

De la délicate question des ressources constitutives des compétences en éducation physique et de leur évaluation

Benoît Lenzen

Université de Genève, Faculté de psychologie et des sciences de l'éducation & Institut universitaire de formation des enseignants

Benoit.Lenzen@unige.ch

Bernard Poussin

Université de Genève, Institut universitaire de formation des enseignants

Bernard.Poussin@unige.ch

Hervé Dénervaud

Université de Genève, Institut universitaire de formation des enseignants

Herve.Denervaud@edu.ge.ch

Adrián Cordoba

Université de Genève, Faculté de psychologie et des sciences de l'éducation & Institut universitaire de formation des enseignants

Adrian.Cordoba@unige.ch

Daniel Deriaz

Université de Genève, Institut universitaire de formation des enseignants

Daniel.Deriaz@unige.ch

Résumé

Formaliser des compétences en éducation physique (EP) et concevoir des outils pour évaluer celles-ci constituent une véritable difficulté professionnelle. A partir d'une revue de la littérature, nous avons synthétisé les principes à adopter pour formaliser et évaluer des compétences en EP. A l'aide de cette grille d'analyse, nous avons analysé les prescriptions genevoises pour l'enseignement obligatoire, afin de mettre en évidence les solutions trouvées par leurs auteurs pour dépasser cette difficulté ainsi que les écueils restant à surmonter. Les prescriptions pour l'enseignement primaire et l'enseignement secondaire ont en commun de prescrire d'évaluer les élèves dans une situation complexe et authentique. Elles se distinguent par le fait que les premières prescrivent aux enseignants primaires d'évaluer des ressources qui renvoient à des situations plutôt qu'à des compétences, tandis que les secondes prescrivent aux enseignants secondaires d'évaluer simultanément des compétences dans l'action et autour de l'action et leurs ressources constitutives.

Mots clés

Education physique ; approche par compétences ; évaluation

Summary

Professionals have difficulties to define competences to be taught and assessed in physical education (PE). Based on a literature review, we synthesised the principles to be observed when defining and assessing competencies in PE. By means of this analysis grid, we analysed Genevan official instructions for compulsory school, in order to highlight solutions to get over these difficulties and obstacles to be overcome. What official instructions for primary and secondary schools have in common is that they both recommend to assess students in a complex and authentic situation. What distinguishes them is that official instructions for primary school ask teachers to assess skills which are related to situations rather than to competencies, while official instructions for secondary school ask teachers to simultaneously assess competencies in and around action and their related skills.

Keywords

Physical education ; competency-based approach ; assessment

Pour citer cet article : Lenzen, B., Poussin, B., Déneraud, H., Cordoba, A. & Deriaz, D. (2017). De la délicate question des ressources constitutives des compétences en éducation physique et de leur évaluation. *Evaluer. Journal international de Recherche en Education et Formation*, 2(1), pp. 47-67.

1. Introduction

Depuis une trentaine d'années, l'approche par compétences n'en finit pas d'alimenter débats et controverses dans les milieux scolaires et académiques. Parmi les nombreux opposants à cette approche, Crahay (2006) va jusqu'à proposer de « *renoncer à la doxa de la compétence, de ressusciter la notion de champ conceptuel chère à Vergnaud (1987) et, ainsi outillé sur le plan théorique, de restaurer et de revivifier le disciplinaire* » (p. 171). En tant que didacticiens de l'éducation physique (EP), nous partageons cette volonté de restaurer le disciplinaire, en préservant l'épistémologie propre aux différentes disciplines scolaires (Astolfi, 2008 ; Schneider-Gilot, 2006). Nous étendons d'ailleurs cette vigilance épistémologique aux activités physiques, sportives et artistiques (APSA) qui constituent les principales pratiques sociales de référence de l'EP (Lenzen, Cordoba, Poussin, Dénervaud & Deriaz, 2013). En revanche, le caractère pratique de cette discipline scolaire s'oppose au recours à des modèles exclusivement centrés sur le savoir-texte (Laparra & Margolinas, 2010 ; Rey, 2007) pour penser l'enseignement et l'évaluation en EP. A la suite d'autres auteurs (e.g., Carette, 2007 ; Ubaldi, 2005), nous proposons donc de poursuivre la réflexion sur la notion de compétence et ses implications sur les questions d'enseignement, d'apprentissage et d'évaluation, à partir d'un point de vue didactique et dans la discipline scolaire EP pour ce qui nous concerne.

En Suisse romande, le curriculum officiel pour l'enseignement obligatoire (Plan d'études romand, PER), entré en vigueur en 2011, s'inscrit résolument dans l'approche par compétences (CIIP, 2010-2016 ; Nanchen, 2008 ; Rey, 2008). Dans le canton de Genève, le Département de l'instruction publique (DIP) a produit des documents de liaison pour contextualiser les prescriptions du PER (DGCO, 2011-2012 ; DGCO, 2011-2013 ; DGEO, 2014). En effet, si le contenu de l'enseignement obligatoire est désormais régi au niveau intercantonal, les prescriptions en matière d'évaluation restent de la responsabilité des cantons. Pour ce qui concerne l'EP, ces prescriptions diffèrent d'un canton romand à l'autre (Allain, Deriaz, Voisard & Lentillon-Kaestner, 2016).

Formaliser des compétences à évaluer et concevoir des instruments pour ce faire constituent une réelle difficulté pour la noosphère et les enseignants en général (Crahay, 2006 ; De Ketele, 2009), et en EP en particulier (Cogérino, 2002 ; Lenzen, Dénervaud & Poussin, 2012 ; López-Pastor, Kirk, Lorente-Catalán, MacPhail & MacDonald, 2013 ; Mougnot, 2016). Dans cette contribution, à partir d'une revue de la littérature non exhaustive et sélectivement orientée par notre projet, nous synthétiserons les principes à adopter pour formaliser et évaluer des compétences en EP. A l'aide de cette grille d'analyse, nous analyserons ensuite les prescriptions genevoises pour l'enseignement obligatoire, afin de mettre en évidence les solutions trouvées par leurs auteurs pour dépasser cette difficulté ainsi que les écueils restant à surmonter.

2. Cadre théorique

2.1 La notion de compétence et son utilisation en éducation physique

Au-delà des controverses autour de la notion de compétence, il nous semble qu'un consensus peut s'établir autour de définitions proches de celle-ci : « *une compétence repose sur la mobilisation, l'intégration et la mise en réseau d'une diversité de ressources [internes et externes] [...] Cette mobilisation des ressources s'effectue dans une situation donnée, dans le but d'agir : la compétence est nécessairement située ; pour autant, elle s'exerce dans une diversité de situations, à travers un processus d'adaptation et pas seulement de reproduction de mécanisme [...] il est difficile de définir la notion de compétence indépendamment de son cadre d'apprentissage et des situations prévues pour son évaluation* » (iGen, 2007, pp. 11-12, cité par

Mottier Lopez & Tessaro, 2010, p. 32). Développer et évaluer des compétences à l'école supposent un important travail curriculaire, de la part de la noosphère (Chevallard, 1985) comme des enseignants. Dans des disciplines praxiques comme l'EP, il convient en effet de (a) repérer et expliciter des pratiques sociales qui ne sont pas immédiatement « lisibles », (b) identifier les compétences en jeu dans ces pratiques, et (c) identifier de quoi elles sont faites, comment elles fonctionnent (Perrenoud, 1998).

En EP, la deuxième étape, consistant à identifier les compétences en jeu dans les APSA, n'est pas chose aisée, en raison de l'amalgame entre les diverses formulations utilisées pour désigner ce qui s'enseigne en EP (conduite motrice, comportement caractéristique, compétence, contenu d'enseignement, savoirs pratique, théorique, technique, etc.), et entre le sens, les problèmes et les techniques de l'APSA, les buts de l'élève et les objectifs de l'enseignant (Ubaldi, 2005). Pour pallier ce qui constitue une véritable difficulté professionnelle, cet auteur préconise une approche pragmatique pour fixer les contours de la notion de compétence en EP, considérant que « *toute action finalisée, contextualisée (en cernant des contraintes précises), efficace (qui produit des effets), observable, voire évaluable, acquise par apprentissage et reproductible dans des contextes proches, constitue une compétence* » (p. 66), et que le critère de globalité, certes imprécis, est néanmoins fondamental dans la mesure où « *il permet de penser la difficile question de la différence entre la compétence et les contenus moteurs présidant à son acquisition* » (p. 66). Ubaldi (2005) propose finalement de formaliser les compétences tactico-techniques dans chaque APSA sous la forme d'un verbe d'action suivi du but visé, par exemple « se démarquer en soutien pour aider son partenaire à conserver le ballon » en basket-ball ou « frapper fort pour rompre l'échange » en tennis de table.

