



HAL
open science

Vers un consensus européen sur les critères d'une EPS de qualité : résultats du projet QualiTePE pour la France

Christophe Schnitzler, Langer Wiebke, Erin Gerlach, Christian Hermann, Angelo Crapa, Christelle Trio, Lisa Lefèvre

► **To cite this version:**

Christophe Schnitzler, Langer Wiebke, Erin Gerlach, Christian Hermann, Angelo Crapa, et al.. Vers un consensus européen sur les critères d'une EPS de qualité : résultats du projet QualiTePE pour la France. *Movement & Sport Sciences - Science & Motricité*, A paraître. <hal-05305069>

HAL Id: hal-05305069

<https://hal.science/hal-05305069v1>

Submitted on 9 Oct 2025

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

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



HAL Authorization

Vers un consensus européen sur les critères d'une EPS de qualité : résultats du projet QualiTePE pour la France

Christophe Schnitzler, Langer Wiebke, Erin Gerlach, Christian Hermann, Angelo Crapa, Christelle Trio, Lisa Lefèvre

Christophe Schnitzler, Professeur Associé, HDR, Haute Ecole Pédagogique du pays de Vaud (HEP Vaud), E3S UR1342. Email : christophe.schnitzler@hepl.ch

Wiebke Langer, Wissenschaftlicher Mitarbeiter, Universität Hamburg, Fakultät für Psychologie und Bewegungswissenschaft, Bewegungs- und Sportpädagogik. E-mail : wiebke.langer@uni-hamburg.de

Erin Gerlach, Professor, Universität Hamburg, Fakultät für Psychologie und Bewegungswissenschaft, Bewegungs- und Sportpädagogik. E-mail: erin.gerlach@uni-hamburg.de

Christian Herrmann, professor, pädagogische Hochschule Zürich. E-mail: christian.herrmann@phzh.ch

Angelo Crapa, Wissenschaftlicher Mitarbeiter, pädagogische Hochschule Zürich. angelo.crapa@phzh.ch

Christelle Trio, professeure d'EPS, INSPE de Strasbourg. E-mail : christelle.trio@ac-strasbourg.fr

Lisa Lefèvre, Professeure associée HEP Vaud, laboratoire E3S UR1342. E-mail : lisa.lefevre@hepl.ch

Auteur de correspondance : Christophe Schnitzler. Adresse : Av. de Cour 33, 1007 Lausanne, Suisse. Tel : 0033 6 61 99 14 08. E-mail : christophe.schnitzler@hepl.ch

Remerciements : Nos remerciements s'adressent aux membres du consortium QualiTePE pour la qualité de leur collecte de données et leur participation active dans les discussions, ainsi qu'aux experts ayant accepté de participer à cette étude.

Résumé :

Cet article examine le consensus européen sur la conception d'une éducation physique de qualité (EPQ) dans le cadre du projet européen Erasmus+ KA220-HED, intitulé QualiTePe. Il explore l'adaptation du modèle "MAIN-TEACH", initialement conçu pour évaluer la qualité de l'enseignement en général, à la didactique de l'EPS. Pour ce faire, un processus de Delphi a été mené, impliquant 301 experts issus de 17 pays, afin d'établir un consensus et d'identifier les caractéristiques essentielles de l'enseignement d'une éducation physique de qualité. Les analyses comparent également les perspectives de 38 experts français avec celles du consensus issu de pays majoritairement européens. Les résultats ont conduit à l'adaptation du modèle MAIN-TEACH à l'EPS, avec la proposition d'un questionnaire comprenant 6 dimensions et 20 sous-dimensions. Ils révèlent également qu'il n'y a pas de véritable « exception française », les différences significatives ne concernant que 3 des 20 sous-dimensions analysées. Cette étude confirme l'applicabilité du modèle MAIN-TEACH à l'EPS. Elle ouvre des perspectives pour le développement d'un outil d'évaluation des leçons d'EPS et met en lumière quelques critiques méthodologiques, notamment concernant la représentativité des experts et l'absence de dimensions liées à l'inclusion et à la prise en compte du contexte local. Néanmoins, elle réaffirme un alignement significatif des experts Français avec leurs homologues du projet européen ERASMUS+ sur les priorités et les dimensions de l'EPS de qualité.

Mots clés : éducation physique, qualité, Étude de Delphi, Europe, consensus

Summary:

This article examines the European consensus on the conception of quality physical education (QPE) within the framework of the European Erasmus+ KA220-HED project, titled QualiTePe. It explores the adaptation of the "MAIN-TEACH" model, originally designed to evaluate the quality of teaching in general, to the didactics of PE. To achieve this, a Delphi process was conducted, involving 301 experts from 17 countries, to establish a consensus and identify the key characteristics of quality physical education teaching. The analyses also compare the perspectives of 38 French experts with those of the European consensus. The results led to the adaptation of the MAIN-TEACH model to PE, with the proposal of a questionnaire comprising 6 dimensions and 20 sub-dimensions. They also reveal that there is no real "French exception," with significant differences found in only 3 of the 20 sub-dimensions analyzed. This study confirms the applicability of the MAIN-TEACH model to PE. It opens up perspectives for the development of a PE lesson evaluation tool and highlights some methodological critiques, particularly regarding the representativeness of the experts and the absence of dimensions related to inclusion and consideration of local contexts. Nevertheless, it reaffirms a significant alignment between European and French experts regarding the priorities and dimensions of quality PE.

Key words: physical education, quality, Delphi, European, consensus

Nos remerciements s'adressent aux membres du consortium QualiTePE pour la qualité de leur collecte de données et leur participation active dans les discussions, ainsi qu'aux experts ayant accepté de participer à cette étude.

Introduction

Au cours des deux dernières décennies, l'attention portée à la qualité de l'éducation physique (EP) s'est intensifiée, marquant un enjeu majeur dans le dialogue international. La notion d'Éducation Physique de Qualité (EPQ) a été précisément esquissée suite au rapport d'un séminaire UNESCO tenu à Porto-Novo, Bénin, en mai 2005, offrant un cadre référentiel destiné à promouvoir une EP de haute qualité. Ce cadre met en avant deux axes principaux : le développement de la littératie physique pour enrichir les compétences multidimensionnelles des élèves et l'inculcation de valeurs et de culture pour encourager une participation durable aux activités physiques. Ces orientations soulignent l'importance des débats sur les stratégies et la mise en œuvre efficace pour atteindre ces objectifs ambitieux. Nombreuses sont les stratégies, approches et pratiques éducatives qui se sont développées, basées sur le concept de « qualité » de l'enseignement (Wang et al., 2011).

La qualité de l'enseignement est reconnue comme un facteur clé de succès dans l'apprentissage scolaire (Herrmann, 2019). Inspirées du modèle d'offre et de demande d'Helmke (2014, 2017), les approches s'intéressant à la qualité de l'enseignement identifient des caractéristiques superficielles liées à l'organisation et aux méthodes pédagogiques. Elles mettent également en évidence une dimension plus profonde de l'enseignement, influencée par les actions de l'enseignant et la qualité des interactions entre l'enseignant et les apprenants (Herrmann, 2019). Ainsi, un enseignement de qualité est évalué sur sa capacité à développer des compétences spécifiques, en s'appuyant sur une approche interdisciplinaire soutenue par des interactions enrichissantes qui produisent des effets éducatifs tangibles.

La qualité d'un enseignement a souvent été abordée sous l'angle des résultats, avec une emphase sur les "bonnes pratiques" identifiées par des analyses statistiques, telles que les méta-analyses (Hattie et Yates, 2013). Cependant, cette approche tend à négliger l'ajustement des méthodes pédagogiques à la diversité des apprenants et aux contextes éducatifs spécifiques, un aspect fondamental pour comprendre la qualité de l'enseignement.

