ACTIVE METHODOLOGIES, PEDAGOGICAL MODELS AND ASSESSMENT IN PHYSICAL EDUCATION IN PRIMARY EDUCATION: REVEALING THE LACK OF PEDAGOGICAL COHERENCE

METODOLOGÍAS ACTIVAS, MODELOS PEDAGÓGICOS Y EVALUACIÓN EN EDUCACIÓN FÍSICA EN PRIMARIA: DESVELANDO LA FALTA DE COHERENCIA PEDAGÓGICA

Alba Arija-Mediavilla[1], *María Luisa Santos-Pastor*[2], *Luis Fernando Martínez-Muñoz*[3], *Pedro Jesús Ruiz-Montero*[4]

[1] Departamento de Educación Física, Deporte y Motricidad Humana, Facultad de Formación del Profesorado y Educación, Universidad Autónoma de Madrid, España

[2] Departamento de Educación Física, Deporte y Motricidad Humana, Facultad de Formación del Profesorado yEducación, Universidad Autónoma de Madrid, España

[3] Departamento de Educación Física, Deporte y Motricidad Humana, Facultad de Formación del Profesorado y Educación, Universidad Autónoma de Madrid, España

[4] Departamento de Educación Física y Deportiva, Facultad de Ciencias del Deporte, Universidad de Granada, España

Autor para la correspondencia:
Alba Arija Mediavilla, alba.arija@estudiante.uam.es

Título Abreviado:
Active Methodologies and Assessment: Pedagogical Incoherence

How to cite this article:
Arija-Mediavilla, A., Santos-Pastor, M.L., Martínez-Muñoz, L.F. & Ruiz-Montero, P.J. (2025). Active Methodologies, Pedagogical models and Assessment in Physical Education in Primary Education: Revealing the lack of pedagogical coherence. *Cultura, Ciencia y Deporte*, *20*(66), 2358. https://doi.org/10.12800/ccd.v20i66.2358

Received: 20 November 2024 / **Accepted:** 27 March 2025

Abstract

The educational needs of contemporary society place us in a context of methodological renewal. This transformation towards more participatory approaches requires formative assessment processes and a reconsideration of the purpose of grading. In this context, this study focuses on determining whether there is pedagogical coherence in the teaching-learning process by analyzing the relationships between assessment and grading when pedagogical models are implemented in Physical Education. The quantitative study design uses a validated questionnaire called Active Methodologies and Assessment in Physical Education. The study involved 310 Primary School Physical Education teachers. Based on the data, multiple mediation relationships were established between the mediator variable Active Methodologies and the variables Formative Assessment and Grading. The results revealed (1) a significant correlation between variables with pedagogical models such as Service Learning, Attitudinal Style, and the Hybrid model. However, (2) no significant relationships were found in models like Project-Based Learning, Teaching Games for Understanding or the Sports Education model. These findings suggest a week connection between active methodologies and formative assessment in Physical Education classes in Primary Education. This prompts us to reflect on the need to reassess the coherence of teaching-learning processes in order to understand and advance toward integrated and effective pedagogical approaches.

Keywords: Formative assessment, grading, physical education, pedagogical models, teaching-learning process.

Resumen

Las necesidades educativas de la sociedad actual nos sitúan en un escenario de renovación metodológica. Esta transformación hacia enfoques más participativos en los que el alumnado es protagonista, requiere de procesos de evaluación formativa y la reconsideración de la finalidad atribuida a la calificación. En este contexto, el presente estudio se centra en determinar si existe coherencia pedagógica en el proceso de enseñanza-aprendizaje, analizando las relaciones que se establecen entre los procesos de evaluación y calificación utilizados cuando se ponen en práctica los modelos pedagógicos en Educación Física en Primaria. El estudio de corte cuantitativo emplea un cuestionario validado denominado Metodologías activas y Evaluación en Educación Física en Primaria. En él participaron 310 docentes en activo de Educación Física de Primaria. A partir de los datos obtenidos se establecieron relaciones de mediación múltiple entre la variable mediadora Metodología Activa, y las variables Evaluación Formativa y Calificación. Los resultados mostraron que: (1) ante modelos pedagógicos como el Aprendizaje-Servicio, el Estilo Actitudinal, y la Hibridación de modelos, existían correlaciones significativas entre variables. Sin embargo, (2) no se encontraron relaciones significativas en modelos como el Aprendizaje Basado en Proyectos, el modelo Comprensivo de Iniciación Deportiva o el modelo de Educación Deportiva. Estos hallazgos nos permiten concluir que la vinculación entre metodologías activas y evaluación formativa en las clases de Educación Física en Primaria es muy débil, lo que nos lleva a reflexionar sobre la necesidad de revisar la coherencia de los procesos de enseñanza-aprendizaje para comprender y avanzar hacia propuestas pedagógicas integradas y efectivas.

Palabras clave: Calificación, evaluación formativa, educación física, modelos pedagógicos, proceso de enseñanza-aprendizaje.

Introduction

The educational paradigm shift initiated in Spain with the General Organization of the Educational System Organic Law (LOGSE, 1990), and continued with the incorporation of competencies into educational curricula (LOE, 2006 and subsequent), advocates for the comprehensive education of students to adapt to new demands and face the challenges of today's society. Currently, the LOMLOE (2020) proposes a new competence and curriculum approach that responds to the demands of the 2030 Agenda and the Sustainable Development Goals (Pérez-Pueyo et al., 2022). In this scenario, to respond to the new challenges of the knowledge society (UNESCO, 2021), a real transformation of the teaching-learning process is necessary through methodological and evaluation approaches in which students are an active part of the process (Coll & Martín, 2021; Otero-Saborido et al., 2023).

Pedagogical Models. The Methodological Change in Physical Education

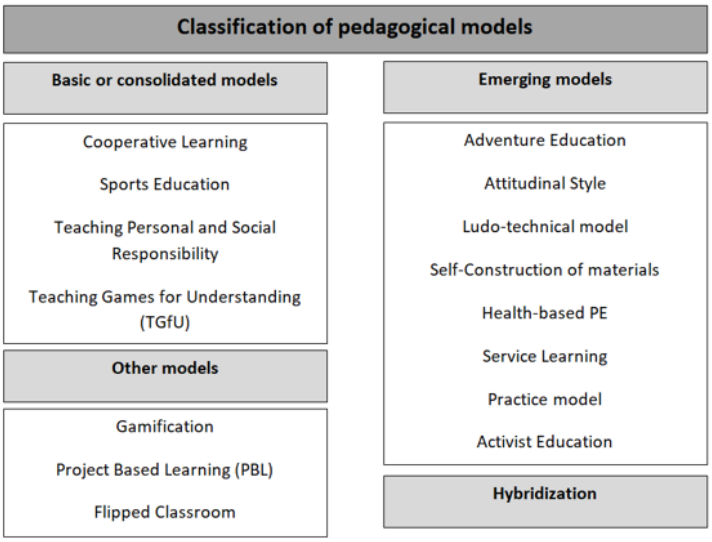
In recent decades, from a methodological standpoint, Physical Education (PE) has undergone changes that have allowed it to adapt to the requirements of today's society (León-Díaz et al., 2023). In this social and regulatory context, teaching approaches centered on directive and analytical didactic strategies are being superseded by alternative teaching-learning approaches that give students a clear leading role. We highlight active methodologies, specified in the pedagogical models within the field of PE (León-Díaz et al., 2020; MacPhail et al., 2018).