Dans la troisième étape, Perrenoud (1998) distingue les *ressources cognitives mobilisées* (informations, théories, concepts, rapport au savoir, méthodes, techniques, procédures, habiletés, attitudes) et les *schèmes opératoires* « *qui permettent, en temps réel, la mobilisation efficace des ressources cognitives* » (p. 506). De son côté, Le Boterf (1999, p. 28, cité par Coulet, 2011) considère « *un double équipement de ressources : des ressources incorporées (connaissances, savoir-faire, qualités personnelles, expérience...) et des réseaux de ressources de son environnement (réseaux professionnels, réseaux documentaires, banques de données...)* ».

Selon Coulet (2011), des « modèles » de ce type, exprimés en termes de « savoirs, savoir-faire et savoir-être » ou reposant sur l'idée d'une combinatoire de ressources, ne permettent pas une description dynamique des processus en jeu. Pour produire une telle description, cet auteur affirme la nécessité de considérer entre autres que (a) il y a lieu de marquer une claire distinction entre *tâche* (but à atteindre et conditions dans lesquelles il doit être atteint) et *activité* (ce qui est mis en œuvre par le sujet pour exécuter la tâche), (b) la tâche que traite le sujet n'est pas forcément celle qui lui est prescrite, (c) l'activité étant organisée de façon hiérarchique, il convient de choisir le bon niveau de description en fonction des objectifs poursuivis, (d) un même sujet peut adopter différentes manières de procéder, en fonction des circonstances et/ou de son niveau de compétence, pour accomplir une même tâche, et (e) la plupart des activités humaines se réalisent via un ensemble d'outils matériels ou symboliques, en référence au concept vygotkien d'instrument psychologique dont « *le langage, les diverses formes de comptage et de calcul, les moyens mnémotechniques, les symboles algébriques, les œuvres d'art, l'écriture, les schémas, les diagrammes, les cartes, les plans, tous les signes possibles, etc.* » (Vygotski, 1930/1985, p. 39, cité par Coulet, 2011, pp. 15-16) en sont des exemples.

2.2 Evaluer des compétences

Selon Crahay (2006), l'évaluation des compétences est problématique dès lors que la complexité inédite est érigée en norme de l'approche par compétences. Evaluer des compétences reviendrait par conséquent à évaluer un *savoir-mobiliser*, ce qui consacrerait le retour du *savoir-transférer*, autrement dit de l'intelligence. Adopter pour critère de compétence la résolution de problèmes à la fois complexes et inédits équivaldrait en outre à confronter les élèves à un niveau d'exigence extrêmement élevé, hors de portée de la majorité d'entre eux. Crahay (2006) souligne encore le risque de confusion majeur entre situations d'apprentissage et d'évaluation. De son côté, Allal (2002) joue sur cette confusion en préconisant d'intégrer, dans une perspective d'évaluation formative de type interactif, des modalités d'autoévaluation, d'évaluation mutuelle et de coévaluation dans une situation complexe. Cette méthode donnerait à l'élève l'occasion de prendre conscience de son fonctionnement et d'apprendre à gérer plus efficacement la situation.

Conscient du risque d'augmentation des exigences inhérent à l'approche par compétences, Carette (2007) propose de réunir trois conditions pour évaluer des compétences à l'école : (a) proposer des tâches inédites aux élèves ; (b) proposer des tâches complexes aux élèves et (c) présenter aux élèves des tâches complexes et inédites qui font appel à un ensemble de procédures effectivement apprises. Pour éviter que cette approche se transforme en terreau pour des « pédagogies élitistes », Carette (2007) propose d'y ajouter une quatrième condition consistant à présenter aux élèves des épreuves diagnostiques. Il préconise, dès lors, une procédure d'évaluation en trois phases : (a) l'enseignant propose aux élèves une tâche complexe et inédite que ces derniers devront résoudre sans aucune aide ; (b) l'enseignant propose à nouveau la même tâche aux élèves mais, cette fois, celle-ci est décomposée en tâches élémentaires dont les consignes sont explicites et qui sont présentées dans l'ordre où elles doivent être accomplies pour parvenir à résoudre la tâche globale ; (c) l'enseignant propose enfin aux élèves une série de tâches simples décontextualisées, dont les consignes sont celles qui sont habituellement utilisées lors de l'apprentissage des procédures élémentaires qui leur correspondent.

Enfin, selon Tardif (1993), le paradigme constructiviste dans lequel s'inscrit l'approche par compétences impose la reconsidération des critères traditionnels de validité de l'évaluation et l'établissement d'une nouvelle hiérarchisation de ces critères. Cet auteur en retient quatre : (a) le respect du critère de *validité de construit* implique le recours à des tâches complètes, complexes et signifiantes. Cela demande de plus qu'aucune composante ne soit isolée de la compétence qui lui octroie sa signification. Toute évaluation de composantes pour elles-mêmes est dès lors à exclure car cela créerait une certaine forme d'invalidité de construit ; (b) pour satisfaire au critère de *validité de contenu*, il importe que l'instrument ou la modalité d'évaluation permette une description exhaustive et adéquate d'une compétence et des interactions de ses composantes. Les questions et les situations doivent être orientées sur des compétences et non sur des composantes, ce qui garantirait la congruence avec l'objectif mesuré ; (c) satisfaire au critère de *validité écologique* implique un haut degré de ressemblance entre les conditions dans lesquelles les élèves se trouvent lors de l'évaluation et celles qui existent à l'extérieur de la classe et de l'institution scolaire ; (d) enfin, le critère de *validité déontologique* ou *validité de conséquence* est défini comme le degré d'exactitude avec lequel un instrument ou une modalité d'évaluation détermine les compétences développées.

2.3 Evaluer en éducation physique

Avant les années 70, l'évaluation ne constituait pas une préoccupation pour la discipline scolaire EP. Des formes d'évaluation que López-Pastor et al. (2013) qualifient de « traditionnelles » ont ensuite émergé, dans un contexte fortement marqué par les théories behavioristes de l'apprentissage (Hay, 2006). Les connaissances théoriques étaient évaluées, comme pour les disciplines scolaires plus académiques, par le biais d'examens oraux, de dissertations ou de questionnaires à choix multiples. Les évaluations pratiques incluaient quant à elles l'usage de tests d'habiletés motrices et de condition physique, de tables de performance dans les APSA « énergétiques » telles que la natation et l'athlétisme, et du jugement subjectif de l'enseignant dans les jeux individuels et collectifs. La participation, l'effort, l'attitude, le comportement et le port d'une tenue de sport faisaient fréquemment partie des critères d'évaluation. Des formes d'évaluation « alternatives » (López-Pastor et al., 2013) se sont ensuite développées en appui sur les théories constructivistes de l'apprentissage (Hay, 2006), amenant les élèves à s'impliquer activement dans la résolution de problèmes complexes en mobilisant diverses ressources (prise d'informations, connaissances, habiletés motrices). Souvent qualifiées d'authentiques par opposition aux situations d'évaluation artificiellement décontextualisées (Hay, 2006 ; López-Pastor et al., 2013), ces formes alternatives d'évaluation se caractérisaient par (a) leur intégration dans le processus d'enseignement/apprentissage, (b) le recours aux élèves en tant que coévaluateurs et (c) leur dimension majoritairement formative (Desrosiers, Genet-Volet & Godbout, 1997). Les instruments d'évaluation « alternatifs » sont peu documentés dans la littérature, à l'exception notable de ceux développés pour évaluer les apprentissages des élèves dans les sports collectifs, tels que le Game Performance Assessment Instrument-GPAI (Oslin, Mitchell & Griffin, 1998) et le Team Sport Assessment Procedure-TSAP (Gréhaigne, Godbout & Bouthier, 1997).

Selon Hay et Penney (2009), quatre conditions doivent être remplies pour une évaluation efficace en EP : (a) l'évaluation doit être pensée pour l'apprentissage (*assessment for learning*), autrement dit formative ; (b) l'évaluation doit être authentique et, si possible, intégrée. Cette condition renvoie au nécessaire alignement entre curriculum, pédagogie et évaluation, ces trois dimensions étant à la base d'une « EP de qualité » (Penney, Brooker, Hay & Gillespie, 2009) ; (c) l'évaluation doit être valide. Cette condition repose sur différents critères, proches de ceux retenus par Tardif (1993) exposés plus haut, et (d) l'évaluation doit être socialement juste.

En France, les situations d'évaluation en EP proposées par les référentiels actuels s'inscrivent partiellement dans l'approche par compétences (Mougenot, 2016). Elles reflètent bien la complexité de l'activité de référence et prennent en compte les rôles sociaux constitutifs de celle-ci (Mascret & Rey, 2011) : observateurs, arbitres, juges, entraîneurs, pareurs, etc. En revanche, le caractère souvent non inédit des situations pour les élèves témoigne d'une difficulté en EP à rencontrer pleinement les « standards » de l'approche par compétences (Mougenot, 2016). Cette difficulté tient en partie, selon nous, à l'épistémologie des APSA. Dans certaines APSA, telles que les sports collectifs, le plus important est de faire en sorte que les joueurs soient capables de gérer d'une façon optimale le désordre (Gréhaigne, 1989 ; Gréhaigne & Godbout, 1995 ; Villepreux, 1987, cités par Gréhaigne, Marle & Zerai, 2013). En ce sens, la complexité et l'inédit sont intrinsèquement constitutifs d'une situation d'évaluation authentique dans ces APSA en EP. Mais, dans d'autres APSA comme la gymnastique ou la danse, la préparation des sportifs et/ou artistes à l'épreuve (concours ou spectacle) consiste précisément à faire en sorte qu'ils maîtrisent les enchaînements ou les chorégraphies qui, s'ils ne sont pas intégralement connus au début de la préparation, le sont

en tout cas avant de les présenter devant les juges et/ou le public. Dans ces APSA en EP, vouloir à tout prix exposer les élèves à des tâches inédites en guise d'évaluation constituerait un contresens épistémologique.