Cadre théorique et question de recherche

Pour combler cette lacune, Charalambous et Praetorius (2020) ont proposé le modèle MAIN-TEACH, basé sur huit dimensions reflétant les opportunités d'apprentissage et leur impact sur les élèves. Le modèle est issu de la synthèse de cadres existants et de l'analyse de douze outils d'observation en classe en mathématiques, utilisés d'abord dans les pays germanophones et puis à l'international¹. Au départ, la préoccupation de Charalambous et Praetorius (2020) était de penser de façon collaborative à travers des cadres synergiques les contours d'un enseignement de qualité.

L'enjeu initial était de clarifier la multiplicité des perspectives pédagogiques par lesquelles envisager la qualité de l'enseignement non pas sous la forme de liste, mais en trois niveaux

¹ Les douze instruments se déclinent en (cadres génériques : CLASS, DMEE, ISTOF, 3 dimensions de base ; cadres spécifiques aux mathématiques : IQA, MQI, M-Scan ; cadres génériques hybrides spécifiques aux mathématiques : TRU, UTOP, enseignement TEDS, MECORS & QoT).

comprenant une ou plusieurs de ces dimensions qui expliquent sa représentation multicouche. Les dimensions ont des fonctions différentes qu'elles soient au niveau inférieur ou supérieur et intègrent des aspects génériques (comme la différenciation et l'adaptation) et spécifiques à chaque discipline (comme la sélection du contenu et l'évaluation formative). Ce cadre a fait l'objet d'une analyse comparative en mathématiques et en sciences naturelles (Praetorius et al., 2020 ; Praetorius & Gräsel, 2021).

Les auteurs soulignent la nécessité de penser des adaptations et des ajouts spécifiques selon les disciplines enseignées. Il s'agit particulièrement de les lier aux objectifs, contenus et méthodes d'enseignements envisagés (Neumann, 2018). En EP, Herrmann et Gerlach (2020) se sont emparés du modèle pour le mettre à l'épreuve et soumettre le projet de recherche dont fait l'objet cet article.

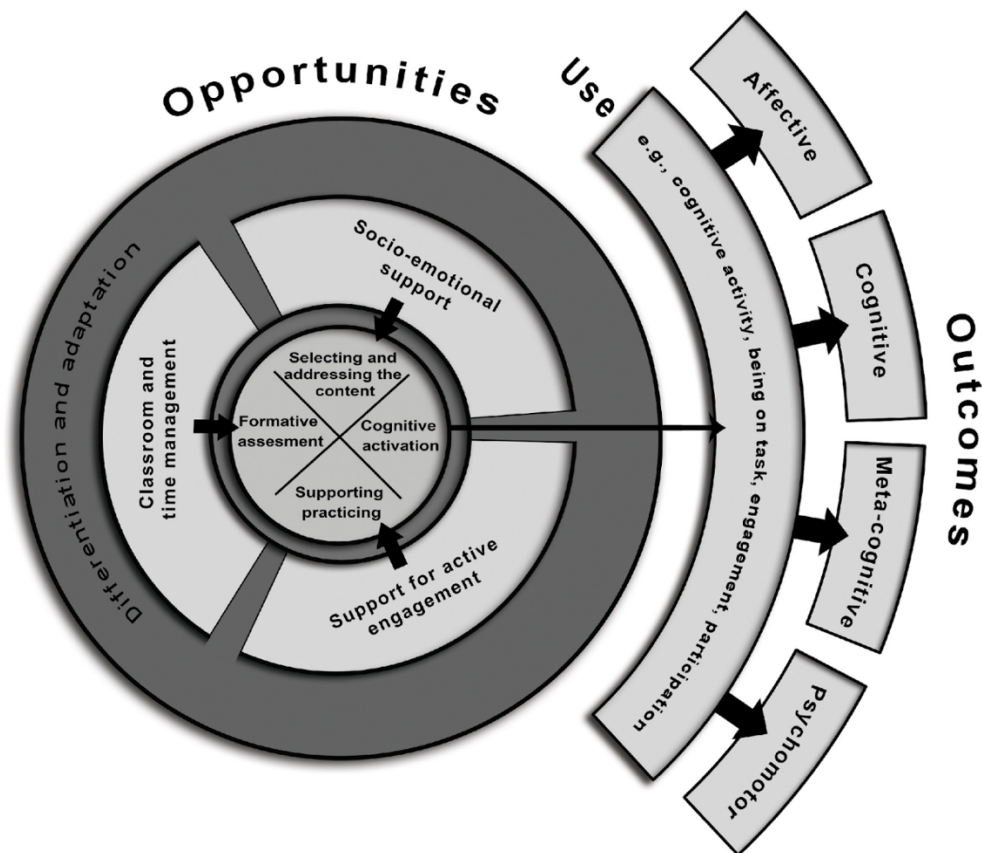
Le modèle offre un cadre pour conceptualiser la qualité de l'enseignement en intégrant des aspects génériques et spécifiques à l'EP, basé notamment sur 3 critères :

- i. La gestion de classe caractérisée par la formulation d'objectifs clairs, de règles explicites d'une utilisation optimale du temps et la mise en place d'un climat sécurisant.
- ii. Un climat d'enseignement centré sur l'élève qui se traduit par des relations soutenant de l'enseignant vis-à-vis de l'élève. Les enseignants favorisent le développement en tenant compte de la diversité des besoins, des intérêts, du vécu antérieur des élèves et des exigences curriculaires. Cette bienveillance basée sur des retours d'information constructifs et personnalisés est alors susceptible de générer des interactions positives au sein de la classe, venant optimiser les apprentissages.
- iii. Une activation cognitive permettant de mobiliser les connaissances antérieures des élèves pour leur permettre d'en explorer et d'en acquérir de nouvelles, tout en développant les capacités réflexives venant soutenir la métacognition (Praetorius et al., 2018).

Ce modèle considère l'enseignement comme un mélange d'opportunités offertes aux étudiants (opportunities) et de l'usage qu'ils font de ces opportunités en termes de produit (appelés outcome, figure 1). Cette utilisation englobe des éléments comme la focalisation des étudiants sur la tâche, leur implication et leur participation à la leçon, ainsi que leur activité cognitive.

Figure 1: The MAIN-TEACH model (Charalambous & Praetorius, 2020)

https://ars.els-cdn.com/content/image/1-s2.0-S0191491X20301425-gr1_lrg.jpg



Ce modèle représente une synthèse des recherches concernant la qualité de l'enseignement général. Comme le montre la Figure 1, il est constitué de huit dimensions génériques et spécifiques à une matière à différents niveaux. Outre les trois dimensions de base et génériques de la qualité de l'enseignement dans la couche intermédiaire (1. Gestion de la classe et du temps, 2. soutien socioémotionnel et 3. soutien pour un engagement actif), ce modèle considère quatre dimensions supplémentaires dans une couche interne (4. Sélectionner et adresser le contenu, 5. l'activation cognitive, 6. le soutien à la pratique et 7. l'évaluation formative). Ces quatre dimensions sont organisées ensemble, puisque toutes visent directement à soutenir les processus d'apprentissage des élèves. Enfin, il existe une dimension transversale (8. Différenciation et adaptation). Cette dimension est particulière en ce qu'elle est entrelacée avec toutes les autres dimensions. Pour les auteurs, « les actions et les interactions des enseignants avec les élèves sont éclairées par des considérations sur les différences entre les élèves ainsi que sur la nécessité de personnaliser l'enseignement en fonction des différents besoins et niveaux de préparation des élèves » (Traduction libre de Charamoulos et Praeorius, 2020, p. 5).