Understood as long-term structured and systematic approaches, focused on achieving learning objectives in four areas (physical, academic, social, and affective), pedagogical models generate deep, meaningful, and lasting learning (Fernández-Río & Iglesias, 2022). Furthermore, they facilitate the transfer of learning by placing students in contact with the problems that surround them (Fernández-Río et al., 2016; Pérez-Pueyo et al., 2021). Thus, we can define them as a theoretical-practical framework that guides the teaching-learning process, based on active pedagogical theories and principles (active methodologies) that seek the comprehensive development of students.

Based on the scientific community's agreement on the principles of pedagogical models, as detailed in Pérez-Pueyo et al. (2021), thirteen distinct types are identified, categorized into two primary blocks: (1) basic or consolidated pedagogical models and (2) emerging models, which respond to the evolution of PE (Fernández-Río et al., 2018). To these two main groups, a third should be added, where we would frame the hybridization of the previous models (Casey & McPhail, 2018; González-Víllora et al., 2019; Haerens et al., 2011; Shen & Shao, 2022). In the first group, we find: Cooperative Learning, Sports Education (Guijarro et al., 2020), Teaching Personal and Social Responsibility and Teaching Games for Understanding (Galeano-Rojas et al. 2023; Kirk & McPhail, 2002). Within the group of emerging models, the following would be encompassed: Adventure Education (Baena-Extremera, 2011), Attitudinal Style (Pérez-Pueyo, 2016), the Ludo-technical model, Self-Construction of materials, Health-based PE (Díaz-Tejerina & Fernández Río, 2024), Service Learning (Calvo-Varela et al., 2019; Chiva-Bartoll & Fernández-Río, 2022), the Practice model, and the Activist Education pedagogical model. Other authors also include, within this classification, other models not specific to the PE area: Consolidated models, such as Project-Based Learning (PBL), or emerging models like Gamification or the Flipped Classroom (Arufe-Giráldez et al., 2023) (Figure 1).

Figure 1

Classification of Pedagogical Models



Note. Figure adapted from Pérez-Pueyo et al. (2021).

Formative Assessment: the First Methodological Change

For the methodological transition, from directive proposals to proposals that actively involve students, to be carried out coherently, it must be accompanied by formative assessment processes linked to a redefinition of the purpose attributed to grading (Pérez-Pueyo et al., 2021). In recent decades, different authors have highlighted the importance of formative and shared assessment (FSA) as the first methodological change to be implemented in the transformation of the teaching-learning process (López-Pastor & Pérez-Pueyo, 2017; Pérez-Pueyo & Hortigüela-Alcalá, 2020; Pérez-Pueyo et al., 2020). In this context, evaluation must be consistent with the methodological approach and allow students to know in advance: what do I have to learn?, where am I in the learning process?, what do I have to do to improve? (López-Pastor, 2006; Hortigüela-Alcalá et al., 2019).

Thus, to implement FSA processes, it is necessary to start from a democratic approach that grants students voice and responsibility (Aarskog, 2021). In other words, that the agents involved in the evaluation process become aware so that students learn more and teachers can adapt their practice to the needs of the classroom (Barrientos-Hernán et al., 2019). Formative assessment is, therefore, a process whose purpose is to obtain information to improve the teaching-learning process (Pérez-Pueyo et al., 2024). To achieve this objective, students must be active agents in the process, participating in self-assessment and peer-assessment proposals that allow them to reflect and self-regulate. Thus, in addition to being formative, the evaluation must be shared. In this FSA process, a series of fundamental steps must be considered: 1) initial presentation to students about learning objectives, assessment activities, assessment instruments, and grading criteria; (2) establishment of a process timeline; (3) formative feedback with shared assessment procedures (self-assessment, peer assessment); (4) conducting assessment activities leaving room for learning improvement; and (5) making the leap to formative grading, understood as the translation of a teaching-learning process into a grade, which will reflect the student's learning process as fairly and rigorously as possible (Hortigüela-Alcalá et al., 2019; Hortigüela-Alcalá et al., 2020; Pérez-Pueyo et al., 2020).

Formative Grading, the end of the Process

The term formative grading is not explicitly found in the bibliography, as it is considered inherent to FSA processes. As noted by López-Pastor and Pérez-Pueyo (2017): The grading system used can either reinforce the designed learning process or undermine it. Therefore, it is advisable to think very carefully about how to do it. For these purposes, when we talk about formative grading, we are referring to a process integrated within a broader approach to assessment for learning, which ultimately aims to improve student learning above the final grade (Brookhart, 2011). Therefore, we must start from the premise that grading is not synonymous with assessing, especially when we are talking about formative assessment

processes. In this sense, for grading to be consistent with the FSA model, aspects such as active student participation (self-grading), consensus and dialogue (peer-grading and dialogic grading) must be reflected inherently in the process. Thus, we can understand grading as the final part of the assessment process and, ultimately, a teaching obligation where the entire process carried out is reflected. The way both processes relate to each other determines the purpose of the pedagogical proposal (whether it's an assessment aimed at learning or an assessment aimed at grading) (Barba-Martín et al., 2020).

Despite advances in the assessment processes used by teachers, the idea of assessment as a tool to measure student performance persists (Moura et al., 2021; Oliver-Álvarez & Martos-García, 2023). It is common to find teachers who carry out FSA processes, but who, when it comes to making the leap to grading, do not take into account the decisions made previously, which causes them to incur summative and final grading processes typical of the traditional assessment model (Hortigüela-Alcalá et al., 2019).

Active Methodologies, Formative Assessment, and Grading in Physical Education

The relationship between these three processes (active methodologies, assessment, and grading), as well as their pedagogical importance from a social and regulatory point of view, is widely recognized in the educational literature, as it has been detailed in the text. However, there is hardly any evidence that reveals what is happening in classrooms.

Hortigüela-Alcalá et al. (2020) y Pérez-Pueyo et al. (2020) propose a way to integrate FSA into each of the pedagogical models. Thus, they detail the basic steps for its integration with specific instruments and examples, adapted to the specific characteristics and foundations of each model.

In this line, but in the university context, Santos-Pastor et al. (2019, 2024) present a successful experience, combining the Service Learning model in physical activity with FSA proposals. Among the main conclusions, the improvement in the quality of the work stands out, thanks to individualized monitoring and feedback, as well as greater significance and functionality of the learning. These results reappear in the study by Abella-García et al. (2020), who analyze a successful experience in the university context, in which they propose a PBL approach with formative assessment strategies. In the Primary Education stage, we find the proposal by Cañas-Encinas et al. (2019), who incorporate the FSA system into a gamified project in PE, obtaining similar conclusions. For their part, Herrero-González et al. (2020), in their study of a group of Primary PE teachers who implemented Cooperative Learning in their classes, demonstrated that they used FSA in their proposals when they implemented this methodology, reinforcing the idea of Alonso et al. (2017) of the integration of FSA into the components and essence of this pedagogical model.

In short, although it is not explicitly mentioned in the research, it seems consistent to affirm that, in a context where pedagogical models are being implemented within the framework of active methodologies and FSA processes, grading, as the final part of the teaching-learning process, should be consistent.

With this research, we intend to approach the practical reality of Primary PE teachers who implement active methodologies, with the aim of understanding what relationship is established between the evaluation and grading processes used when they put pedagogical models into practice, and therefore, whether they connect with FSA processes, associated with democratic and shared grading models. In essence, the research aims to establish if consistency exists between the methodology and the assessment processes employed in Primary Physical Education.

Material y Methods

Design

This study is framed within a quantitative correlational research approach, using a validated questionnaire as a data collection instrument aimed at identifying the methodological approach and the evaluation and grading processes in Primary PE (Active Methodologies and Assessment in Primary PE questionnaire (Spanish acronym: MAEV-EFP). The research has at all times taken into account the ethical and data processing procedures approved after review by the ethics committee of the Autonomous University of Madrid (CEI-131-2718).