Si le modèle d'évaluation proposé par Carette (2007) pour évaluer des compétences à l'école semble constituer une piste intéressante pour lutter contre l'échec scolaire, il est, selon nous, peu applicable en EP en raison de la spécificité de l'activité motrice dans la définition du milieu didactique en EP. Celui-ci est difficilement modélisable indépendamment de ce que vont y produire les élèves et les solutions des problèmes auxquels il confronte les élèves sont rarement canoniques (Amade-Escot, 2005 ; Le Paven, Roesslé, Roncin, Loquet & Léziart, 2007). En EP, il n'est souvent pas possible de découper une tâche complexe, telle qu'un match de sport collectif par exemple, en tâches élémentaires qui, présentées dans un ordre chronologique avec des consignes appropriées (deuxième phase du modèle de Carette), aboutiraient in fine à la réalisation de la tâche globale. En effet, « *un jeu repose rarement sur l'application simple de combinaisons tactiques apprises précédemment pendant l'entraînement* » (Gréhaigne, Marle & Zerai, 2013, p. 12).

Enfin, en EP, le critère de *validité écologique* défini par Tardif (1993) nécessite d'être relativisé au regard de la notion de « forme de pratique scolaire » (Dhellemmes, 2004). Inscrite au cœur de trois cultures (celle de l'APSA, celle de l'école et celle des élèves et de leur développement), celle-ci peut en effet « *modifier et dépasser la 'forme' de la pratique sociale pour mieux en faire vivre le fond' aux élèves, au sens culturel du terme* » (Mascret & Dhellemmes, 2011, p. 114).

3. Grille d'analyse

Le tableau 1 illustre notre tentative de formaliser une grille d'analyse à partir d'une sélection des apports théoriques présentés ci-dessus, que nous avons jugés compatibles avec l'épistémologie de la discipline scolaire EP.

Tableau 1. Grille d'analyse

Formalisation des compétences attendues des élèves <ul style="list-style-type: none">• Critère de globalité ?• Type de ressources ?
Tâches/situations d'évaluation <ul style="list-style-type: none">• Situations complexes ?• Situations inédites ?
Outils impliqués dans les situations d'évaluation <ul style="list-style-type: none">• Outils matériels ?• Outils symboliques ?
Activité attendue des élèves <ul style="list-style-type: none">• Hiérarchisation ?• Choix des procédures ?
Critères de validité <ul style="list-style-type: none">• Validité de construit ?• Validité de contenu ?• Validité écologique ?• Validité déontologique ?

4. Méthodologie

4.1 Corpus de données

Pour le cycle moyen de l'enseignement primaire (8-12 ans), les prescriptions cantonales que nous avons analysées sont constituées de deux documents : (a) le *Document de liaison. Enseignement et évaluation. Cycle moyen*, rédigé par la Direction générale de l'enseignement obligatoire (DGEO, 2014), a pour but d'aider les enseignants à intégrer leurs pratiques d'enseignement et d'évaluation dans le cadre légal et réglementaire de l'enseignement primaire genevois ; (b) le document *Evaluer les apprentissages en éducation physique* (SCP, 2013) est un document pédagogique interne à la DGEO destiné à illustrer le document de liaison.

Pour le cycle d'orientation (CO)¹ (12-15 ans), nous avons analysé trois documents distincts. Les deux premiers émanent du niveau de référence macro-structurel (Briot, 1999), représenté par la Direction générale du CO : (a) comme son nom l'indique, le document *Clarifications dans le domaine de l'évaluation* (DGCO, 2011-2012) apporte des clarifications dans le domaine de l'évaluation aux enseignants de ce degré scolaire quelle que soit la discipline ; (b) le *Document de liaison. Education physique* (DGCO, 2011-2013) est destiné à faciliter l'application concrète du PER dans les classes genevoises pour ce qui concerne l'EP. (c) Le troisième document est un projet d'enseignement en basket-ball rédigé par l'équipe pédagogique d'un collège genevois à l'intention d'élèves de 10^{ème} année². Il concrétise, à un niveau méso-structurel (Briot, 1999) – celui de l'établissement scolaire – et pour une APSA déterminée, les prescriptions intercantionales et cantonales.

4.2 Analyse des données

Le corpus de données a été analysé selon une démarche inductive en quatre temps : (a) l'auteur principal a extrait, de l'ensemble des documents constituant le corpus de données, les éléments qui faisaient écho aux composantes de la grille d'analyse et a produit une analyse préliminaire ; (b) cette analyse préliminaire a été mise à l'épreuve des analyses des coauteurs ; (c) l'auteur principal a retravaillé l'analyse préliminaire en intégrant les remarques, commentaires, nuances des coauteurs ; (d) l'analyse ainsi retravaillée a été finalisée et validée collectivement lors d'une séance en équipe.

5. Résultats

5.1 Les prescriptions officielles pour le cycle moyen de l'enseignement primaire

Le *Document de liaison* inscrit pleinement et explicitement l'évaluation en EP dans l'approche par compétences. Ainsi, il précise notamment :

L'évaluation en éducation physique consiste à porter un regard sur les compétences développées par les élèves et pas uniquement sur les performances. La compétence en éducation physique consiste à mobiliser, en situation complexe, trois types de ressources principales :

¹ Terminologie genevoise pour désigner l'enseignement secondaire I, dernier degré de l'enseignement obligatoire en Suisse.

² Ancienne terminologie, correspondant à des élèves de 13-14 ans.

- des connaissances (savoirs) ;
- des habiletés (savoir-faire) ;
- des attitudes (savoir-être).

Ces ressources sont les composantes de la discipline. L'évaluation ne consiste pas en une série d'exercices à réussir de manière stéréotypée, spécifique ou décontextualisée (dribbler, passer, etc.), mais en une seule situation complexe sollicitant l'intelligence des élèves (jouer au basket en situation) (DGEO, 2014, pp. 82-83).

Le critère de *globalité* inhérent à la notion de compétence est présent à travers la *tâche* ou *situation d'évaluation* qui doit être *complexe* et solliciter l'intelligence des élèves, aptitude que Crahay (2006) associait de manière assez critique au *savoir-transférer*. « Jouer au basket en situation » constitue le degré le plus élevé de compétence. Cette formulation extrêmement large ne permet pas d'identifier ce qui est effectivement attendu des élèves dans l'APSA basket-ball à ce stade de leur scolarité. Ce travail de formalisation reste à charge des enseignants. De même, les enseignants sont invités à caractériser l'*activité* attendue des élèves selon le classique triptyque de ressources savoirs – savoir-faire – savoir-être (Coulet, 2011).

Il est à noter qu'à l'école primaire à Genève, l'EP est une discipline sans notes, dans laquelle « pour les activités bilan [...], l'enseignant atteste du niveau de connaissances et de compétences avec des appréciations en termes d'objectifs 'atteints' ou 'non atteints, à reprendre' » (DGEO, 2014, p. 22).

Le *Document de liaison* apporte des précisions quant aux modalités de l'évaluation à mettre en œuvre par les enseignants³ :

L'évaluation annuelle est basée sur au moins trois activités bilan (qui sont des observations critériées), complétées par d'autres prises d'information. Ces activités relèvent de plusieurs axes ou domaines d'activité (courir-sauter-lancer ; se maintenir en équilibre, grimper, tourner ; jouer ; etc.)⁴ répartis au cours de l'année. Pour que les objectifs de l'activité bilan soient considérés comme atteints, les trois composantes (connaissances, habiletés, attitudes) doivent être elles-mêmes atteintes. Une composante « très bien atteinte » ne peut donc pas compenser une composante « non atteinte ». Le nombre total de critères doit se situer entre quatre et huit. Ils doivent se répartir sur les trois composantes. Celle des habiletés doit toujours dominer. Pour qu'une composante soit considérée comme atteinte, la moitié au moins des critères d'évaluation doit être atteinte (DGEO, 2014, p. 83).

Pour l'ensemble des disciplines scolaires, ce document précise en outre que :

les objectifs de l'activité et les critères d'évaluation doivent être établis à l'avance et clairement définis afin d'assurer une équité de traitement entre les élèves et une transparence concernant l'évaluation. Les critères d'évaluation des compétences des élèves peuvent être de nature qualitative ou quantitative (DGEO, 2014, p. 19).