Ce 'squelette' de huit dimensions semble offrir une base solide pour élaborer un référentiel commun de l'enseignement de l'EP de qualité. Il reste tout de même incomplet au regard des spécificités de la discipline, comme les dimensions motrices qui caractérisent l'EPS. Herrmann et Gerlach (2020) ont pointé l'importance de considérer l'engagement de l'élève au travers de l'activation de leur potentiel cognitivo-moteur, mais aussi de penser les processus

d'enseignement et d'apprentissage en lien avec les tâches motrices. Dans la comparaison du modèle entre certaines disciplines par Prateorius et al. (2020), il est également mentionné la dimension spécifique de la gestion de classe en EPS où un plus grand nombre de blessures peut survenir. Aussi la sécurité à travers la gestion des espaces d'apprentissages, du matériel ou des localisations des élèves est également à considérer.

Cependant il n'y a pas à notre connaissance de consensus sur ce qui constitue les dimensions d'une leçon d'EPQ en France comme dans l'Union européenne. Cet article vise à définir les dimensions et les critères essentiels d'une EP de qualité, en s'appuyant sur le modèle MAIN-TEACH tout en intégrant les spécificités de l'enseignement de l'EP. Nous explorerons comment ce modèle peut être adapté pour mieux refléter les spécificités motrices de l'EP et répondre aux besoins diversifiés des élèves, dans le but de proposer un cadre enrichi pour l'évaluation et l'amélioration de la qualité de l'enseignement en EP. Dans un second temps, nous examinerons comment l'adaptation de ce modèle MAIN-TEACH offre la possibilité, d'une part, de dégager un consensus sur les dimensions, sous-dimensions et critères d'une EPS de qualité, et d'autre part, de s'adapter à l'EP(S) française.

Méthodes

La technique de l'étude de Delphi

Cette technique a été développée pour la première fois dans les années 1950 par Dalkey et Helmer (1963) dans le but d'obtenir un consensus d'experts fiable. L'approche – du nom de *l'oracle grec antique de Delphes*, pouvait prédire l'avenir – favorisait l'anonymat et évitait la confrontation potentielle et directe entre experts (Okoli et Pawlowski, 2004). Une approche de Delphi est une méthode de recherche qui consiste à demander à des experts leur point de vue sur un aspect particulier de leur domaine d'expertise (Carr-Chellmann et Kroth, 2018). Ce processus itératif en plusieurs étapes permet à chacun de reconsidérer son point de vue à la lumière de la contribution des autres, tout en maintenant son anonymat, ce qui encourage l'authenticité et l'honnêteté et minore les relations de pouvoir. Les seuils d'accord entre les experts dépendent en fait du champ considéré, et les bornes acceptables pour considérer un niveau de consensus acceptable varient entre 51% et 100% selon les études.

Dans le contexte de ce projet Erasmus+, la méthode Delphi a été utilisée pour obtenir l'avis et les idées des experts des dix-sept pays pour la plupart européens participants afin de développer un consensus sur les critères d'une EP de qualité. Pour cette étude, les étapes suivantes ont été proposées :

1. Étape 1 : choix du questionnaire initial

La première étape est la phase Pré-Delphi au cours de laquelle la liste initiale des dimensions et sous-dimensions a été créée par le consortium du projet ERAMUS+ QualiTePE (QualiTePE n°KA220-HED-E2D5617B) composé d'au moins un membre issu de chacun des 19 pays membres de l'Union européenne. Ce consortium s'est appuyé sur le modèle interdisciplinaire MAIN-TEACH proposé par Charalambous & Praetorius (2018) comprenant initialement 8

dimensions. Les membres du consortium ont été invités à rapporter sur l'état actuel de l'art et la recherche concernant la qualité du processus dans l'enseignement de l'EPS dans leurs pays respectifs, puis à discuter des huit dimensions du modèle MAIN-TEACH pour élaborer une première proposition de caractéristiques d'un enseignement d'EPS de qualité, puis à proposer les sous-dimensions qui composent chaque dimension identifiée. Enfin, pour chaque sous-dimension, une série de questions a été élaborée pour intégrer le questionnaire pilote.

2. Étape 2 : Sélection des dimensions et sous-dimensions du questionnaire QualiTePE

La seconde étape consiste en l'évaluation, par des experts indépendants du projet ERASMUS+, sélectionnés par le consortium du projet ERASMUS+, de la pertinence du contenu de ce questionnaire. Les experts des différents pays sont sollicités pour répondre à un questionnaire en ligne. Ils avaient la possibilité de laisser des commentaires à l'issue du questionnaire. Les experts ont été sélectionnés selon trois critères d'expertise (Williams et Lee, 2020) : des connaissances et compétences avancées combinées à des pratiques pédagogiques efficaces, démontrées par au moins 10 ans d'expérience professionnelle ; une pratique réflexive attestée par leur capacité à participer activement à un groupe de réflexion ; et une adaptabilité dans la gestion de contextes éducatifs diversifiés. Les membres du consortium ont pu ainsi recruter 301 experts dans 17 différents pays européens et extraeuropéens. Le nombre d'experts par pays est présenté dans le Tableau 1.

Pays	Nombre d'experts répondant round 1	Pourcentage du total	Nombre d'experts répondant round 2	Pourcentages du total
Allemagne	43	14,3%	42	22,0%
Belgique	11	3,6%	11	5,8%
Espagne	13	4,3%	4	2,1%
Estonie	8	2,6%	4	2,1%
Finlande	11	3,6%	10	5,2%
France	38	12,6%	28	14,7%
Grèce	6	2,0%	5	2,6%
Irlande	12	4,0%	11	5,8%
Luxembourg	7	2,3%	0	0,0%
République tchèque	26	8,6%	24	12,6%
Slovaquie	3	1,0%	0	0,0%
Slovénie	21	6,9%	16	8,4%
Suède	24	7,9%	16	8,4%
Macédoine du Nord	29	9,6%	1	0,5%
Serbie	2	0,7%	1	0,5%
Suisse	22	7,3%	17	8,9%
Turquie	25	8,3%	1	0,5%
Total	301	100%	191	100%

Tableau 1 : composition du panel d'experts pour les rounds 1 et 2 de l'étude

Certains pays pourraient, mais ne font cependant pas partie de l'Union européenne (Macédoine du Nord, Suisse, Turquie, Serbie). Leur présence s'explique cependant par une volonté de se rapprocher, voire rejoindre l'UE et l'intérêt apporté au projet et aux enjeux d'une EPS de Qualité.