Participants

For the selection of participants, a non-probabilistic incidental sampling was used. A sample of 310 PE teachers (59% men and 41% women) was obtained, all of whom were active (inclusion criterion), with representation from all Autonomous Communities in the national territory. Contact with the study participants was made via email, after prior contact with the educational centers. As Table 1 shows, male teacher participation predominates in the present study, while approximately 55% of the total sample is over 41 years old. Likewise, more than 55% of the total sample has been teaching for more than 13 years. Regarding the grade levels in which teaching is carried out, it has been considered that each participant may be assigned teaching in more than one grade, therefore the percentages are cumulative. Ultimately, the overwhelming majority of educational institutions where the participating teachers in this study work are public schools.

Table 1

Sociodemographic Characteristics of the Participants (n = 310)

Variable	Number (Percentage)
Gender	
Female	127(41)
Male	183(59)
Age	
20-30	43(13.9)
31-40	98(31.6)
41-50	124(40)
Over 51 years old	45(14.5)
Teaching experience	
Less than 1 year	8(2.6)
1-3 years	29(9.4)
4-6 years	35(11.3)
7-12 years	64(20.6)
13-18 years	65(21)
More than 18 years	109(35.2)
Grade level they teach a	
1st	170(54.8)
2nd	185(59.7)
3rd	189(61)
4th	201(64.8)
5th	205(66.1)
6th	202(65.2)
School ownership	
Public school	258(83.2)
Charter school	48(15.5)
Private school	4(1.3)

a. Each teacher may teach in more than one grade.

Instruments

For the development of the research, the «ad hoc» questionnaire called Active Methodologies and Assessment in Primary PE (Spanish acronym: MAEV-EFP) (Arija-Mediavilla et al., 2024) was used. Following an extensive review of the scientific literature on active methodologies, pedagogical models, evaluation, and grading, a list of references and characteristics was developed. This allowed us to define the foundation of each of the variables and to make adjustments according to the purpose of the questionnaire. The first version of the questionnaire underwent a validation process by 11 experts following the Delphi method validation structure. The experts were selected according to four inclusion criteria: (1) willingness and interest to participate in the study; (2) availability; (3) level of experience regarding the subject of study; and (4) level of knowledge about the topic (López-Gómez, 2018). The resulting version of the questionnaire underwent a double pilot process to calculate the initial reliability and validity of the instrument. Firstly, a small sample of active Primary PE teachers (16 teachers) was asked to assess the suitability of the items based on three parameters: (1) understanding of the question and its purpose; (2) correct and clear wording; and (3) simple answer. Once the suitability of the questionnaire was verified in

a real-world application context, a single-case validation test was conducted. The resulting questionnaire and the validation process can be consulted in Arija-Mediavilla et al. (2024).

This questionnaire consists of four dimensions (personal data, methodology, assessment, and grading) and 12 variables. The questions are designed to identify the frequency of use of each subvariable by the teachers. In the present research, from the various dimensions and variables of the questionnaire, the following have been analyzed: all subvariables of the Personal data dimension (gender, age, experience, teaching grade level, and school ownership); the active methodologies used (M) variable from the Methodology dimension; the assessment model (X) variable from the Assessment dimension; and the grading model (Y) variable from the Grading dimension (Table 2).

Table 2

Dimensions, Variables, Analyzed Subvariables, and Response Typology

Dimension	Variable	Subvariables	Response
Methodology	Active methodologies (M)	M1= Cooperative Learning, M2= Project-Based Learning (PBL), M3= Flipped Classroom, M4= Learning Environments, M5= Service Learning, M6= Hybrid model, M7= Attitudinal Style, M8= Teaching Games for Understanding, M9=Self-Construction of materials, M10= Adventure Education, M11= Health-based PE, M12= Sports Education	Likert Scale (1 Never - 4 Always)
Assessment	Assessment model (X)	Learning-oriented assessment, where the goal, through assessment, is for students to be aware of their learning and improve it (formative, continuous, integrated into daily activities, and shared with students)	Likert Scale (1 Never - 4 Always)
Grading	Grading model (Y)	Y1=Formative Grading by assigning grades to various instruments used and averaging by didactic programming unit (DPU), Y2= Formative Grading through a consensual dialogue with students to make a common decision, Y3= Formative Grading by assigning grades to Self-Assessment instruments, attributing a percentage to the final grade, Y4=Formative Grading by assigning grades to peer-assessment instruments and attributing a percentage to the final grade	Likert Scale (1 Never - 4 Always)

Note: DPU: didactic programming unit

The subvariables selected in the active methodologies variable and that related to the FSA model respond to the consensus of the expert group and the research team, taking as a reference the classification collected in León-Díaz et al. (2020) and its adaptation to the Primary stage. The subvariables related to the grading model were, in turn, derived from the foundational principles of the formative assessment process, to ensure the grading transition is consistent and true to those principles. In this sense, the aim is that grading is not only a measurement instrument, but that it remains aligned with the democratic and pedagogical foundations that characterize the FSA. This implies that grading criteria not only reflect student academic performance, but also their learning process, promoting their participation in the process.

Procedure

The data used in the research were obtained through two rounds of questionnaire distribution, one carried out in the 2021/2022 academic year and the other in the 2022/2023 academic year. For the selection of the sample, the census of educational centers of Early Childhood and Primary Education from the different Autonomous Communities was used. Emails were sent to the institutional addresses of the educational centers, asking for the questionnaire to be distributed among their active PE teachers. To maximize participant reach, the questionnaire was administered via the Google Forms platform, using a single response format. Participants were informed at all times about confidentiality and data processing protocols, and were required to give their consent to proceed with completing the questionnaire.

Statistical Analysis

All statistical analyses were performed using the SPSS 25 statistical package for Social Sciences (IBM SPSS for MAC, Armonk, NY).

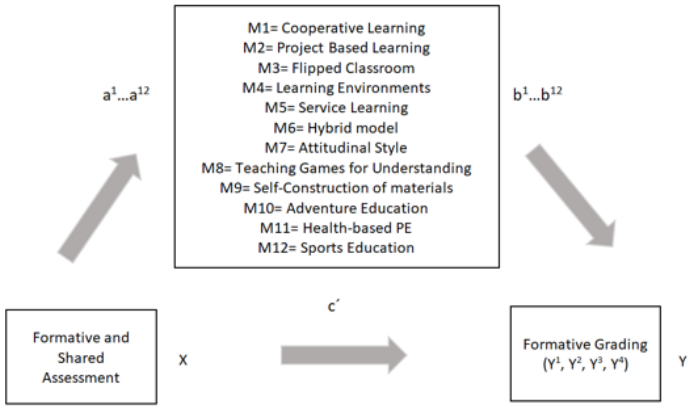
Descriptive statistical analysis. To show the sociodemographic characteristics of the participating teachers regarding the assessment and grading used in their daily practice, response frequencies (number and percentage) were calculated according to gender, age group (20-30 years, 31-40 years, 41-50 years, and > 51 years), teaching experience (< 1 year, 1-3 years, 4-6 years, 7-12 years, 13-18 years, and > 18 years), the grade level they teach in Primary Education (1st, 2nd, 3rd, 4th, 5th, and 6th), and the ownership of the educational center where they work (public, subsidized, or private).

Non-parametric statistical analysis and multiple correlation. The Kolmogorov-Smirnov test was used to determine the normality of the data, with the result being non-parametric. Correlation analyses were performed using a Spearman correlation test to determine the associations between the FSA variables, the four types of formative grading explained in the method section, and the 12 active methodologies used by the participants. The significance level was set at $p < .005$, showing strong and positive correlation in values close to 1 and strong and negative correlation in values close to -1.