A la lecture des extraits précédents, on constate que le *Document de liaison* définit le *niveau hiérarchique* de l'*activité* de l'élève auquel l'évaluation est censée s'opérer, en l'occurrence celui des ressources que l'enseignant aura préalablement définies et auxquelles il aura associé des critères d'évaluation qualitatifs ou quantitatifs. Par un jeu de moyenne, les ressources validées

³ A ce niveau de scolarité, l'enseignement et l'évaluation en EP sont partagés entre des maîtres spécialistes de la discipline et des enseignants généralistes.

⁴ Cette terminologie fait référence aux Manuels fédéraux d'éducation physique édités par la Confédération suisse, qui constituent les principaux moyens d'enseignement à disposition des enseignants d'EP en Suisse.

détermineront la validation ou non des composantes (trois types de ressources principales), laquelle validera ou non *in fine* l'activité bilan.

Le document *Evaluer les apprentissages en éducation physique* comporte un exemple de fiche d'évaluation dans l'activité « s'opposer, lutter ». La situation d'évaluation est décrite par la formule « *en prenant part à des jeux de lutte...* ». Les trois composantes prescrites par le *Document de liaison* y sont respectivement déclinées en deux connaissances (« *tu connais les principales articulations à échauffer pour l'activité de lutte et tu les colories sur ce dessin* » ; « *tu sais abaisser ton centre de gravité et tu es bien sur tes appuis* »), quatre habiletés (« *tu te déplaces pour garder ton équilibre et déséquilibrer ton partenaire* » ; « *tu utilises à ton profit la force de ton partenaire* » ; « *tu sais tirer, pousser* » ; « *tu sais soulever, retourner ton partenaire pour l'immobiliser* »), et deux attitudes (« *tu t'engages dans la lutte et acceptes le contact* » ; « *tu respectes la règle d'or* ») (SCP, 2013, p. 1).

L'évaluation porte donc sur les ressources – prédéterminées par l'enseignant – que l'élève est supposé mobiliser dans l'activité bilan et aucun résultat ou performance n'est associé à cette activité, consistant ici à *jouer au basket en situation* ou *[prendre] part à des jeux de lutte*. Par conséquent, l'élève a peu le choix des procédures à utiliser pour résoudre les problèmes posés par la situation complexe – et inédite dans ces APSA d'opposition – dans laquelle il est évalué. Par ailleurs, à l'exception du dessin sur lequel l'élève est invité à colorier les articulations à échauffer pour l'activité de lutte et de la fiche d'évaluation proprement dite, les documents officiels ne prescrivent pas d'outils matériels particuliers. L'enseignant pourrait toutefois en élaborer selon ses besoins. Le langage des élèves (outil symbolique) semble peu sollicité dans la situation d'évaluation.

De manière générale, ces prescriptions et leur illustration en lutte nous semblent satisfaire au critère de *validité de construit* exigeant le recours à des tâches complexes et signifiantes. En revanche, le manque d'indications, quant au résultat à atteindre par l'élève, ne leur permet pas de satisfaire au critère de *validité de contenu*, l'évaluation portant sur les composantes de la/des compétence(s) à mettre en œuvre pour réussir plutôt que sur la/les compétence(s) elle(s)-même(s) qui n'est/ne sont pas définie(s). Nous manquons d'éléments pour pouvoir véritablement nous prononcer sur le critère de *validité écologique*, même si les exemples de situations d'évaluation proposés dans les différents documents analysés (*jouer au basket en situation*, *[prendre] part à des jeux de lutte*) correspondent à des pratiques sociales de référence significatives. Enfin, le critère de *validité déontologique* repose entièrement sur le travail curriculaire de l'enseignant, à qui il revient d'élaborer les critères d'évaluation qualitatifs ou quantitatifs lui permettant de valider ou non les différentes connaissances, habiletés et attitudes qu'il aura définies. Nous pouvons nous poser la question de savoir, par exemple, à partir de quel moment l'enseignant considérera que l'élève « *sait tirer, pousser* ».

Le tableau 2 synthétise notre analyse des prescriptions officielles pour le cycle moyen de l'enseignement primaire.

Tableau 2. Analyse des prescriptions officielles pour le cycle moyen de l'enseignement primaire

Formalisation des compétences attendues des élèves	
Critère de globalité ?	Présent à travers le type de tâche/situation d'évaluation prescrit
Type de ressources ?	Connaissances (savoirs) / habiletés (savoir-faire) / attitudes (savoir-être)
Tâches/situations d'évaluation	
Situations complexes ?	Oui, une seule situation complexe sollicitant l'intelligence des élèves
Situations inédites ?	Selon l'APSA
Outils impliqués dans les situations d'évaluation	
Outils matériels ?	Non prescrits mais au bon vouloir de l'enseignant
Outils symboliques ?	Non
Activité attendue des élèves	
Hierarchisation ?	Le niveau hiérarchique auquel s'opère l'évaluation est celui des ressources
Choix des procédures ?	L'élève a peu le choix des procédures à utiliser
Critères de validité	
Validité de construit ?	Oui, tâches complexes et significantes
Validité de contenu ?	Non, évaluation portant sur les composantes plutôt que sur la/les compétence(s)
Validité écologique ?	<i>A priori</i> oui
Validité déontologique ?	Dépend du travail curriculaire de l'enseignant

5.2 Les prescriptions macro-structurelles et leur concrétisation au niveau méso-structurel pour le cycle d'orientation

Le document *Clarifications dans le domaine de l'évaluation* se réfère lui aussi explicitement à l'approche par compétences. Il précise notamment que « *l'évaluation doit rester représentative de l'ensemble des objectifs d'apprentissage et doit permettre de tester des progressions et des compétences travaillées en classe* » (DGCO, 2011-2012, p. 3). Ce qui demeure fondamental dans cette période de grands changements, c'est une évaluation :

- conçue comme un élément intégrant du parcours de formation,
- qui s'inscrit dans les différentes étapes de l'apprentissage et dont les règles sont connues des élèves,
- ajustée au fur et à mesure entre les maître-esse-s que ce soit dans les écoles ou dans les groupes de disciplines,
- qui cherche à mettre en valeur les acquis des élèves (DGCO, 2011-2012, p. 5).

L'idée de *tester des progressions et des compétences travaillées en classe* et de *mettre en valeur les acquis des élèves* tendrait à marquer une certaine distance avec la norme de la complexité inédite propre à une acception forte de l'approche par compétences (Mougenot, 2016).

Le *Document de liaison* indique, quant à lui, dans la logique de l'approche par compétences, qu'il s'agit pour les élèves de « *mobiliser des habiletés motrices, des connaissances, des habiletés* »

méthodologiques, des attitudes pour réussir en situation » (DGCO, 2011-2013, p. 3). En matière d'évaluation, il prescrit entre autres de mettre en place une démarche évaluative en trois temps : (a) situation de référence en début de cycle (évaluation à fonction diagnostique) ; (b) évaluation formative en cours de cycle afin de proposer des situations d'apprentissage et/ou de remédiation ; (c) évaluation finale (certificative) pour évaluer les compétences dans l'action et autour de l'action des élèves. Il fixe également les principes que doit respecter la situation d'évaluation finale :

Cette évaluation comporte trois dimensions : la performance collective ou individuelle en référence aux aspects culturels des APSA ; la maîtrise de l'activité qui valide les apprentissages réalisés au cours du cycle d'apprentissage. Ces deux dimensions valident les compétences « dans l'action ». Les compétences « autour de l'action » sont validées en référence aux rôles sociaux inhérents à l'APSA enseignée (DGCO, 2011-2013, p. 4).

Un critère de *globalité* est présent à travers l'idée d'une situation de référence qui devient une situation d'évaluation en fin de cycle. Si ces deux situations, l'une à vocation diagnostique et l'autre à vocation sommative sont identiques, cela renforce notre analyse d'un écart autorisé par les prescriptions macro-structurelles par rapport à la norme de la complexité inédite. Le triptyque savoirs – savoir-faire – savoir être (Coulet, 2011) est ici décliné en quatre types de ressources afin de caractériser l'*activité* attendue des élèves, conformément aux propositions de Delignières et Garsault (2004) : habiletés motrices, connaissances, habiletés méthodologiques et attitudes. Les habiletés méthodologiques renvoient aux ressources constitutives des compétences « autour de l'action » tandis que la catégorie « habiletés motrices » contient les ressources constitutives des compétences « dans l'action ». Les connaissances et les attitudes peuvent être impliquées dans l'un ou l'autre de ces types de compétences. La *performance collective ou individuelle en référence aux aspects culturels de l'APSA* faisant partie des dimensions prises en compte dans l'évaluation de l'*activité* de l'élève, on peut considérer que le *niveau hiérarchique* auquel s'opère l'évaluation est, pour partie en tout cas, celui de la/des compétence(s) dans l'action qui balise(nt) le cycle d'enseignement-apprentissage.