La Figure 1 précise la répartition des professions

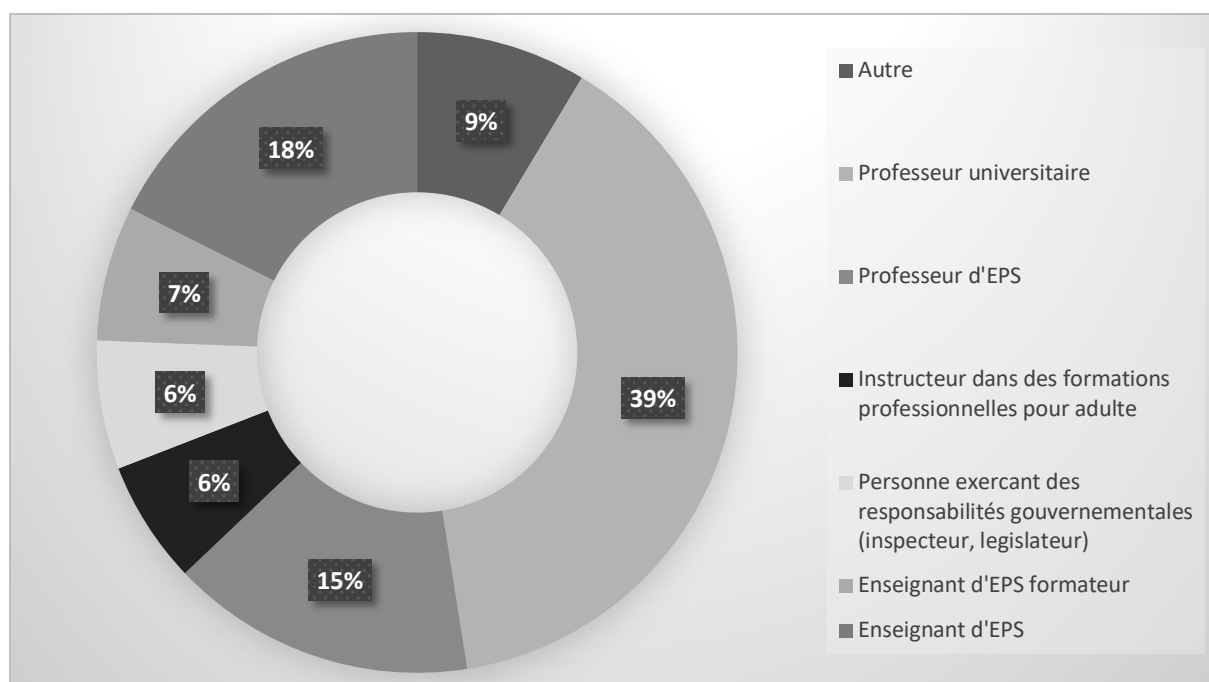


Figure 1 : Origine professionnelle des experts sollicités

La collecte des données s'est étalée sur cinq mois, de septembre 2022 à janvier 2023.

Le panel d'experts retenus est sélectionné au niveau européen selon les critères suivants : *chercheurs/professeurs d'université, instructeurs d'établissements d'enseignement supérieur, personnes ayant des responsabilités au sein du gouvernement de l'éducation dans le domaine de l'éducation et de la formation tout au long de la vie, personnes impliquées dans la formation pratique des enseignants d'éducation physique, personnes engagées dans les syndicats et associations d'enseignants d'éducation physique et sportive, enseignants d'éducation physique qui agissent en tant que mentors ou superviseurs pendant les stages des étudiants*. En France, les experts sont sélectionnés pour leur expertise académique dans le domaine de l'EPS ainsi qu'une expérience professionnelle significative supérieure à 10 années en tant que professionnels. Les personnes visées possèdent une connaissance approfondie des problématiques et des défis spécifiques liés à l'EPS, tant du point de vue des pratiques pédagogiques que de la formation des enseignants. 50 experts sont sollicités dans toute la France en respectant la répartition suivante : un tiers d'enseignants engagés dans la formation ou le tutorat, un tiers de chercheurs ou universitaires, un tiers d'inspecteurs académiques. Sur ce nombre, 38 ont répondu au 1^{er} tour de l'étude Delphi, et 28 lors du second tour. Les questionnaires sont anonymes et validés par une commission éthique de l'université de Hambourg qui a porté le projet Erasmus+ QualiTePE (autorisation n°AZ: 2022_0589).

Évaluation des dimensions et sous-dimensions

Dans un premier temps, les experts avaient pour charge d'évaluer la pertinence de chacune des sous-dimensions à l'aide d'une échelle de Likert à sept niveaux allant de 1 (pas du tout important) à 7 (tout à fait important). Ces échelles permettent de noter l'importance de chaque sous-dimension. Toute sous-dimension jugée insuffisamment pertinente est éliminée, selon les règles décidées en amont par le consortium de chercheurs responsable de l'étude de Delphi. Ces règles sont les suivantes : nous avons évalué le pourcentage de réponses portant une note supérieure ou égale à 5. Lorsque ce pourcentage dépassait 90%, la sous-dimension était conservée. Dans le cas contraire, elle était enlevée.

À l'issue de ce processus, nous avons examiné le nombre de sous-dimensions restant par dimension. Pour être conservée, le consortium a décidé qu'une dimension devait être composée d'au moins trois sous-dimensions. Lorsque le nombre de sous-dimensions retenu était inférieur à ce seuil, le consortium déterminait la façon de traiter ces sous-dimensions, en cherchant à les distribuer sur une autre dimension.

Étape 3 : classification des dimensions et sous-dimensions par ordre d'importance

Dans cette étape, les experts devaient hiérarchiser, par ordre d'importance, les sous-dimensions composant chaque dimension sélectionnée lors de l'étape précédente.

Pour résumer, ces trois étapes ont consisté à 1. Adapter le modèle MAIN-TEACH aux spécificités de l'EP ; 2. Sur la base de cette proposition, de sélectionner les sous-dimensions qui font consensus auprès d'un large panel d'experts européen ; 3. De hiérarchiser, les dimensions et, au sein de chaque dimension, de hiérarchiser ses sous-dimensions par ordre décroissant d'importance (1 étant très important). Les rangs moyens pour chaque dimension et sous-dimension ont été relevés et analysés.

Étape 4 : Analyses des différences entre les experts européens et Français

Au cours de cette étape, les scores attribués à chaque sous-dimension lors de l'étape 2 (sur des échelles de Likert allant de 1 à 7), ainsi que la manière dont les experts européens et français ont classé les 6 dimensions retenues lors de l'étape 3 par ordre d'importance, ont été comparés statistiquement. Après avoir vérifié la normalité de la distribution des scores associés à chaque question, nous avons effectué une série de tests t et de comparaison des moyennes pour échantillons indépendants afin de déterminer l'existence d'une éventuelle différence significative. Nous avons ainsi comparé :

- (i) Les différences de scores attribués aux sous-dimensions à l'aide des échelles de Likert entre les répondants français et européens lors de la seconde étape du processus de Delphi.
- (ii) Les rangs moyens accordés aux dimensions entre la France et l'Europe dans la seconde partie de l'étude.

Pour l'ensemble de ces analyses, les différences étaient considérées comme significatives avec un seuil de 5 % ($p < 0,05$).

D'un point de vue qualitatif, nous avons ensuite recueilli les remarques faites par les experts français sur le questionnaire et en proposons une synthèse.

Résultats

Étape 1 : Choix du questionnaire initial

Les membres du consortium QualiTePE ont tout d'abord élaboré une série de questions rattachées chacune aux dimensions du modèle MAIN-TEACH. Il est cependant apparu au cours de cette démarche que la dimension 8 « différenciation et adaptation » était réellement transversale, et amenait à poser des questions redondantes avec d'autres dimensions (notamment « soutien à la pratique » et « soutien à l'engagement actif »). Le questionnaire initial n'inclut donc pas nommément la dimension 8, et intègre au final sept dimensions et 28 sous-dimensions. C'est ce questionnaire ainsi adapté qui a été soumis pour l'étude Delphi présenté dans l'étape 2.

Étape 2 : Sélection des dimensions et sous-dimensions

Lors du « Delphi Round 1 », 301 experts ont évalué les 28 sous-dimensions sur une échelle de Likert à sept niveaux contenant des notes de 1 – pas du tout important à 7- très important. Seules les sous-dimensions ayant obtenu plus de 90 % de réponses avec une note de 5, 6 ou 7 ont été conservées pour le « Round 2 ». Ainsi, plusieurs sous-dimensions, dont 6.1 « alignement constructif », ont été éliminées, ce qui a conduit à la suppression complète de la dimension 6 « évaluation formative », le nombre de questions restantes étant insuffisantes. Le contenu de la sous-dimension 6.2 « feedback constructif » a été intégré à la sous-dimension 5.1 « soutien individuel à l'apprentissage », réduisant le questionnaire à 6 dimensions et 20 sous-dimensions.