Multiple mediation statistical analysis. This statistical method seeks to examine how several simultaneous mediating variables (active methodologies) explain the relationship between an independent variable (formative assessment) and a dependent variable (formative grading), decomposing the total effect into direct effects (relationship between independent and dependent variables, without intervention of mediators) and indirect effects (relationship between independent and dependent variables passing through mediating variables), providing a more detailed understanding of the underlying causal mechanisms in the relationships between variables (MacKinnon et al., 2012). This statistical analysis, using a bootstrap correction, was performed to analyze the mediation of the active methodologies used in PE class (M1...M12) in the relationship between the FSA (X) and four different types of formative grading (Y1...Y4) (Figure 2).

Figure 2

Hypothetical Model of the Relationship Between FSA (X) and Four Types of Formative Grading (Y1, Y2, Y3, Y4) in the Present Study



Note: a=Indirect path between FSA (a1...a12) and each of the 12 active methodologies (M1...M12); b=Indirect path between each of the 12 active methodologies (M1...M12) and formative grading (b1...b12); c´=Direct path of relationship between FSA and formative grading.

A 5,000 bootstrap was used to analyze multiple mediation, with a bias-corrected and accelerated 95% confidence interval (CI). The bootstrap technique is used to calculate estimates of direct and indirect effects in multiple mediation models. Bootstrap percentile-based CIs were examined to determine the type of mediation applied. Lack of significance is observed

when the confidence interval, defined by its minimum and maximum values, includes zero (Preacher & Hayes, 2008). Additionally, collinearity was checked with VIF (Variance Inflation Factor) values ranging from 1,135 to 1,467, and the Durbin-Watson autocorrelation coefficient with values never below 1. These data highlight the absence of significant collinearity or autocorrelation problems in the model, reinforcing the robustness of the statistical analysis assumptions.

Results

Multiple correlation statistical analysis. Table 3 presents the results of the correlation between FSA, 12 active methodologies used in the present study, and four types of formative grading. In general terms, most variables show correlation with each other ($p < .005$).

Table 3

Correlation Established Between FSA, Active Methodologies Used, and Formative Grading

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1. AM1																
2. AM2	.184**															
3. AM3	.231**	.196**														
4. AM4	.254**	.261**	.258**													
5. AM5	.296**	.358**	.223**	.338**												
6. AM6	.109*	.253**	.231**	.300**	.310**											
7. AM7	.121*	.138*	.173**	.321**	.317**	.288**										
8. AM8	.176**	.177**	.161**	.237**	.217**	.180**	.351**									
9. AM9	.101	.260**	.243**	.194**	.248**	.135*	.090	.210**								
10. AM10	.172**	.246**	.142*	.298**	.298**	.108*	.301**	.165**	.277**							
11. AM11	.240**	.111*	.132*	.262**	.221**	.143	.278**	.189**	.261**	.354**						
12. AM12	.080	-.096	.071	.192**	.067	-.019	.153**	.374**	.077	.147*	.309**					
13. FSA	.141*	.116*	.089	.075	.161**	.157**	.127*	.142*	.205**	.118*	.286**	.083				
14. FG1	.158**	.086	.146*	.199**	.325**	.063	.237**	.046	.061	.214**	.219**	.112*	.025			
15. FG2	.195**	.160**	.207**	.265**	.261**	.131*	.252**	.198**	.057	.132*	.112*	.098	.188**	.317**		
16. FG3	.165**	.153**	.240**	.204**	.217**	.032	.149**	.153**	.140*	.147**	.114*	.095	.074	.392**	.411**	
17. FG4	.198**	.225**	.289**	.225**	.268**	.078	.152**	.200**	.237**	.172**	.140*	.118*	.147**	.309**	.447**	.762***

Note. AM: Active Methodology; AM1: Cooperative Learning; AM2: PBL; AM3: Flipped Classroom; AM4: Learning Environments; AM5: Service Learning; AM6: Hybrid model; AM7: Attitudinal Style; AM8: Teaching Games for Understanding; AM9: Self-Construction of materials; AM10: Adventure Education; AM11: Health-based PE; AM12: Sports Education; FSA: Formative and shared assessment; FG: Formative Grading; FG1: Grade different instruments; FG2: Consensus-based dialogue with students ; FG3: Assign grades to self-Assessment instruments; FG4: Assign grades to peer-assessment instruments.

Multiple mediation statistical analysis. Table 4 presents the results obtained from the multiple mediation analysis, detailing the indirect effect of FSA (X) on the four subvariables that reflect the formative transition from assessment to grading (Y1-Y4), through the 12 active methodologies used by the teachers (M).

We present the results obtained, collected in Table 4, in three blocks: (1) path from FSA (X) to the mediators of the 12 active methodologies (M); (2) path from the 12 active methodologies (M) to the 4 types of grading (Y1-Y4); and (3) mediation of the effect of FSA (X) on the four types of grading mediated by the active methodologies.

(1) On one hand, the results obtained from the path from FSA (X) to the mediators (M) corresponding to each of the 12 active methodologies used show significance ($p < .05$) in six active methodologies: Service Learning (AM5), Hybrid model (AM6), Attitudinal Style (AM7), Self-Construction of materials (AM9), Adventure Education (AM10) and Health-based PE (AM11). On the other hand, the mediation relationships were not significant in the following methodologies: PBL (AM2), Flipped Classroom (AM3), Teaching Games for Understanding (AM8) and Sports Education (AM12).

(2) In the path between the 12 active methodologies (M) towards each of the four types of formative grading (Y1, Y2, Y3, Y4), significance was again shown in six types of methodologies ($p < .05$), where Service Learning (AM5) showed the highest frequency of significance.

(3) Finally, regarding the total mediation of the estimated effect of FSA (X) on the four types of grading (Y), as measured by the 12 active methodologies, the results were as follows: Regarding Formative Grading assigning grades to different instruments used and averaging by DPU (Y1), the total mediation of the estimated effect was 70.5% in Service Learning (AM5), 46.5% in Attitudinal Style (AM7), 36.67% in Adventure Education (AM10), and 70.8% in Health-based PE (AM11). Concerning Formative Grading via collaborative dialogue with students for shared decision-making (Y2), the estimated mediation was 12.3% in Service Learning (AM5) and 11.2% in Attitudinal Style (AM7). As for the third variable, Formative Grading by assigning grades to self-assessment instruments, attributing a percentage to them in the final grade (Y3), the total mediation of the estimated effect was 34.4% for Service-Learning (AM5) and 28.1% for Hybrid model (AM6). Finally, regarding Formative Grading by assigning grades to peer-assessment instruments and attributing a percentage to them in the final grade (Y4), the estimated mediation was 18.9% in Service Learning (AM5), 12.4% in Hybrid model (AM6), and 20.7% of the total effect of FSA (X) in Self-Construction of materials (AM9).

These results highlight the existence of a substantial coherence between formative assessment processes and grading when the following active methodologies are implemented: Service Learning (AM5), Hybrid model (AM6), Attitudinal Style (AM7), Self-Construction of materials (AM9), Adventure Education (AM10) and Health-based PE (AM11).