Sans aller jusqu'à proposer un exemple de fiche d'évaluation dans une APSA donnée comme au primaire, le *Document de liaison* fournit néanmoins un canevas précis à respecter en termes (a) de présentation générale (le projet doit notamment être défini en termes de compétences dans l'action et autour de l'action), (b) d'intitulé (organisation, matériel, aspects pratiques et indications pédagogiques, références éventuelles), (c) de fiche d'observation (outil pratique permettant l'observation des élèves), (d) d'évaluation (fiche d'évaluation) et (e) d'organisation de l'épreuve. Parmi les thématiques pour la formation continue, il propose finalement « *la construction des indicateurs nécessaires à l'évaluation des compétences motrices et des aspects propres aux capacités transversales et à la formation générale* » (DGCO, 2011-2013, p. 4).

En termes d'outils matériels, outre la fiche d'évaluation proprement dite, le *Document de liaison* prescrit l'usage de fiches d'observation, outils pratiques permettant l'observation des

⁵ Les compétences dans l'action (Lenzen et al., 2012) correspondent à celles du pratiquant de l'APSA, tandis que les compétences autour de l'action (Lenzen et al., 2012) renvoient à celles mobilisées pour exercer les rôles sociaux (Mascret & Rey, 2011) ou aux savoirs d'accompagnement en relation avec la régulation de toute activité physique : mise en condition physique, échauffement, entraînement, sécurité, réglementation, etc. (Cogérino, 2002).

⁶ Le PER est en effet organisé selon trois entrées : (a) cinq domaines disciplinaires, dont le domaine *Corps et mouvement* qui réunit l'EP et l'éducation nutritionnelle ; (b) les capacités transversales, qui doivent permettre à l'élève d'améliorer sa connaissance de lui-même et concourent à réguler ses apprentissages ; et (c) la formation générale, qui met l'accent sur les aspects éducatifs nécessaires au futur citoyen.

élèves à l'aide d'indicateurs prédéterminés. Comme pour le primaire, un travail de conception, certes bien cadré, reste à charge des enseignants du CO, à titre individuel ou collectif (équipe pédagogique au sein d'un établissement, cf. projet de basket-ball ci-dessous), afin de définir ces indicateurs (critères d'évaluation).

Le projet d'enseignement en basket-ball cible deux compétences, une dans l'action et une autour de l'action : (a) « *alterner jeu placé lorsque la défense protège son panier et jeu de contre-attaque lorsque les configurations de jeu sont favorables pour terminer chaque action par un tir* » ; (b) « *des observateurs/entraîneurs observent le jeu pour prendre/donner des informations sur la qualité du jeu et le comportement des joueurs* ». Le projet liste ensuite les ressources à mobiliser dans l'épreuve, catégorisées selon les prescriptions du document de liaison en (a) habiletés motrices, (b) connaissances, (c) habiletés méthodologiques et (d) attitudes. Les ressources sont elles-mêmes formulées sous la forme d'un verbe – pas nécessairement d'action – suivi du but visé. Dans certains cas, il s'agit de compétences d'un niveau de globalité moindre que les deux compétences principales (ex : « *organiser ses déplacements à la périphérie ou dans les couloirs de jeu direct pour tirer seul sans opposition* »). Dans d'autres cas, il s'agit de compétences élémentaires ou opérations (Carette, 2007) que nous ne considérons pas comme des compétences selon les définitions que nous avons retenues dans l'introduction (ex : « *maintenir une posture équilibrée dans le tir en foulée ou à l'arrêt pour atteindre la cible avec un maximum de réussite* »).

Le projet d'enseignement détaille les trois dimensions de l'évaluation conformément aux prescriptions du *Document de liaison* : (a) la performance collective de l'équipe lors du tournoi final (b) la maîtrise personnelle de l'activité en fonction des niveaux des élèves ; (c) les compétences autour de l'action. La performance est à évaluer sur la base des points marqués par les équipes et de la participation des joueurs à la prise de risque de leur équipe. La maîtrise de l'activité est à apprécier au regard d'indicateurs, au nombre de sept, associés à un barème : balle reçue (BR) ; balle conquise (BC) ; balle perdue (BP) ; passe offensive (PO) ; tir raté (TR) ; tir cadré (TC) ; panier (PA). Ces indicateurs renvoient explicitement, mais de façon non mutuellement exclusive, aux ressources listées dans le projet. Ainsi, le nombre de ballons joués par chaque joueur (somme des ballons reçus BR et conquis BC) est supposé constituer un indicateur de l'habileté à se démarquer. L'habileté à prendre des décisions tactiques offensives et défensives efficaces serait quant à elle évaluée sur la base des indicateurs « passe offensive » (PO) dans le rôle de passeur, « tir cadré » (TC) et « panier » (PA) dans le rôle du tireur, et « balle conquise » (BC) dans le rôle du défenseur. Cette évaluation de la maîtrise de l'activité repose fortement sur la contribution des élèves qui jouent le rôle d'observateurs. En contrepartie, il appartient à l'enseignant de valider la compétence autour de l'action ciblée par le projet, via trois critères d'évaluation qui renvoient peu ou prou à des connaissances, habiletés méthodologiques et attitudes. Un règlement précise les conditions de déroulement de l'épreuve (nombre de joueurs, durée des matchs, rôle des officiels, violations et sanctions, etc.).

A un niveau de référence méso-structurel, ce projet d'enseignement illustre le caractère *complexe* – et *inédit* dans une APSA d'opposition telle que le basket-ball – de la situation prévue pour évaluer les apprentissages réalisés par les élèves durant le cycle d'enseignement-apprentissage. La *tâche*, qualifiée d'épreuve, est précisément définie par un règlement. L'*activité* que l'élève devra mettre en œuvre dans cette épreuve est définie à travers (a) les *compétences* dans l'action et autour de l'action qui orientent le projet et (b) les ressources de types habiletés motrices, connaissance, habiletés méthodologiques et attitudes (cf. *Document de liaison*) qui relèvent de différents *niveaux hiérarchiques* (compétences, opérations). Plusieurs outils matériels sont à utiliser dans cette situation d'évaluation : une fiche d'observation dans laquelle les élèves observateurs consignent le résultat de leur observation d'un pair ; une fiche

d'évaluation dans laquelle l'enseignant synthétise les niveaux de performance, de maîtrise et de compétence autour de l'action de chaque élève.

Ces prescriptions de niveau macro-structurel et leur concrétisation au niveau méso-structurel pour une APSA donnée nous semblent satisfaire au critère de *validité de construit* exigeant le recours à des tâches complexes et significatives. Le critère de *validité de contenu* peut être considéré comme étant également rempli par ces prescriptions qui, à travers un outillage complexe reposant fortement sur le rôle social d'observateur (Mascret & Rey, 2011) confié aux élèves, incitent les (collectifs d') enseignants à formuler une description exhaustive et adéquate des compétences à évaluer et des interactions de leurs composantes. Enfin, la concrétisation, au niveau méso-structurel et pour l'APSA basket-ball, des prescriptions macro-structurelles pour le CO témoigne du potentiel de ces dernières à susciter des projets d'enseignement et d'évaluation valides tant d'un point de vue *écologique* que *déontologique*. Néanmoins, pour le premier critère, la validité est tributaire des choix curriculaires des (collectifs d') enseignants. Pour le second, la validité dépend fortement du sérieux et de la compétence avec lesquels les élèves observateurs s'acquitteront de leur tâche de co-évaluation.

Le tableau 3 synthétise notre analyse des prescriptions macro-structurelles et de leur concrétisation au niveau méso-structurel pour le CO.

Tableau 3. Analyse des prescriptions macro-structurelles et de leur concrétisation au niveau méso-structurel pour le cycle d'orientation

Formalisation des compétences attendues des élèves	
Critère de globalité ?	Présent à travers l'idée d'une situation de référence qui devient une situation d'évaluation en fin de cycle
Type de ressources ?	Habilités motrices / connaissances / habiletés méthodologiques / attitudes
Tâches/situations d'évaluation	
Situations complexes ?	Oui, évaluation finale (certificative) pour évaluer les compétences dans l'action et autour de l'action des élèves
Situations inédites ?	Mise à distance de la norme de la complexité inédite mais les situations d'évaluation peuvent se révéler inédites selon l'APSA
Outils impliqués dans les situations d'évaluation	
Outils matériels ?	Usage prescrit de fiches d'observation à confier aux élèves
Outils symboliques ?	L'utilisation des fiches d'observation par les élèves dans la phase intermédiaire d'évaluation formative sollicite le langage
Activité attendue des élèves	
Hiéarchisation ?	Le niveau hiérarchique auquel s'opère l'évaluation est celui des compétences, via des indicateurs qui renvoient de façon non mutuellement exclusive aux ressources
Choix des procédures ?	Partiellement, les indicateurs contribuant à la validation de la maîtrise de l'activité sont nombreux et ils ne doivent pas être atteints individuellement