Étape 3 : Classification par ordre d'importance

Dans le « Delphi Round 2 », les experts ont classé les dimensions par ordre d'importance afin de déterminer si certaines présentaient des scores d'importance trop marginaux au sein des dimensions. Les mêmes questions ont été posées concernant les sous-dimensions au sein des dimensions. Pour ce faire, les experts ont été sollicités à deux reprises *via* des questionnaires en entre août 2022 et janvier 2023. À la fin de ce processus, il a été décidé de conserver toutes les sous-dimensions sans classement interne significatif, car les écarts entre les moyennes étaient minimales. Ces 20 sous-dimensions servent désormais d'indicateurs pour les 6 dimensions retenues.

L'ensemble des résultats des étapes pré-Delphi, étape 1 et étape 2 sont présentées dans le tableau 2 :

	Dimension et sous-dimensions questionnaire initial	Delphi Round 1		Dimension et sous-dimensions questionnaire après round 1	Delphi round 2
1	Sélection et présentation du contenu et traitement des objectifs connexes	% réponses >5/7	1.	Sélection et présentation du contenu et traitement des objectifs connexes	Classement sous-dimensions
1.1	Définition des objectifs (à court et long terme)	94.2	1.1	Définition des objectifs (à court et long terme)	2.5
1.2	Sélection d'un contenu qui a du sens au regard des conditions de vie des élèves	92.6	1.2	Sélection d'un contenu qui a du sens au regard des conditions de vie des élèves	2.8
1.3	Sélection appropriée de contenus exigeants	93.2	1.3	Sélection appropriée de contenus exigeants	3.0
1.4	Communication claire et compréhensible	92.9	1.4	Communication claire et compréhensible	3.1
1.5	Ordre logique des contenus pour la progression de l'apprentissage	93.5	1.5	Ordre logique des contenus pour la progression de l'apprentissage	3.4
1.6	Prise en compte claire des objectifs divers et variés de l'EP	89.2			
2.	Gestion de la classe		2.	Gestion de la classe	
2.1	Clarté des règles de conduite, des routines et des attentes.	93.2	2.1	Clarté des règles de conduite, des routines et des attentes.	2.5
2.2	Gestion des comportements	89.2			
2.3	Gestion efficace du temps	84.2			
2.4	Organisation bien structurée	93.2	2.2	Organisation bien structurée	2.4
2.5	Présence et accessibilité de l'enseignant	91.4	2.3	Présence et accessibilité de l'enseignant	2.4
2.6	Gestion de l'espace et des équipements	86.7			
2.7	La santé et la sécurité pendant le cours	94.4	2.4	La santé et la sécurité pendant le cours	2.6
3.	Soutien socioaffectif et climat social		3.	Soutien socioaffectif et climat social	
3.1	Climat d'enseignement centré sur l'élève	94.7	3.1	Climat d'enseignement centré sur l'élève	2.3
3.2	Empathie de l'enseignant	96.6	3.2	Empathie de l'enseignant	2.7
3.3	Culture du feedback positif	96	3.3	Culture du feedback positif	2.9
3.4	Promotion d'un climat de classe positif	97	3.4	Promotion d'un climat de classe positif	2.0
4.	Activation cognitive et motrice		4.	Activation cognitive et motrice	
4.1	Fait usage de buts et d'objectifs transparents pour l'activation cognitive et motrice	86.7			
4.2	Consignes et tâches d'apprentissage exigeantes	91	4.1	Consignes et tâches d'apprentissage exigeantes	1.7
4.3	Promotion de l'activité motrice des élèves	90.4	4.2	Promotion de l'activité motrice des élèves	1.9
4.4	Promotion de l'activité cognitive de l'élève	91.0	4.3	Promotion de l'activité cognitive de l'élève	2.3
5.	Soutien à la pratique		5.	Soutien à la pratique	
5.1	Soutien individuel à l'apprentissage	93.0	5.1	Soutien individuel à l'apprentissage	1.5
5.2	Évolution du soutien à la pratique	89.1			
5.3	Approche constructive des difficultés des élèves	93.2	5.2	Approche constructive des difficultés des élèves	1.4
6	Évaluation formative				
6.1	Alignement entre objectif d'apprentissage et évaluation	89.3			
6.2	Feedback constructif	92.0			
7.	Soutien à l'engagement actif		6	Soutien à l'engagement actif	
7.1	Promotion de la participation et de l'engagement actif de tous les élèves	93.2	6.1	Promotion de la participation et de l'engagement actif de tous les élèves	1.4
7.2	Inclusion de tous les élèves dans toutes les activités d'apprentissage	94.7	6.2	Inclusion de tous les élèves dans toutes les activités d'apprentissage	1.5

Tableau 2 : Évolution des dimensions et sous-dimensions issues des étapes du processus de Delphi. Les sous-dimensions sous le seuil critique de 90% sont grisées.

Étape 3 : Deuxième round de l'étude de Delphi :

L'analyse de la hiérarchisation des sous-dimensions à l'intérieur de chaque dimension ne laisse pas apparaître de différences significatives entre les rangs moyens des sous-dimensions à l'intérieur des dimensions.

Étape 4 : Analyses des différences entre les experts européens et français

Comparaison des réponses concernant les sous-dimensions

La Figure 2 présente le résultat de l'étape 4, c'est-à-dire la comparaison entre les réponses des experts français et la moyenne du consensus européen concernant les sous-dimensions.