Table 4

Indirect Effect of FSA (X) on Four Different Types of Formative Grading (Y1, Y2, Y3, Y4), Through 12 Active Methodologies, all Used by the Participating Teachers in the Present Study

Mediator	Effect of X on M (a1-a12)	p-value	SE	Effect of M on Y (b1- b12)	p-value	SE	Bootstrap estimate	SE	95% CI	
									Below	Above
	Y1 = Formative Grading by assigning grades to different instruments used and averaging by DPU.									
AM1	.133	.005	.047	.058	.582	.105	.007	.015	-.019	.046
AM2	.111	.069	.061	-.078	.348	.083	-.008	.012	-.045	.006
AM3	.107	.074	.060	.111	.178	.082	.012	.012	-.003	.049
AM4	.082	.208	.065	.140	.078	.079	.011	.013	-.005	.053
AM5	.191	.004	.066	.392	.000	.079	.075	.033	.018	.152
AM6	.214	.005	.075	-.079	.234	.066	-.016	.016	-.063	.006
AM7	.161	.029	.074	.202	.005	.071	0.32	0.21	.003	.087
AM8	.195	.008	.074	-.087	.238	.073	-.017	.016	-.064	.007
AM9	.216	.000	.061	.005	.945	.083	.001	.018	-.033	.041
AM10	.144	.031	.066	.178	.025	.079	.025	0.18	.000	.079
AM11	.292	.000	0.58	.170	.067	.092	0.49	.028	.002	.117
AM12	.116	.085	.067	0.58	.456	0.78	.006	.012	-.009	.047

Table 4 (cont.)

Indirect Effect of FSA (X) on Four Different Types of Formative Grading (Y1, Y2, Y3, Y4), Through 12 Active Methodologies, all Used by the Participating Teachers in the Present Study

Mediator	Effect of X on M (a1-a12)	p-value	SE	Effect of M on Y (b1- b12)	p-value	SE	Bootstrap estimate	SE	95% CI	
									Below	Above
	Y2= Formative Grading through a consensual dialogue with students to reach a common decision.									
AM1	.133	.005	.047	.094	.333	.097	.012	.014	-.011	.050
AM2	.111	.069	.061	.008	.916	.076	.000	.010	-.016	.025
AM3	.107	.074	.060	.120	.115	.076	.013	.012	-.002	.050
AM4	.082	.208	.065	.205	.005	.073	.017	.018	-.008	.067
AM5	.191	.004	.066	.168	.022	.073	.032	.021	.003	.092
AM6	.214	.005	.075	-.036	.556	.061	-.007	.015	-.043	.020
AM7	.161	.029	.074	.182	.006	.066	.029	.019	.003	.084
AM8	.195	.008	.074	.094	.169	.068	.018	0.16	-.005	.062
AM9	.216	.000	.060	-.081	.293	.077	-.017	.019	-.065	.014
AM10	.144	.031	.066	.097	.186	.073	.014	.014	-.004	.058
AM11	.292	.000	.058	-.030	-.719	.085	-.009	.026	-.060	.045
AM12	.116	.085	.067	.044	.541	.073	.005	.011	-.009	.043
	Y3= Formative Grading by assigning grades to self-assessment instruments, attributing a percentage to them in the final grade.									
AM1	.133	.005	.047	.070	.465	.096	.009	.013	-.012	.044
AM2	.111	.069	.061	.046	.540	.076	.005	.010	-.008	.037
AM3	.107	.074	.060	.263	.000	.075	.028	.020	-.001	.081
AM4	.082	.208	.065	.136	.061	.072	.011	.012	-.005	.049
AM5	.191	.004	.066	.179	.014	.072	.034	.019	.006	.087
AM6	.214	.005	.075	-.130	.031	.060	-.028	.018	-.076	-.002
AM7	.161	.029	.074	.091	.172	.067	.014	.014	-.003	.058
AM8	.195	.008	.074	.066	.334	0.68	.013	.015	-.010	.055
AM9	.216	.000	.060	.088	.258	.077	.019	.019	-.012	.065
AM10	.144	.031	.066	.083	.264	.074	.012	.014	-.006	.052
AM11	.292	.000	.058	-.002	.973	.086	-.001	.026	-.052	.055
AM12	.116	.085	.067	.059	.421	.073	.006	.011	-.008	.042
	Y4= Formative Grading by assigning grades to peer-assessment instruments and attributing a percentage to them in the final grade.									
AM1	.133	.005	.047	.089	.338	.093	.012	.013	-.008	.047
AM2	.111	.069	.061	.102	.166	.073	.011	.011	-.002	.049
AM3	.107	.074	.060	.277	.000	.073	.029	.020	-.001	.085
AM4	.082	.208	.065	.124	.078	.070	.010	.012	-.004	.049
AM5	.191	.004	.066	.207	.003	.070	.039	.022	.006	.096
AM6	.214	.005	.075	-.121	.038	.058	-.026	.018	-.077	-.001
AM7	.161	.029	.074	.105	.106	.065	.017	.015	-.001	.061
AM8	.195	.008	.074	.089	.180	.066	.017	.017	-.004	.068
AM9	.216	.000	.060	.201	.008	.075	.043	.022	.010	.101
AM10	.144	.031	.066	.069	.337	.072	.010	.013	-.007	.047
AM11	.292	.000	.058	-.035	.674	.084	-.010	.025	-.062	.037
AM12	.116	.085	.067	.065	.361	.071	.007	.011	-.006	.045

Note. Based on a 5,000-bootstrap sample. X= FSA; Y= Formative Grading categorized into four types (Y1, Y2, Y3, Y4); M= Mediator; CI= Confidence Interval; SE= Standard Error; DPU= Didactic Programming Unit; AM: Active Methodology; AM1: Cooperative Learning; AM2: PBL; AM3: Flipped Classroom; AM4: Learning Environments; AM5: Service Learning; AM6: Hybrid model; AM7: Attitudinal Style; AM8: Teaching Games for Understanding; AM9: Self-Construction of materials; AM10: Adventure Education; AM11: Health-based PE; AM12: Sports Education. If CI contains zero, it is interpreted as not significant.

Discussion

The implementation of active methodologies through pedagogical models, both in education in general and in PE in particular, responds to the requirements of current society and a curriculum articulated around the competency

development of students. However, the inclusion of these proposals lacks meaning if it is not implemented with FSA processes accompanied by a congruent grading system as the final part of the process (Pérez-Pueyo et al., 2021). This study has reviewed the mediation of active methodologies, through pedagogical models implemented by Primary PE teachers, in the relationship between FSA and grading processes.

The obtained results open two lines of discussion. On one hand, we find (1) low significance in the relationships between certain active methodologies and FSA processes, highlighting a lack of coherence in the process, and on the other hand, (2) the active methodologies that do mediate significantly.

The didactic strategies of PBL (AM2) and Flipped Classroom (AM3) do not show significant relationships with the FSA model, leading us to infer the use of traditional assessment models and, consequently, a summative and final conception of grading (Hortigüela-Alcalá et al., 2019). PBL is one of the methodologies that gives the most prominence to students, providing them with a real-world learning context and directly and actively involving them in the process (Abella-García et al., 2020). Aspects such as autonomy and self-regulation, key in this methodology, require self-assessment and peer-assessment processes typical of FSA. In this regard, this disconnection found in the results concludes in a lack of coherence in the process that may be due to factors such as: lack of training, students' lack of experience, uncertainty when delegating part of the responsibility, or the prioritization of the result over the process (León-Díaz et al., 2020). In the specific case of the Flipped Classroom, this methodology does not inherently provide meaning to learning and can be understood as a tool that can be integrated into other models as it focuses on the presentation of content and its subsequent work in the classroom, rather than on learning itself (Willis et al., 2023). In this regard, the design of formative assessment in these contexts can be more challenging and less evident in its implementation, which could explain why it is not systematically observed in teaching practice.