Critères de validité	
Validité de construit ?	Oui, tâches complexes et significantes
Validité de contenu ?	Oui, description exhaustive et adéquate des compétences à évaluer et des interactions de leurs composantes
Validité écologique ?	Potentiellement, selon les choix curriculaires des (collectifs d') enseignants
Validité déontologique ?	Potentiellement, selon la validité de la co-évaluation

6. Discussion

Formaliser des compétences en EP et concevoir des outils pour évaluer celles-ci constituent une véritable difficulté professionnelle (Cogérino, 2002 ; Lenzen et al., 2012 ; López-Pastor et al., 2013 ; Mougenot, 2016 ; Ubaldi, 2005). En croisant les apports théoriques que nous avons recensés autour de la notion de compétence et de son utilisation en EP, de l'évaluation des compétences et de l'évaluation en EP, nous avons construit une grille d'analyse (point 3). Nous avons utilisé celle-ci pour analyser les prescriptions genevoises pour l'enseignement obligatoire afin de mettre en évidence les solutions trouvées par leurs auteurs pour dépasser cette difficulté professionnelle ainsi que les écueils restant à surmonter. A partir de cette analyse de productions locales, nous espérons pouvoir produire de nouvelles connaissances utiles à la noosphère et aux enseignants d'EP, voire d'autres disciplines scolaires. Nous allons successivement discuter (a) l'élaboration de notre grille d'analyse, (b) le statut des références curriculaires que nous avons analysées, et (c) les solutions et les écueils pour évaluer en EP dans une approche par compétences.

Dans un premier temps, une compétence se définissant notamment par un critère de *globalité* (Ubaldi, 2005) et la mobilisation d'une diversité de *ressources* (iGen, 2007, cité par Mottier-Lopez & Tessaro, 2010 ; Le Boterf, 1999, cité par Coulet, 2011 ; Perrenoud, 1998), nous avons retenu ces deux critères afin d'analyser comment étaient formalisées les compétences attendues des élèves dans notre corpus de données. L'évaluation de compétences nécessitant, selon plusieurs auteurs, de confronter les élèves à des tâches *complexes* et *inédites* (Carette, 2007 ; Mougenot, 2016), nous avons retenu ces deux critères pour caractériser les tâches/situations d'évaluation préconisées dans les documents que nous avons analysés, quand bien même cette norme de la complexité inédite est sujette à controverse (Crahay, 2006) et doit selon nous être relativisée au regard de l'épistémologie des APSA (cf. section 2.3 ci-dessus). Considérant avec Coulet (2011) que la plupart (sinon la totalité) des activités humaines étaient instrumentées, nous avons jugé pertinent de répertorier les *outils matériels* et les *outils symboliques* impliqués dans les situations d'évaluation prescrites ou construites par une équipe pédagogique que nous avons analysées. Coulet (2011) soulignant par ailleurs la nécessité de prendre en compte l'organisation hiérarchique de l'activité, par exemple entre activité, actions et opérations, nous avons intégré cette considération dans notre grille d'analyse, afin d'analyser (a) le *niveau hiérarchique* auquel s'opérait l'évaluation dans notre corpus de données et (b) si cette évaluation laissait ou non le choix à l'élève des *procédures* à utiliser pour accomplir la tâche prescrite. Enfin, nous avons retenu dans notre grille d'analyse les quatre *critères de validité* propres à l'évaluation dans le paradigme constructiviste dans lequel s'inscrit, selon Tardif (1993), l'approche par compétences.

D'autres choix étaient sans doute possibles et c'est à l'aune des résultats produits par nos analyses – que nous discuterons dans le troisième temps de cette discussion – que le lecteur pourra jauger l'intérêt et les limites de notre grille d'analyse. Parmi les conditions à réunir pour évaluer des compétences selon notre revue de la littérature, nous n'avons volontairement pas intégré à notre grille d'analyse la condition proposée par Carette (2007) de présenter aux élèves des épreuves diagnostiques pour deux raisons : (a) il ne s'agit pas à proprement parler d'une condition constitutive de l'évaluation dans l'approche par compétences, mais d'une condition à y ajouter pour que cette évaluation soit formative et socialement juste (Carette, 2007 ; Hay & Penney, 2009) ; (b) le caractère plus ou moins diagnostique de l'évaluation peut être envisagé comme une conséquence des caractéristiques de celle-ci selon les critères de notre grille d'analyse. Nous discuterons également ce point dans le troisième temps de cette discussion.

Dans un deuxième temps, il nous paraît utile de discuter le statut des références curriculaires constituant notre corpus de données, dans la mesure où il diffère selon le degré scolaire considéré et a des conséquences sur le travail de transposition didactique interne (Chevallard, 1985) à charge des enseignants pour concevoir une évaluation valide selon le paradigme constructiviste. Pour le cycle moyen de l'enseignement primaire, le *Document de liaison* (DGEO, 2014) et le document *Evaluer les apprentissages en éducation physique* (SCP, 2013) relèvent tous deux du niveau de référence macro-structurel, « *plan supérieur officiel de type prescriptif : celui des programmes et instructions officielles* » (Briot, 1999, p. 75). Au sein des écoles primaires genevoises, il n'y a pas d'équipe pédagogique en EP mais un maître spécialiste de la discipline qui prend en charge deux périodes d'EP hebdomadaires pour chaque classe et des enseignants généralistes qui assurent la troisième période hebdomadaire de leurs classes respectives. Cela signifie que la concrétisation des références curriculaires macro-structurelles se fait directement au niveau micro-structurel, « *plan terminal individuel de type actualisé : celui des enseignants* » (Briot, 1999, p. 75). Les références macro-structurelles étant rédigées de manière suffisamment large pour garantir une certaine liberté pédagogique aux enseignants (Maulini et al., 2012), il est normal que le critère de validité déontologique ne puisse s'analyser à ce niveau mais que son analyse soit renvoyée aux niveaux méso- ou micro-structurels. En l'occurrence ici, il s'agit du second, la responsabilité d'élaborer les connaissances, habiletés et attitudes à évaluer, puis les critères d'évaluation qualitatifs ou quantitatifs permettant de les valider revenant aux enseignants. Ces derniers (spécialiste et généralistes) peuvent bien entendu collaborer entre eux mais les conditions structurelles s'y prêtent peu à l'école primaire genevoise (beaucoup de temps passé devant les élèves, maîtres spécialistes voltigeant entre plusieurs établissements, etc.). Par ailleurs, comme le *Document de liaison* (DGCO, 2011-2013) le propose pour le CO, une partie des formations continues (niveau de référence méso-structurel) gérées par la DGEO pourrait être consacrée à cette problématique afin d'aider les enseignants primaires à accomplir ce travail de transposition qui leur incombe.

Pour le CO, le document *Clarifications dans le domaine de l'évaluation* (DGCO, 2011-2012) et le *Document de liaison* (DGCO, 2011-2013) relèvent tous deux du niveau de référence macro-structurel, tandis que le projet d'enseignement en basket-ball relève du niveau de référence méso-structurel, « *plan intermédiaire structurel de type incitatif : celui des projets d'établissement et d'équipe pédagogique ; celui des structures de formation continue* » (Briot, 1999, p. 75). L'existence d'équipes pédagogiques en EP dans chaque établissement du CO genevois permet en effet de concrétiser les références macro-structurelles à ce niveau intermédiaire, même si cela ne se fait pas dans certaines écoles. Nous avons donc pu analyser le critère de validité déontologique à ce niveau méso-structurel, avec comme résultat la validité potentielle du projet d'enseignement en basket-ball, même s'il reste une condition indépendante de notre

analyse à satisfaire pour que la validité soit garantie. Comme bon nombre de propositions actuelles en matière d'évaluation en EP (Butler & Hodge, 2001 ; Johnson, 2004 ; Kolovelonis & Goudas, 2012 ; López-Pastor et al., 2013 ; Nadeau, Richard & Godbout, 2008 ; Patton & Griffin, 2008 ; Richard, Godbout & Gréhaigne, 2000 ; Ward & Lee, 2005), le dispositif analysé ici mobilise l'évaluation par les pairs. Par conséquent, la validité déontologique de l'évaluation certificative dépendra fortement du sérieux et de la compétence avec lesquels les élèves observateurs s'acquitteront de leur tâche de co-évaluation, ceux-ci en étant néanmoins tout à fait capables (Kolovelonis & Goudas, 2012 ; Nadeau et al., 2008 ; Richard et al., 2000).