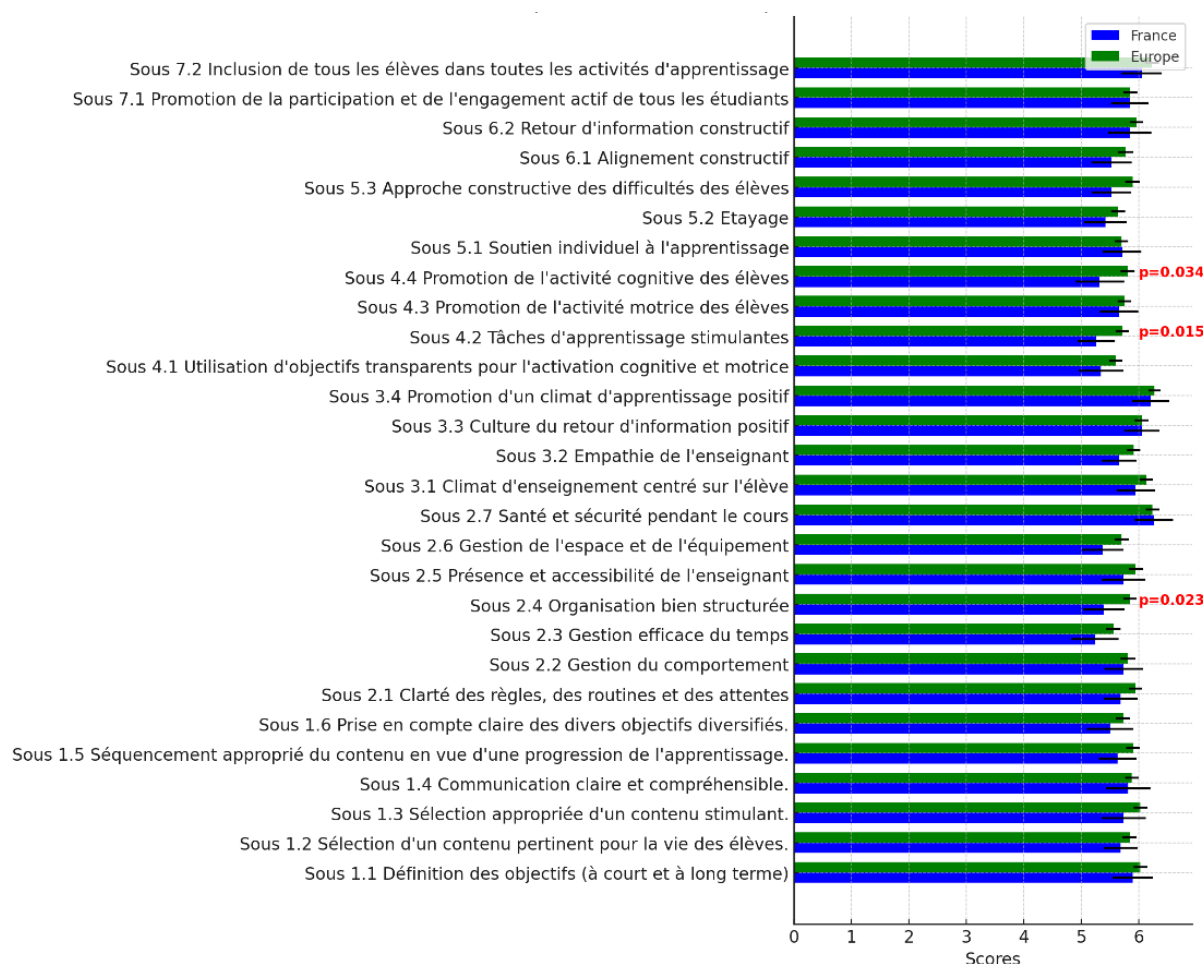


Figure 2 : Comparaison des scores entre les experts français et le consensus de l'ensemble des experts participant au programme Erasmus+ QualiTePE sur les sous-dimensions retenues

Lorsqu'on compare ces réponses, les experts français montrent des scores significativement inférieurs par rapport au consensus de l'ensemble des experts sur les sous-dimensions 2.4 'Organisation bien structurée', 4.2 'Tâches d'apprentissage stimulantes', et 4.4 'Promotion de

l'activité cognitive des élèves'. Ces éléments constitutifs d'une EPS de qualité méritent une attention particulière dans une démarche visant à améliorer les pratiques professionnelles d'enseignement.

Comparaison du rang d'importance des dimensions

Les résultats concernant la comparaison classement des dimensions par ordre d'importance sont présentés dans la Figure 3

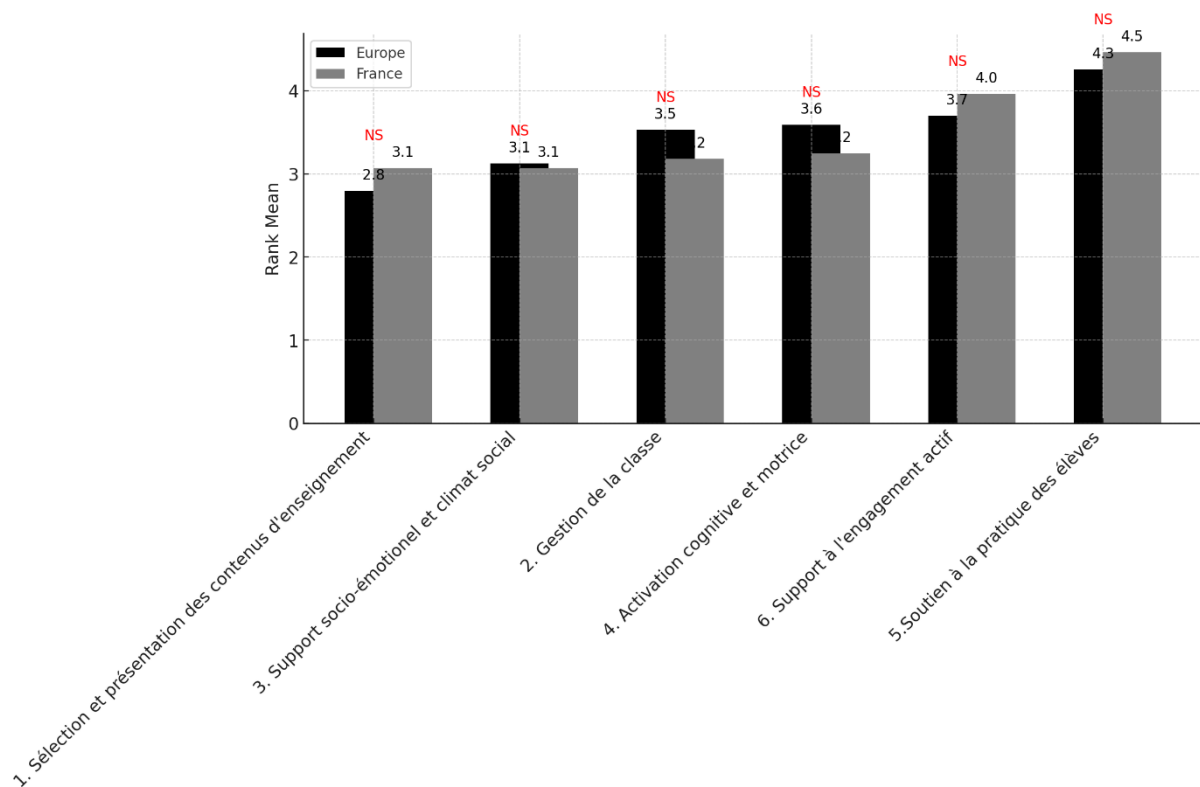


Figure 3 : Analyse comparative des rangs moyens attribués à différentes dimensions évaluant la qualité de l'éducation physique et sportive (EPS) entre la France et l'ensemble de l'Europe

L'analyse statistique ne révèle aucune différence statistiquement significative au seuil de 5% entre les scores moyens d'importance de chaque dimension pour la France et l'ensemble des experts du projet QualiTePE, ce qui indique que l'ordre d'importance de ces dimensions ne diffère pas significativement.

Synthèse des remarques faites par les experts Français

38 experts ont rempli le document visant à évaluer la pertinence du questionnaire. Les réponses des experts fournissent des perspectives intéressantes sur

- La sélection et présentation du contenu: Les experts ont suggéré certaines modifications, telles que l'inclusion de comportements liés au fair-play, la considération de l'inclusion des handicaps et l'importance d'ajouter des aspects liés à la psychologie sportive.
- Support pour l'engagement actif : Certains experts ont proposé d'inclure des paragraphes sur le travail de groupe pour favoriser l'engagement actif.

Discussion

L'objectif de ce projet, inscrit dans une démarche collaborative à travers le programme Erasmus+, était de clarifier le phénomène complexe de l'enseignement de qualité en EPS à travers une étude de consensus en utilisant le modèle MAIN-TEACH. Il s'agissait également, dans un second temps, d'identifier les aspects sur lesquels les experts français adhéraient à ce consensus. Nos résultats montrent, d'abord, l'adéquation du modèle MAIN-TEACH pour l'EPS, puis d'autre part, un consensus européen concernant les dimensions et sous-dimensions du questionnaire QualiTePE auprès d'un panel de 301 experts européens recrutés, et enfin, qu'il n'existe pas de « spécificité française » dans la conception de l'EPS de qualité. Les résultats montrent non pas une cacophonie de perspectives, mais une harmonieuse convergence des voix d'experts.

Cette étude en quatre étapes a cherché à dégager un consensus sur une définition de l'EPS de qualité en Europe, en appuyant son design sur un modèle théorique récent qu'est le MAIN-TEACH (Charalambous et Praetorius, 2020).

Le modèle MAIN-TEACH a été adapté à l'EPS en utilisant un processus de Delphi impliquant 301 experts de 17 pays, ce qui prouve sa solidité pour conceptualiser la qualité de l'enseignement dans un contexte multidimensionnel. L'application de ce modèle dans le domaine particulier de l'éducation physique est validée par le maintien de six dimensions essentielles (sélection du contenu, gestion, soutien socio-émotionnel, activation cognitivo-motrice, soutien à la pratique, soutien à l'engagement). Cependant, il est important de mettre en évidence certaines critiques méthodologiques, telles que l'absence d'une dimension spécifiquement consacrée à l'inclusion et aux particularités locales, deux éléments pourtant essentiels dans la création d'une EPS inclusive et contextualisée.