Following this first line of discussion, we find that Teaching Games for Understanding (AM8) and Sports Education (M12) do not show significant relationships with the FSA model either. Although the scientific literature shows a paradigm shift in sports teaching, the traditional approach remains latent in the teaching of sports in school PE. Fernández-Río and Iglesias (2022) note that the Sports Education model is the most frequently used educational proposal for teaching and learning sports. However, in this model, the acquisition of sports skills in content-focused game situations takes precedence over competency-based learning (Arufe-Giráldez et al., 2023). For its part, although the Teaching Games for Understanding model prioritizes personal and social development (Kirk & McPhail, 2002; Pérez-Pueyo et al., 2020), its teaching focuses on performance and the development of tactical skills and knowledge based on game situations (Galeano-Rojas et al., 2023). This orientation of sports learning models, where value is placed on the acquisition of sports skills over competency-based learning, could support the fact that the most frequently used assessment models are traditional in nature. As Moura et al. (2021) point out, PE teachers seem to use assessment to grade motor progress but not motor competence. In this context, where performance is measured and practice is valued from a quantitative orientation, FSA processes are relegated to the background.

In addition to the specific factors of each model that may be conditioning the lack of significance in the relationships with FSA, the lack of teacher training regarding the implementation of FSA processes (Herrero-González et al., 2021) leads to the application of methodological strategies lacking purpose by intentionally not framing the assessment process as an initial step (Pérez-Pueyo et al., 2021).

On the other hand, the results report significant mediations between FSA and the linked grading when six active methodologies are used, specifically: Service Learning (AM5), Hybrid model (AM6), Attitudinal Style (AM7), Self-Construction of materials (AM9), Adventure Education (AM11) and Health-based PE (AM12).

The Service Learning model shows the highest mediation results with an effect on all four types of grading (70.5%, 12.3%, 34.4%, and 18.9% respectively). These results are supported by the very essence and structure of the methodology, which gives students an active role, involving them in all decisions and actions of the process. As Santos-Pastor et al. (2019, 2024) point out, in Service Learning projects, assessment cannot be reduced to a final grade; assessment must accompany students throughout all phases of the process (diagnosis, design, implementation, reflection, learning evaluation, and service evaluation). Self-assessment, peer-assessment, and shared dialogue processes are essential in this methodology to enhance reflection on experiences, contributing to the improvement of the teaching-learning process and, ultimately, the quality of the project (Calvo-Varela et al., 2019). On the other hand, the relationship of this methodology with grading processes

in which different instruments are used would align with formative actions to link the academic and the social contexts, allowing the leap from FSA processes to grading (Chiva-Bartoll & Fernández-Río, 2022).

Meanwhile, the Attitudinal Style acts as a mediator with an impact on grading when grades are assigned to the various instruments employed, with the average determined per DPU or learning situation (46.6%), and in the grading discussed with students (11.2%). This methodology is structured on three components, with a clear foundation of constant feedback, which seeks to address students' errors and needs, with the aim of providing positive and successful experiences to achieve greater student motivation (Pérez-Pueyo, 2016). Formative assessment is a fundamental pillar of the Attitudinal Style, given the importance of student participation in their learning process (self-assessment), in the learning process of others (peer-assessment), and in collective decision-making based on dialogue (Álvarez-Sánchez et al., 2025; Herrero-González et al., 2020; Pérez-Pueyo et al., 2020). These considerations support the rationale for merging active methodologies and FSA in order to give meaning to learning and the acquisition of competencies. However, in terms of grading, they do not completely align with the results found in the study. Despite the presence of dialogue-based assessment and grading processes, the results do not show significant relationships when self-assessment and peer-assessment processes are applied. Grading is associated with the use of different instruments (although in an isolated and partially meaningful way), which could be explained by a lack of knowledge about the practical implementation of an integrated FSA system consistent with the aims of an active methodology that focuses on the competency development of students (MacPhail et al., 2018).

The Adventure Education and Health-based PE models mediate by 36.7% and 70.8% respectively, when grades are given to different instruments used and the average is calculated for didactic units or learning situations. Firstly, this mediation relationship should be questioned in the absence of a qualitative analysis of the data that would allow us to know the type of instruments teachers use and how they use them, since an assessment instrument can be formative or not depending on the purpose attributed to it in the process of use (López-Pastor & Pérez-Pueyo, 2017). The predominance of experiential learning in a context of risk (objective or subjective), based on motor challenges with different levels of difficulty within the Adventure Education model, implies a greater sequencing of achievement levels and the adoption of a greater number of safety measures. This aspect can influence the assessment process, leading teachers to opt for a more summative than formative approach, focused on the final result rather than on the process (Baena-Extremera, 2011). The ultimate goal of the Health-based PE model is for students to value having an active life and to transfer the acquired learning to their leisure time. For this to happen, it is necessary that, according to the psychological theory of self-determination, the needs for autonomy, competence, and relatedness are satisfied. Student-centered approaches and assessment proposals will be essential to allow students to acquire a sense of competence (Díaz-Tejerina & Fernández-Río, 2024). This model is one of the most recent and, therefore, with less development, as it was in 2011 when Haerens et al. (2011) laid its foundations. This aspect directly influences teacher training due to a smaller number of successful experiences and available scientific research, which leads us to question the validity of the results obtained (Fernández-Río et al., 2016).

The mediation of the Self-Construction of materials model with FSA on formative grading, when grades are given to peer-assessment instruments and assigned a percentage in the final grade, is reinforced by three principles of constructivism on which it is based: students learn by doing, in groups, and by designing and inventing (Fernández-Río et al., 2016). Thus, given the time constraints and the characteristics of the construction process and the material, several of the proposals based on the Self-Construction of materials model are carried out in groups. The potential for hybridization of this model with others leads to its introduction as an initial group activity linked to material and process construction and evaluation strategies based on cooperative learning. Furthermore, the construction process inherently involves formative assessment processes by requiring a verification, quality check, and rethinking or improvement process (formative feedback) (Pérez-Pueyo et al., 2021). Although the validity of these results is grounded in the characteristics of the pedagogical proposal, it is worth questioning whether the pedagogical shift to shared grading is based on coherent agreements and has been carried out when students have gained awareness and responsibility for the process (Hortigüela-Alcalá et al., 2019; López-Pastor & Pérez-Pueyo, 2017).

Finally, in the specific case of Hybrid model, it mediates the relationship between FSA and grading with self-assessment and peer-assessment instruments. However, significant relationships are also evident between this pedagogical model and the use of traditional assessment and grading models. These results reflect the reality of the classroom where Hybrid model is a methodological catch-all referred to when different mixed methodological proposals are used, which are not always coherent (González-Víllora et al., 2019; Shen & Shao, 2022). As Pérez-Pueyo et al. (2021) note, the hybridization of models

should only be approached when the models to be combined are mastered and as long as it implies a better learning outcome.

Conclusions

The study's results reflect an unstable and unclear link between active methodologies and FSA in Primary PE. It is confirmed that proposals such as Service Learning or the Attitudinal Style are consistently integrating FSA processes. However, the research also highlights some inconsistencies that suggest the use of traditional assessment models within some active methodologies, resulting in a summative and final assessment focused on grading. Based on these results, and keeping in mind the need to link active methodologies with the FSA model to ensure a coherent teaching-learning process, it would be important to have access to successful experiences for teachers that clearly demonstrate the connection between the methodology and FSA processes, (with concrete examples).

Regarding the study's limitations, firstly, the sample used focused exclusively on Primary PE teachers, so the conclusions might not be generalizable to other educational levels or subjects. Secondly, the research was based on a quantitative analysis, which might have limited the understanding of the qualitative factors underlying teachers' decisions regarding formative assessment. Finally, the limited availability of empirical evidence in emerging models, such as Health-based PE, has reduced the depth of the analysis in some less established methodologies.

