolaire](#). *ContrePied, Hors-série n° 30*, 46-48.
- Laroche, J.Y. et Pion, I. (1995). [Des grilles comportementales aux indices chiffrés](#). *Revue EPS*, 252, 35-37.
- Mascret, N. (2009). [Compétition et compétence : l'utilisation du score en badminton](#). *Contrepied*, 23, 21-23.
- Observatoire des pratiques en éducation prioritaire (2021). [L'EPS et le sport en éducation prioritaire](#). Les rapports de l'académie de Créteil.
- Perrenoud, P. (2004). [Évaluer des compétences](#). *Éducateur, n° spécial « La note en pleine évaluation »*, 8-11.
- Perrenoud, Ph. (2005). [L'évaluation des élèves : outil de pilotage ou pare angoisse ?](#) *Cahiers pédagogiques*, 438, 14-16.
- Portes, M. (1996). [1987-1995. Où va donc « l'appropriation culturelle » en EPS ?](#) *Revue Hyper*, 192, 15-17.
- Portes, M. (1999). [Nouveaux programmes d'EPS et rénovation des pratiques](#). *Revue EPS*, 278, 27-30.
- Portes, M. (2006). [Les promesses d'une pratique studieuse du handball](#). *Les cahiers du CEDRE/CEDREPS*, 6, 66-73.

- Röösli, W. (2017). [L'évaluation par indicateurs de compétence avec EPIC](#). Café pédagogique, 9 février 2017.
- SNEP (2005). [S'il te plait écris-moi un programme](#). Texte publié sur le site SNEP-FSU, daté du 26 juin 2005.
- Testevuide, S. (2020). [Introduction](#). *Les Cahiers du CEDREPS*, 17, 7-11.
- Ubaldi, J.-L. & Philippon, S. (2003). [Quelle EPS ? Une illustration en basket-ball](#). *Revue EPS*, 299, 67-72.
- Ubaldi, J.L., Coston, A., Coltice, M., Philippon, S. (2006). [Cibler, habiller, intervenir pour permettre à nos élèves d'apprendre en EPS](#). *Les cahiers du CEDREPS*, n°6, p. 7-31.
- Wiggins, G. (1989). [A true test: Toward more authentic and equitable assessment](#). *Phi Delta Kappa*, 70, 703-714.