Dans l'étape préliminaire, le consortium du projet Erasmus+ a adapté le modèle MAIN-TEACH aux spécificités de l'EPS en considérant la différenciation et l'adaptation comme des éléments transversaux intégrés dans les sous-dimensions, ce qui a conduit à la définition de sept dimensions et 28 sous-dimensions.

Au cours de la deuxième étape, le questionnaire portant sur ces sous-dimensions a été présenté à un groupe de 301 experts provenant de 17 pays. Une sous-dimension devait avoir un taux d'approbation supérieur à 90 % et un score minimal de cinq sur une échelle de Likert pour être retenue. Lors de cette sélection rigoureuse, certaines sous-dimensions ont été supprimées, notamment celle liée à l'évaluation formative, afin de ne garder que six dimensions essentielles (sélection du contenu, gestion, soutien socio-émotionnel, activation cognitivo-motrice, soutien à la pratique et soutien à l'engagement), qui sont en accord avec l'ensemble.

Ces attributs s'alignent sur ceux typiquement mis en avant dans les approches holistiques de l'EPS, notamment celle proposée par l'UNESCO (2015). Ils sont également en accord avec les facteurs soulignés par plusieurs auteurs : Whitehead (2001) point l'importance de développer la littératie physique, c'est-à-dire la motivation, la confiance, la compétence physique, ainsi que la connaissance et la compréhension, recoupant 4 des 6 dimensions proposées ; Mali et Kumar (2017) ont mis l'accent sur la valeur sociale et la signification individuelle, venant soutenir l'importance de la sélection d'un contenu pertinent et le soutien socio-émotionnel ; Mottet et Beebe (2006) ont suggéré des facteurs liés au domaine affectif, notamment le développement des attitudes, des croyances et des valeurs, venant également renforcer l'importance des dimensions liées au soutien de la participation et de l'engagement.

Les scores attribués par le groupe d'experts ont cependant conduit, compte tenu de nos critères, à éliminer différentes sous-dimensions. Il en est ainsi de la sous-dimension "évaluation formative". Étonnant de prime abord, ce choix ne minimise cependant pas son importance dans l'enseignement et l'apprentissage de qualité. Au contraire, cela reflète une décision méthodologique visant à capturer les aspects les plus critiques et les plus directement mesurables de la qualité de l'EPS selon les objectifs spécifiques de l'étude ou du cadre d'évaluation. Plusieurs hypothèses peuvent permettre d'expliquer ce choix. On pourrait en effet supposer que les experts peuvent avoir privilégié des dimensions plus directement observables et mesurables dans le cadre d'une éducation physique. L'évaluation formative, bien qu'essentielle, peut être plus subtile et variée dans ses méthodes et moins immédiatement apparente dans une observation de cours d'EPS. Ou encore que les autres dimensions sélectionnées couvrent déjà largement les principes de l'évaluation formative, conduisant à évaluer les questions composant cette dimension comme redondantes.

Malgré cela le haut pourcentage d'accord entre experts des différentes nations souligne l'adéquation de ce modèle pour l'évaluation de l'EPS, dans un contexte où des conditions de conservation des items particulièrement restrictives, une dimension ne pouvant être conservée si son pourcentage de réponse supérieur à 5/7 était supérieur à 90%.

La troisième étape a impliqué de trier les sous-dimensions sélectionnées en fonction de leur importance, mettant ainsi en évidence l'importance de chaque sous-dimension dans la définition d'une EPS de qualité.

Finalement, l'étape 4 a établi une comparaison entre les réponses des experts français et celles de leurs homologues européens, ce qui a révélé un consensus global. Malgré des scores légèrement inférieurs en France pour trois sous-dimensions (2.4 organisation bien structurée, 4.2 tâches d'apprentissage stimulantes, 4.4 promotion de l'activité cognitive des élèves), l'ordre des priorités n'a pas été modifié, ce qui témoigne de l'importance donnée aux critères d'évaluation de l'EPS entre la France et le reste de l'Europe. Cela suggère que les experts français considèrent la dimension motrice comme prédominante, laissant de côté d'autres aspects vus comme moins essentiels. Cependant, ces désaccords avec le consensus européen semblent bien maigres si l'on examine l'ordre des priorités, strictement identique entre celle des experts Français et celui de l'ensemble du projet QualiTePE. La figure 3 montre dans ce contexte que la priorité semble être donnée aux aspects préparatoires à l'enseignement et de

gestion. Si le primat semble être, au niveau des notes données aux sous-dimensions, à la dimension motrice, celle-ci semble s'effacer en importance sur des savoirs liés à la gestion de classe.

Analyse des différences entre experts européens et français

Des différences entre experts européens et français ont cependant été objectivées sur trois sous-dimensions. L'analyse qualitative apporte quelques éléments explicatifs. Dans la dimension 4 « activation cognitive et motrice », la sous-dimension « Promotion de l'activité cognitive des élèves » est initialement classée sous la moyenne des autres pays européens, puis en dernière position par les experts français au second round. Cette tendance reflète la conception française de l'activité cognitive. Les experts soulignent l'importance d'objectifs clairs, adaptés aux progrès des élèves, et d'un enseignement explicite pour soutenir leurs apprentissages. Cependant, ils mettent en garde contre une activation cognitive excessive qui pourrait réduire le temps dédié à l'activité motrice, jugée essentielle. Lors du second round, la priorité est accordée à l'engagement moteur, considéré comme central, tandis que la dimension cognitive est davantage intégrée à travers des actions concrètes et des retours spécifiques sur le processus d'apprentissage de chaque élève.

Concernant la dimension 5, « soutien à la pratique des élèves », la sous-dimension « approche constructive des difficultés des élèves » obtient également des scores inférieurs à la moyenne du projet QualiTePE. Les experts Français soulignent l'importance du retour d'information pour une éducation physique de qualité, mais privilégient une réponse individualisée aux difficultés des élèves, plutôt que de promouvoir des interactions entre pairs. Une relation enseignant-élève de qualité est considérée comme cruciale, bien que certains experts regrettent le manque de considération des interactions entre élèves, qui favoriseraient le développement de compétences transversales.

Ce questionnaire, centré sur l'analyse de la leçon, adopte une approche plus spécifique que les approches globales qui intègrent des aspects politiques et environnementaux (Ho et al., 2019 et le projet Eupeo, 2021). Ce choix méthodologique permet d'observer de manière approfondie les interventions pédagogiques en EPS. Cependant, le modèle MAIN-TEACH, fondé sur une approche constructiviste, fait l'objet de critiques, notamment épistémologiques, car il tend à minimiser l'importance des interactions sociales, pourtant essentielles à l'apprentissage.

En outre, le fait que les activités ludiques, l'adaptation environnementale et la contextualisation locale ne soient pas interrogées – pourtant reconnues comme des éléments essentiels d'une EPS de qualité (Naspe, 2004) – restreint la portée de l'évaluation.

La composition du groupe d'experts pose également question : le choix entre les différents acteurs de la profession semble relativement équilibré, mais on ne peut en dire autant au niveau de la représentativité du nombre d'experts par nation. La composition du groupe d'experts soulève également des interrogations : la répartition des différents acteurs de la profession (scientifiques, enseignants, inspecteurs, etc.) semble relativement homogène entre les pays,

mais il n'en va pas de même pour la représentativité du nombre d'experts par nation. En effet, ce nombre varie considérablement, allant de 1 à 43 experts selon les pays. De plus, seuls 10 des 17 pays représentés dépassent le seuil critique de 12 experts au round 1 ; ils ne sont plus que 6 pays à atteindre ce seuil au round 2.