Finally, regarding future lines of research opened up by the study, it would be valuable to conduct qualitative studies that explore teachers' perceptions and experiences when implementing FSA in combination with pedagogical models, in order to identify practical obstacles and pedagogical attitudes that influence its application. In line with this, it would be worth analyzing how PE teachers utilize pedagogical models and how this links with FSA, based on their age and years of experience, as well as the causes and consequences this entails. In addition, it would be valuable to develop guides or teacher training programs aimed at strengthening the use of formative assessments in active methodologies, particularly in those that show less coherence, such as PBL, Sports Education, Teaching Games for Understanding, or the Hybrid model.

Ethics Committee Statement

The study was conducted in accordance with the Declaration of Helsinki and was approved by the Ethics Committee: UNIVERSIDAD AUTÓNOMA DE MADRID (CEI 131-2718, 16/05/2023).

Conflict of Interest Statement

The authors declare no conflicts of interest, including any personal circumstances or interests that could be perceived as influencing the presentation or interpretation of the results.

Funding

This research received no funding as it was conducted within the framework of a doctoral thesis.

Authors' Contribution

For multi-authored research articles, you must detail their individual contributions. For this purpose, use the CRedit taxonomy (https://credit.niso.org/): Conceptualization Arija-Mediavilla, A., Santos-Pastor, M.L. & Martínez-Muñoz, L.F.; Methodology Arija-Mediavilla, A., Santos-Pastor, M.L. & Ruiz-Montero, P.J; Software Ruiz-Montero, P.J; Validation Arija-Mediavilla, A., Santos-Pastor, M.L. & Martínez-Muñoz, L.F.; Formal Analysis Ruiz-Montero, P.J.; Investigation Arija-Mediavilla, A., Santos-Pastor, M.L. & Martínez-Muñoz, L.F.; Resources Arija-Mediavilla, A.; Data Curation Ruiz-Montero, P.J. & Arija-Mediavilla, A.; Writing – Original Draft Arija-Mediavilla, A.; Writing – Review & Editing Santos-Pastor, M.L. & Martínez-Muñoz, L.F.; Visualization Arija-Mediavilla, A.; Supervision Santos-Pastor, M.L.; Project Administration Santos-Pastor, M.L. All authors have read and agreed to the published version of the manuscript.

Data Availability Statement

The data that support the findings of this study are available on request from the corresponding author alba.arija@estudiante.uam.es

References

Abella-García, V., Ausín-Villaverde, V., Delgado-Benito, V., & Casado-Muñoz, R. (2020). Aprendizaje basado en proyectos y estrategias de evaluación formativas: Percepción de los estudiantes universitarios. *Revista Iberoamericana de Evaluación Educativa, 13*(1), 93-110. https://doi.org/10.15366/riee2020.13.1.004

Aarskog, E. (2021). 'No assessment, no learning': Exploring student participation in assessment in Norwegian physical education. *Sport, Education and Society, 26*(8), 875-888. https://doi.org/10.1080/13573322.2020.1791064

Alonso, G., Alonso, M., & Echarri, C. (2017). La evaluación en la cooperación como camino de mejora. En J. V. Ruiz-Omeñaca (Ed.), *Aprendizaje Cooperativo en Educación Física* (pp. 213-233). Editorial CCS.

Álvarez-Sánchez, J. L., Gutiérrez-García, C., & Hortigüela-Alcalá, D. (2025). Assessing the development of the attitudinal style as Pedagogical Model in Physical Education: A scoping review. *Retos. Nuevas Tendencias en Educación Física, Deportes y Recreación, 62*, 765–776. https://doi.org/10.47197/retos.v62.109744

Arija-Mediavilla, A., Santos-Pastor, M. L., Martínez-Muñoz, L. F., & Ruiz-Montero, P. J. (2024). Diseño y validación de un cuestionario mediante Método Delphi para valorar las relaciones entre metodologías activas y evaluación formativa en Educación Física en Primaria. *Retos. Nuevas Tendencias en Educación Física, Deportes y Recreación, 51*, 1442-1451. https://doi.org/10.47197/retos.v51.101502

Arufe-Giráldez, V., Sanmiguel-Rodríguez, A., Ramos-Álvarez, O., & Navarro-Patón, R. (2023). News of the Pedagogical Models in Physical Education. A Quick Review. *International Journal of Environmental Research and Public Health, 20*(3), 2586. https://doi.org/10.3390/ijerph20032586

Baena-Extremera, A. (2011). Programas didácticos para Educación Física a través de la Educación de Aventura. *Espiral. Cuadernos del Profesorado, 4*(7), 3-13. https://doi.org/10.25115/ecp.v4i7.914

Barba-Martín, R. A., Hortigüela-Alcalá, D., & Pérez-Pueyo, A. (2020). Evaluar en Educación Física: Análisis de la tensiones existentes y justificación del empleo de la evaluación formativa y compartida. *Educación Física y Deporte, 39*(1), 23-47. https://doi.org/10.17533/udea.efyd.v39n1a03

Barrientos-Hernán, E., López-Pastor, V. M., & Pérez-Brunicardi, D. (2019). ¿Por qué hago evaluación formativa y compartida y/o evaluación para el aprendizaje en EF? La influencia de la formación inicial y permanente del profesorado. *Retos. Nuevas Tendecias en Educación Física, Deportes y Recreación, 36*, 37-43. https://doi.org/10.47197/retos.v36i36.66478

Brookhart, S. M. (2011). *Grading and learning: Practices that support student achievement*. Solution Tree Press.

Calvo-Varela, D., Sotelino-Losada, A., & Rodríguez-Fernández, J. E. (2019). Aprendizaje-servicio e inclusión en educación primaria. Una visión desde la Educación Física. Revisión sistemática. *Retos. Nuevas Tendencias en Educación Física, Deporte y Recreación, 36*, 611–617. https://doi.org/10.47197/retos.v36i36.68972

Cañas-Encinas, M., García-Martín, N., Pinedo-González, R., & Caballero-San José, C. (2019). Aplicación del sistema de evaluación formativa y compartida a un proyecto gamificado. *Revista Infancia, Educación y Aprendizaje, 5*(2), 212-219. https://doi.org/10.22370/ieya.2019.5.2.1649

Casey, A., & McPhail, A. (2018). Adopting a models-based approach to teaching physical education. *Physical Education and Sport Pedagogy, 23*(3), 294-310. https://doi.org/10.1080/17408989.2018.1429588

Chiva-Bartoll, O., & Fernández-Rio, J. (2022). Advocating for service-learning as a pedagogical model in Physical Education: Towards an activist and transformative approach. *Physical Education and Sport Pedagogy, 27*(5), 545-558. https://doi.org/10.1080/17408989.2021.1911981

Coll, C., & Martín, E. (2021). La LOMLOE, una oportunidad para la modernización curricular. *Avances en Supervisión Educativa, 35*, 1-22. https://doi.org/10.23824/ase.v0i35.731

Díaz-Tejerina, D., & Fernández-Río, J. (2024). El modelo pedagógico de educación física relacionado con la salud. Una revisión sistemática siguiendo las directrices PRISMA*. Retos. Nuevas Tendencias en Educación Física, Deportes y Recreación, 51*, 129–135. https://doi.org/10.47197/retos.v51.101068

Fernández-Río, J., Calderón, A., Hortigüela-Alcalá, D., Pérez-Pueyo, A., & Aznar, M. (2016). Modelos pedagógicos en educación física: Consideraciones teórico-prácticas para docentes. *Revista Española de Educación Física y Deportes, 413*, 55-75. https://doi.org/10.55166/reefd.v0i413.425