Dans un troisième temps, et sous réserve du statut différent des documents analysés pour l'enseignement primaire et le CO discuté ci-dessus, nous pouvons mettre en perspective les solutions trouvées par la noosphère et/ou les enseignants de ces deux degrés scolaires et les écueils rencontrés pour formaliser et évaluer des compétences en EP. Pour le cycle moyen de l'enseignement primaire, la noosphère a fait le choix de prescrire aux enseignants d'évaluer des ressources qui renvoient à des situations (*jouer au basket en situation, [prendre] part à des jeux de lutte*) plutôt qu'à des compétences. Cette prescription pose un problème en termes de validité de contenu (Tardif, 1993) et du peu de choix laissé aux élèves quant aux procédures à utiliser (Coulet, 2011) pour résoudre les problèmes posés par la situation complexe et possiblement inédite dans laquelle ils sont évalués. Elle possède en revanche un fort potentiel diagnostique, à l'instar de ce que visent les phases 2 et 3 du modèle de Carette (2007). Ce constat peut être mis en relation avec la logique curriculaire progressive du PER pour l'EP (CIIP, 2010-2016). Pour les cycles 1 et 2 (enseignement primaire), les attentes fondamentales⁷ y sont essentiellement formulées en termes de ressources (exemples issus de l'axe thématique « Jeux individuels et collectifs » : lance et réceptionne une balle ; reconnaît et respecte les rôles dans les jeux pratiques). Pour le cycle 3 (CO), les attentes fondamentales se rapprochent de véritables compétences (exemples issus du même axe thématique : adapte son comportement à la situation de jeu ; participe activement à trois jeux orientés ou de renvoi en ayant un rôle déterminé ; arbitre ses pairs).

En accord avec cette interprétation, nous avons constaté que l'évaluation prescrite par la noosphère aux enseignants du CO et sa concrétisation dans le projet d'enseignement en basket-ball satisfaisaient aux critères de validité de construit et de contenu. En contrepartie, leur potentiel diagnostique est moindre. Certes, des indicateurs tels que proposés dans le projet d'enseignement en basket-ball – conçus en relation mais de façon non mutuellement exclusive et non exhaustive avec les compétences et ressources listées dans le projet – peuvent permettre, à l'enseignant, jusqu'à un certain point, de diagnostiquer si les élèves en échec ou en difficulté le sont parce qu'ils ne possèdent pas les ressources nécessaires ou parce qu'ils sont incapables de les mobiliser en situation. Mais ce potentiel diagnostique a ses limites dans une discipline telle que l'EP où il est couramment admis qu'on puisse savoir ce qu'il faudrait faire pour être plus performant sans savoir le faire (Darnis & Lafont, 2008). Ainsi, par exemple, prétendre évaluer l'habileté à prendre des décisions tactiques offensives efficaces à partir des indicateurs « passe offensive » (PO), « tir cadré » (TC) et « panier » (PA) se heurte au rôle majeur de la technique dans la concrétisation des intentions tactiques. Un élève pourrait performer dans un test de simulation vidéo, mais échouer en situation de jeu réelle à cause d'un répertoire technique déficient. De même, dans un sport collectif, l'habileté à se démarquer ne garantit pas au joueur qui possède cette ressource un nombre élevé de ballons joués, si ses partenaires l'ignorent trop fréquemment. La valorisation, dans la situation

⁷ Les attentes fondamentales décrivent les acquisitions qui sont nécessaires à la poursuite des apprentissages dans les disciplines concernées. Elles déclinent ce que les élèves doivent au moins maîtriser au cours, mais au plus tard à la fin du cycle (CIIP, 2010-2016).

d'épreuve, de la participation de chaque joueur à la prise de risque de l'équipe et, dans une moindre mesure, de la performance collective (résultat de l'équipe au tournoi), la construction d'équipes de niveaux hiérarchisés en leur sein pour espérer préserver ainsi l'égalité des chances de chaque équipe en contrôlant le rapport de force, ainsi que la règle spécifique interdisant d'utiliser plus de cinq dribbles, sont cependant de nature à réduire ce type de biais.

A partir de notre analyse, en EP, il ne semble pas possible d'évaluer exhaustivement les compétences ciblées dans un projet d'enseignement/apprentissage et l'ensemble des ressources qui les constituent, dans une seule et même situation d'épreuve authentique. Cependant, tout au long du cycle d'enseignement/apprentissage, l'enseignant peut observer à titre diagnostique les ressources de ses élèves (techniques, énergétiques, décisionnelles...) dans les situations d'apprentissage décontextualisées ou les exercices (Ubaldi & Olinger, 2006) qu'il ne manquera sans doute pas de programmer.

7. Conclusion

Face au constat de la difficulté professionnelle consistant à formaliser des compétences en EP et à concevoir des outils pour évaluer celles-ci, nous avons tenté d'analyser les prescriptions genevoises pour l'enseignement obligatoire, afin de mettre en évidence les solutions imaginées par la noosphère pour dépasser cette difficulté professionnelle ainsi que les écueils restant à surmonter. Entre l'enseignement primaire et l'enseignement secondaire, nous avons identifié deux solutions présentant des points communs mais également des différences notables. Au niveau des points communs, les propositions de la noosphère préconisent d'évaluer les élèves dans une situation complexe et authentique (validité de construit). Par ailleurs, un important travail de formalisation de compétences et/ou de ressources et d'élaboration de critères d'évaluation reste à la charge des enseignants, individuellement ou collectivement, pour garantir la validité déontologique de l'évaluation. Au niveau des différences, la noosphère a fait le choix de prescrire aux enseignants primaires d'évaluer des ressources qui renvoient à des situations plutôt qu'à des compétences (invalidité de contenu). A l'inverse, nous avons constaté que l'évaluation prescrite par la noosphère et sa concrétisation dans le projet d'enseignement en basket-ball d'un établissement genevois portaient à la fois sur des compétences dans l'action et autour de l'action et sur leurs ressources constitutives (validité de contenu), avec un recours systématique à l'évaluation par les pairs.

Il reste qu'un écart important peut exister entre ce que prescrivent les instructions officielles (niveau macro-structurel), ce qu'élaborent, à partir des références macro-structurelles ou pas, les équipes pédagogiques au niveau des établissements (niveau méso-structurel) et ce que font les enseignants à l'intérieur des murs de leur gymnase (niveau micro-structurel). Il serait dès lors intéressant d'étudier, au regard des contraintes de l'approche par compétences, les pratiques d'évaluation réelles des enseignants, dont on sait qu'elles sont des arrangements avec les prescriptions officielles, avec les élèves et avec eux-mêmes (David, 2000).

8. Références

- Allain, M., Deriaz, D., Voisard, N. & Lentillon-Kaestner, V. (2016). Perceptions des élèves au niveau du soutien de l'enseignant et de leur compétence dans trois contextes au mode de communication de l'évaluation sommative différente en éducation physique et sportive (EPS). *STAPS*, 113, 73-87.
- Allal, L. (2002). Acquisition et évaluation des compétences en situation scolaire. In J. Dolz & E. Ollagnier (Eds.), *L'énigme de la compétence en éducation* (pp. 75-94). Bruxelles : De Boeck.

- Amade-Escot, C. (2005). Milieu, dévolution, contrat. Regard de l'éducation physique. In M.-H. Salin, P. Clanché & B. Sarrazy (Eds.), *Sur la théorie des situations didactiques* (pp. 91-98). Grenoble : La pensée sauvage.
- Astolfi, J.-P. (2008). *La saveur des savoirs. Disciplines et plaisir d'apprendre*. Paris : ESF.
- Briot, M. (1999). Les stratégies des enseignants d'EPS dans le choix de leurs contenus d'enseignement. *Revue française de pédagogie*, 129, 73-85.
- Butler, S.A. & Hodge, S.R. (2001). Enhancing student trust through peer assessment in physical education. *Physical Educator*, 58(1), 30-39.
- Carette, V. (2007). L'évaluation au service de la gestion des paradoxes liés à la notion de compétence. *Mesure et évaluation en éducation*, 30(2), 49-71.
- Chevallard, Y. (1985). *La transposition didactique. Du savoir savant au savoir enseigné*. Grenoble : La Pensée Sauvage.
- Cogérino, G. (2002). Les difficultés de l'évaluation en EP : le cas des savoirs d'accompagnement. *STAPS*, 59, 23-42.
- Conférence intercantonale de l'instruction publique de la Suisse romande et du Tessin (CIIP) (2010-2016). *Plan d'études romand*. Consulté le 7 février 2017 depuis la plateforme du Plan d'études romand dans <http://www.plandetudes.ch>.
- Colet, J.-C. (2011). La notion de compétence : un modèle pour décrire, évaluer et développer les compétences. *Le travail humain*, 74, 1-30.
- Crahay, M. (2006). Dangers, incertitude et incomplétude de la logique de compétence. *Revue française de pédagogie*, 154, 97-110.
- Darnis, F. & Lafont, L. (2008). Effets de la dissymétrie de compétence pour un apprentissage coopératif en dyades en éducation physique et sportive. *Les cahiers internationaux de psychologie sociale*, 79, 69-83.
- David, B. (2000). *Équité et arrangements évaluatifs. Certifier en éducation physique et sportive*. Consulté le 7 février 2017 sur le site de : ife.ens-lyon.fr/publications/edition-electronique/documents-travaux-recherche-education/david/equite/pdf/comp_dav.pdf.
- Delignières, D. & Garsault, C. (2004). *Libres propos sur l'éducation physique*. Paris : Editions Revue EP.S.
- Desrosiers, P., Genet-Volet, Y. & Godbout, P. (1997). Teachers' assessment practices viewed through the instruments used in physical education classes. *Journal of Teaching in Physical Education*, 16(2), 211-228.
- De Ketele, J.-M. (2009). L'approche par compétences : au-delà du débat d'idées, un besoin et une nécessité d'agir. In M. Ettayebi, P. Jonnaert & R. Operti (Eds.), *Logique de compétences et développement curriculaire. Débats, perspectives et alternatives pour les systèmes éducatifs* (pp. 61-78). Paris : L'Harmattan.
- Dhellemmes, R. (2004). Des situations de « Ré(v)érence » aux formes de pratiques scolaires. *Les cahiers du CEDRE*, 4, 67-78.
- Direction générale du cycle d'orientation (DGCO) (2011-2012). *Clarifications dans le domaine de l'évaluation* (v.2, juin 2012). Genève : DIP.
- Direction générale du cycle d'orientation (DGCO) (2011-2013). *Document de liaison. Éducation physique*. Genève : DIP.
- Direction générale de l'enseignement obligatoire (DGEO) (2014). *Document de liaison. Enseignement et évaluation. Cycle moyen*. Genève : DIP.
- Gréhaigne, J.-F. (1989). *Football de mouvement. Vers une approche systémique du jeu*. Thèse de doctorat (non publiée). Université de Bourgogne, Dijon.
- Gréhaigne, J.-F. & Godbout, P. (1995). Tactical knowledge in team sports from a constructivist and cognitivist perspective. *Quest*, 47, 490-505.
- Gréhaigne, J.-F., Godbout, P. & Bouthier, D. (1997). Performance assessment in team sports. *Journal of Teaching in Physical Education*, 16, 500-516.
- Gréhaigne, J.-F., Marle, P. & Zerai, Z. (2013). Modèles, analyse qualitative et configurations prototypiques dans les sports collectifs. *eJRIEPS*, 30, 5-25.