Cependant il faut noter que l'objectif de cet article n'est pas d'interroger la pertinence du questionnaire, mais d'évaluer si les consensus français et celui des experts issus du programme Erasmus+ QualiTePE sont équivalents eu égard au modèle sélectionné. Les résultats montrent en effet un remarquable alignement culturel des Français sur ces aspects par rapport à leurs collègues européens.

Perspectives

Se pose alors la question de l'utilisation pratique qui peut être tirée de cette étude. Le modèle MAIN-TEACH amène en effet un regard nouveau permettant de caractériser les dimensions d'une EPS de qualité. Pourrait-il servir de base à l'évaluation des leçons d'EPS, et dans l'affirmative, selon quel contexte ? Cette proposition se heurte à plusieurs écueils.

Le premier est lié au caractère situé de l'enseignement de l'EPS. Aucun enseignant ne prétendra développer simultanément ces 28 sous-dimensions (ou caractéristiques) de l'EPS de qualité au sein d'un même cours. L'enseignant peut, en revanche, orienter ses contenus de manière à ce qu'ils répondent aux besoins relevés chez ses élèves, le conduisant à mettre en avant une ou plusieurs dimensions spécifiques dans son enseignement. Ainsi une utilisation modulaire de cet outil, basé sur les objectifs sélectionnés par l'enseignant lui-même, pourrait permettre une observation plus ciblée et des discussions plus riches entre les acteurs de cette évaluation.

Cette remarque nous permet d'initier une réflexion sur la nature des acteurs concernés. En effet le caractère situé d'une leçon impose d'adopter un point de vue phénoménologique inhérent aux interactions humaines. Qualifier une leçon d'EPS de bonne ou de mauvaise dépend souvent du point de vue de l'agent, qui peut avoir des rôles différents (enseignant, formateur, élève...). En l'absence de repères objectifs témoignant du succès ou de l'échec de l'enseignement, il conviendrait de proposer un outil d'analyse qui permette ce croisement des regards comme des focus groups ou qui prendrait en compte les besoins éducatifs particuliers des élèves dans l'appropriation de l'outil.

Conclusion

Cette recherche met en évidence la pertinence du modèle MAIN-TEACH pour définir les contours d'une leçon d'éducation physique et sportive (EPS) de qualité à travers une étude rigoureuse et méthodiquement structurée. La forte adhésion parmi un large panel d'experts engagés dans un projet européen Erasmus+ et l'absence notable d'une "exception française" confirment l'universalité et l'applicabilité du modèle, tout en reconnaissant des nuances spécifiques à chaque contexte national.

Malgré certaines critiques portant sur les choix épistémologiques et la représentativité des experts, les résultats de notre étude soulignent un alignement culturel et professionnel significatif dans la conception de l'EPS de qualité en Europe. Ces constatations encouragent l'application pratique du modèle MAIN-TEACH pour l'évaluation de l'EPS, tout en recommandant une approche modulaire et adaptable qui respecte la diversité des contextes éducatifs et des besoins des élèves.

Ainsi, cette recherche réaffirme l'importance d'une approche holistique et inclusive dans l'évaluation de l'EPS, permettant de mieux répondre aux défis contemporains de l'éducation physique.

Bibliographie

Carr-Chellman, D. J., & Kroth, M. (2018). Profound learning and living: An exploratory Delphi study. In *Adult education research conference (AERC) 2018 conference proceedings*. Manhattan, KS: New Prairie Press, Kansas State University. Retrieved July 2, 2020, from <https://newprairiepress.org/aerc/2018/papers/25>.

Charalambous, C. Y., & Praetorius, A.-K. (2020). Creating a forum for researching teaching and its quality more synergistically. *Studies in Educational Evaluation*, 67, 100894. <https://doi.org/10.1016/j.stueduc.2020.100894>

Dalkey, N., & Helmer, O. (1963). An experimental application of the Delphi method to the use of experts. *Management science*, 9(3), 458-467.

EUPEA. (2021). EUPEO: European Physical Education Observatory: Final Report. Disponible sur: <https://eupea.com/portfolio/eupeo/>

Hattie, J., & Yates, G. C. (2013). *Visible learning and the science of how we learn*. Routledge.

Helmke, A. (2014). *Unterrichtsqualität und Lehrerprofessionalität*. Seelze-Velber: Klett.

Helmke, A. (2017). *Unterrichtsqualität und Lehrerprofessionalität. Diagnose, Evaluation und Verbesserung des Unterrichts*. Bd. 7. Seelze: Klett.

Herrmann, C. (2019). Evaluation der Unterrichtsqualität im Sportunterricht mit dem QUALLIS-Instrument. *Bewegung und Sport*, 73(2), 12–17.

Herrmann, C., & Gerlach, E. (2020). Unterrichtsqualität im Fach Sport – Ein Überblicksbeitrag zum Forschungsstand in Theorie und Empirie. *Unterrichtswissenschaft*, 48(3), 361–384. <https://doi.org/10.1007/s42010-020-00080-w>

Ho, W. K. Y., Ahmed, Md. D., Khoo, S., Tan, C. H., Dehkordi, M. R., Gallardo, M., Lee, K., Yamaguchi, Y., Tao, Y., & Shu, C. (2019). Towards developing and validating Quality Physical Education in schools—The Asian physical education professionals' voice. *PLOS ONE*, *14*(8), e0218158. <https://doi.org/10.1371/journal.pone.0218158>

Mali, A., & Kumar, P. (2017). Emotional intelligence among secondary school students in relation to their socio-demographic variables as the element for enhancing the development of personality. *International Journal of Advanced Education and Research*, *2*(4), 120-124.

Mottet, T. P., & Beebe, S. A. (2006). The relationships between student responsive behaviors, student socio-communicative style, and instructors' subjective and objective assessments of student work. *Communication Education*, *55*(3), 295-312. <https://doi.org/10.1080/03634520600748581>

National Association for Sport and Physical Education. (copyright 2004). Moving into the future: National standards for physical education (2nd ed.). Reston, VA: Author, pp.12-14.

Neumann, P. (2018). Multiperspective physical education: A phantom of physical education practice? *Sportunterricht*, *67*(7), 290–296. Hofmann-Verlag GmbH & Co. KG.

Okoli C., & Pawlowski S.D. (2004). The Delphi method as a research tool: an exemple, design considerations and applications. *Inf Manag.* *42*, 15–29. <https://doi.org/10.1016/j.im.2003.11.002>

Praetorius, A. K., & Gräsel, C. (2021). Noch immer auf der Suche nach dem heiligen Gral: Wie generisch oder fachspezifisch sind Dimensionen der Unterrichtsqualität?. *Unterrichtswissenschaft*, *49*(2), 167-188. <https://doi.org/10.1007/s42010-021-00119-6>

Praetorius, A. K., Herrmann, C., Gerlach, E., Zülsdorf-Kersting, M., Heinitz, B., Nehring, A. (2020). Unterrichtsqualität in den Fachdidaktiken im deutschsprachigen Raum—zwischen Generik und Fachspezifik. *Unterrichtswissenschaft: Zeitschrift für Lernforschung*, *48*(3), 409-446. <https://doi.org/10.1007/s42010-021-00119-6>

Wang, J., Lin, E., Spalding, E., Klecka, C. L., & Odell, S. J. (2011). Quality teaching and teacher education: A kaleidoscope of notions. *Journal of teacher education*, *62*(4), 331-338. <https://doi.org/10.1177/0022487111409>

Whitehead 1, M. (2001). The concept of physical literacy. *European Journal of Physical Education*, *6*(2), 127-138. <https://doi.org/10.1080/1740898010060205>

Williams, B., & Lee, J. (2020). Experts, expertise and health and physical education teaching: A scoping review of conceptualisations. *The Curriculum Journal*, *32*(1), 14–27. Portico. <https://doi.org/10.1002/curj.94>