- Hay, P. (2006). Assessment for learning in physical education. In D. Kirk, D. MacDonald & M. O'Sullivan (Eds.), *Handbook of Physical Education* (pp. 312-325). London: Sage.
- Hay, P. & Penney, D. (2009). Proposing conditions for assessment efficacy in physical education. *European Physical Education Review*, 15(3), 389-405.
- Inspection générale de l'éducation nationale (iGen) (2007). *Les livrets de compétences: nouveaux outils pour l'évaluation des acquis*. Paris : iGen, rapport n° 2007-048.
- Johnson, R. (2004). Peer assessment in physical education. *Journal of Physical Education, Recreation & Dance*, 75(8), 33-40.
- Kolovelonis, A. & Goudas, M. (2012). Students' recording accuracy in the reciprocal and the self-check teaching styles in physical education. *Educational Research and Evaluation*, 18(8), 733-747.
- Laparra, M. & Margolinas, C. (2010). Milieu, connaissance, savoir. Des concepts pour l'analyse des situations d'enseignement. *Pratiques*, 145-146, 141-161.
- Lenzen, B., Cordoba, A., Poussin, B., Déneraud, H. & Deriaz, D. (2013). L'intervention en éducation physique à l'école primaire et secondaire à Genève, entre discipline(s) de référence, demandes sociales et discipline scolaire. In J.-L. Dorier, F. Leutenegger & B. Schneuwly (Eds.), *Didactique en construction, constructions des didactiques* (pp. 109-127). Bruxelles : De Boeck.
- Lenzen, B., Déneraud, H. & Poussin, B. (2012). Les compétences personnelles et sociales en question. Illustration à travers deux outils d'évaluation en éducation physique et sportive. *Actes du 24^e colloque de l'ADMEE-Europe*, University of Luxembourg.
- Le Boterf, G. (1999). *L'ingénierie des compétences*. Paris : Les Editions d'organisation.
- Le Paven, M., Roesslé, S., Roncin, E., Loquet, M. & Léziart, Y. (2007). La dévolution dans les activités physiques sportives et artistiques non scolaires. *Education & Didactique*, 1(3), 9-30.
- López-Pastor, V.M., Kirk, D., Lorente-Catalán, E., MacPhail, A. & MacDonald, D. (2013). Alternative assessment in physical education: a review of international literature. *Sport, Education and Society*, 18(1), 57-76.
- Mascret, N. & Dhellemmes, R. (2011). Culture sportive et culture sportive des APSA. In M. Travert & N. Mascret (Eds.), *La culture sportive* (pp. 99-115). Paris : Editions EP&S.
- Mascret, N. & Rey, O. (2011). Culture sportive et rôles sociaux. In M. Travert & N. Mascret (Eds.), *La culture sportive* (pp. 81-98). Paris : Editions EP&S.
- Maulini, O., Capitanescu Benetti, A., Mugnier, C., Perrenoud, M., Progin, L., Veuthey, C. & Vincent, V. (2012). Qu'est-ce qu'une « bonne pratique » ? Raison pédagogique et rapport à l'efficacité chez les futurs enseignants. *Questions Vives*, 6(18), 39-56.
- Mottier-Lopez, L. & Tessaro, W. (2010). Evaluation des compétences à l'école primaire genevoise : entre prescriptions et pratiques. *Mesure et évaluation en éducation*, 33(3), 29-53.
- Mougenot, L. (2016). *Pour une éthique de l'évaluation. Conceptions et pratiques en éducation physique*. Paris : PUR.
- Nadeau, L., Richard, J.-F. & Godbout, P. (2008). The validity and reliability of a performance assessment procedure in ice hockey. *Physical Education and Sport Pedagogy*, 13(1), 65-83.
- Nanchen, M. (2008). *Mandat d'évaluation de la version 1 du Plan d'études romand « PER »*. Consulté le 7 février 2017 sur le site de la CIIP dans <http://www.ciip.ch/documents/showFile.asp?ID=2695>.
- Oslin, J.L., Mitchell, S.A. & Griffin, L.L. (1998). The game performance assessment instrument (GPAD): development and preliminary validation. *Journal of Teaching in Physical Education*, 17(2), 231-243.
- Penney, D., Brooker, R., Hay, P. & Gillespie, L. (2009). Curriculum, pedagogy and assessment: three message systems of schooling and dimensions of quality physical education. *Sport, Education and Society*, 14(4), 421-442.
- Perrenoud, P. (1998). La transposition didactique à partir de pratiques : des savoirs aux compétences. *Revue des sciences de l'éducation*, 24(3), 487-514.
- Patton, K. & Griffin, L.L. (2008). Experiences and patterns of change in a physical education teacher development project. *Journal of Teaching in Physical Education*, 27, 272-291.

- Rey, B. (2007). Pourquoi l'école s'obstine-t-elle à vouloir faire acquérir des savoirs ? In M. Durand & M. Fabre (Eds.), *Les situations de formation entre savoirs, problèmes et activités* (pp. 171-190). Paris : L'Harmattan.
- Rey, B. (2008). *Rapport d'évaluation de la version 1 du Plan d'études romand « PER »*. Consulté le 7 février 2017 sur le site de la CIIP dans <http://www.ciip.ch/documents/showFile.asp?ID=2692>.
- Richard, J.-F., Godbout, P. & Gréhaigne, J.-F. (2000). Students' precision and interobserver reliability of performance assessment in team sports. *Research Quarterly for Exercise and Sport*, 71(1), 85-91.
- Schneider-Gilot, M. (2006). Quand le courant pédagogique « des compétences » empêche une structuration des enseignements autour de l'étude et de la classification des questions parentes. *Revue française de pédagogie*, 154, 85-96.
- Service de la coordination pédagogique (SCP) (2013). *Evaluer les apprentissages en éducation physique*. Genève : DIP.
- Tardif, J. (1993). L'évaluation dans le paradigme constructiviste. In R. Hivon (Ed.), *L'évaluation des apprentissages. Réflexions, nouvelles tendances et formation* (pp. 27-56). Sherbrooke : Editions du CRP.
- Ubaldi, J.-L. (2005). Les compétences en EPS : mode ou réel changement ? In J.-L. Ubaldi (Ed.), *Les compétences* (pp. 61-79). Paris : Editions Revue EP.S.
- Ubaldi, J.-L. & Olinger, J.-P. (2006). Des options collectives. In J.-L. Ubaldi (Ed.), *L'EPS dans les classes difficiles* (pp. 24-34). Paris : Editions Revue EP.S.
- Villepreux, P. (1987). *Rugby de mouvement et disponibilité du joueur*. Paris : INSEP.
- Vygotski, L. (1985). La méthode instrumentale en psychologie. In B. Schneuwly & J.-P. Bronckart (Eds.), *Vygotsky aujourd'hui* (pp. 39-47). Neuchâtel : Delachaux & Niestlé.
- Ward, P. & Lee, M.-A. (2005). Peer-assisted learning in physical education: a review of theory and research. *Journal of Teaching in Physical Education*, 24, 205-225.