Fernández-Río, J., Hortigüela-Alcalá, D., & Pérez-Pueyo, A. (2018). Revisando los modelos pedagógicos en Educación Física. Ideas clave para incorporarlos al aula. *Revista Española de Educación Física y Deportes, 423*, 57-80. https://doi.org/10.55166/reefd.vi423.695

Fernández-Río, J., & Iglesias, D. (2022). What do we know about pedagogical models in physical education so far? An umbrella review. *Physical Education and Sport Pedagogy, 29*(2), 190-205. https://doi.org/10.1080/17408989.2022.2039615

Galeano-Rojas, D., León-Reyes, B., Ortiz-Franco, M., Farías-Valenzuela, C., Ferrari, G., & Valdivia-Moral, P. (2023). Utilización del teaching games for understanding en deportes de equipo en el contexto de la educación física: Una revisión sistemática. *Journal of Sport and Health Research, 15*(1), 27-44. https://doi.org/10.58727/jshr.102695

González-Víllora, S., Evangelio, C., Sierra, J., & Fernández-Río, J. (2019). Hybridizing pedagogical models: A systematic review. *European Physical Education Review, 25*(4), 1056-1074. http://dx.doi.org/10.1177/1356336X18797363

Guijarro, E., Rocamora, I., Evangelio, C., & González-Víllora, S. (2020). El modelo de Educación Deportiva en España: una revisión sistemática. *Retos. Nuevas Tendencias en Educación Física, Deportes y Recreación, 38*, 886–894. https://doi.org/10.47197/retos.v38i38.77249

Haerens, L., Kirk, D., Cardon, G., & De Bourdeaudhuij, I. (2011). Toward the development of a pedagogical model for health-based physical education. *Quest, 63*(3), 321-338. https://doi.org/10.1080/00336297.2011.10483684

Herrero-González, D., López-Pastor, V. M., & Manrique-Arribas, J. C. (2020). La Evaluación formativa y compartida en contextos de aprendizaje cooperativo en educación física en Primaria. *Cultura, Ciencia y Deporte, 15*(44), 213–222. https://doi.org/10.12800/ccd.v15i44.1463

Herrero-González, D., Manrique Arribas, J. C., & López-Pastor, V. M. (2021). Incidencia de la formación inicial y permanente del profesorado en la aplicación de la evaluación formativa y compartida en educación física. *Retos. Nuevas Tendencias en Educación Física, Deportes y Recreación, 41*, 533–543. https://doi.org/10.47197/retos.v0i41.86090

Hortigüela-Alcalá, D., Pérez-Pueyo, A. & González-Calvo, G. (2019). Pero... ¿A qué nos referimos realmente con la evaluación formativa y compartida?: Confusiones habituales y reflexiones prácticas. *Revista Iberoamericana de Evaluación Educativa, 12*(1), 13-27. https://doi.org/10.15366/riee2019.12.1.001

Hortigüela-Alcalá, D., Pérez-Pueyo, A., & Fernández-Río, J. (2020). Evaluación formativa y modelos pedagógicos: Modelo de responsabilidad personal y social y de autoconstrucción de materiales. *Revista Española de Educación Física y Deportes, 430,* 23-41. https://doi.org/10.55166/reefd.vi430.919

Kirk, D., & MacPhail, A. (2002). Teaching games for understanding and pedagogical knowledge: The case of the TGfU model in physical education. *European Physical Education Review, 8*(3), 279-291. https://doi.org/10.1177/1356336X020083003

León-Díaz, O., Arija-Mediavilla, A., Martínez-Muñoz, L. F., & Santos-Pastor, M. (2020). Las metodologías activas en Educación Física. Una aproximación al estado actual desde la percepción de los docentes en la Comunidad de Madrid. *Retos. Nuevas Tendencias en Educación Física, Deportes y Recreación, 38,* 587-594. https://doi.org/10.47197/retos.v38i38.77671

León-Díaz, O., Martínez-Muñoz, L. F., & Santos-Pastor, M. (2023). Metodologías activas en la Educación Física. Una mirada desde la realidad práctica. *Retos. Nuevas Tendencias en Educación Física, Deportes y Recreación, 48*, 647-656. https://doi.org/10.47197/retos.v48.96661

López-Gómez, E. (2018). El método Delphi en la investigación actual en educación: Una revisión teórica y metodológica. *Educación XX1, 21*(1), 17-40. https://doi.org/10.5944/educXX1.15536

López-Pastor, V. M. (2006). L*a Evaluación en Educación Física: Revisión de los modelos tradicionales y planteamiento de una alternativa: La evaluación formativa y compartida.* Miño y Dávila.

López-Pastor, V. M., & Pérez-Pueyo, A. (2017). *Evaluación formativa y compartida en Educación: Experiencias de éxito en todas las etapas educativas*. Universidad de León. http://buleria.unileon.es/handle/10612/5999

Ley Orgánica 1/1990, de 3 octubre de ordenación general del sistema educativo, Boletín Oficial del Estado, 238 (1990). https://www.boe.es/eli/es/lo/1990/10/03/1

Ley Orgánica 2/2006, de 3 de mayo de Educación, Boletín Oficial del Estado, 106 (2006). https://www.boe.es/eli/es/lo/2006/05/03/2/con

Ley Orgánica 3/2020, de 29 de diciembre, por la que se modifica la Ley Orgánica 2/2006, de 3 de mayo, de Educación, 340 (2020). https://www.boe.es/eli/es/lo/2020/12/29/3

MacKinnon, D. P., Cheong, J., & Pirlott, A. G. (2012). Statistical mediation analysis. In H. Cooper, P. M. Camic, D. L. Long, A. T. Panter, D. Rindskopf, & K. J. Sher (Eds.), *APA handbook of research methods in psychology, Vol. 2. Research designs: Quantitative, qualitative, neuropsychological, and biological* (pp. 313–331). American Psychological Association. https://doi.org/10.1037/13620-018

MacPhail, A., Halbert, J., & O'Neill, H. (2018). The development of assessment policy in Ireland: A story of junior cycle reform. *Assessment in Education: Principles, Policy & Practice, 25*(3), 310-326. https://doi.org/10.1080/0969594X.2018.1441125

Moura, A., Graça, A., McPhail, A., & Batista, P. (2021). Aligning the principles of assessment for learning to learning in physical education: A review of literature. *Physical Education and Sport Pedagogy, 26*(4), 388-401. https://doi.org/10.1080/17408989.2020.1834528

Oliver-Álvarez, M., & Martos-García, D. (2023). La evaluación tradicional y sus consecuencias. Un caso en la Educación Física hegemónica. *Cultura, Ciencia y Deporte, 18*(58), 27-49. https://doi.org/10.12800/ccd.v18i58.1976

Otero-Saborido, F. M., González-Calvo, G., Hortigüela-Alcalá, D., & Vázquez- Ramos, F. J. (2023). Formative and shared assessment in primary school PE curriculum: Teachers' perceptions. *Cultura, Ciencia y Deporte, 18*(55), 79-89. https://doi.org/10.12800/ccd.v18i55.1945

Pérez-Pueyo, A. (2016). El estilo actitudinal en educación física: Evolución en los últimos 20 años. *Retos. Nuevas Tendencias en Educación Física, Deportes y Recreación, 29*, 207-215. https://doi.org/10.47197/retos.v0i29.38720

Pérez-Pueyo, A., & Hortigüela-Alcalá, D. (2020). ¿Y si toda la innovación no es positiva en Educación Física? Reflexiones y consideraciones prácticas. *Retos, 37*(37), 579-587. https://doi.org/10.47197/retos.v37i37.74176

Pérez-Pueyo, A., Hortigüela-Alcalá, D., & Fernández-Río, J. (2020). Evaluación formativa y modelos pedagógicos: Estilo actitudinal, aprendizaje cooperativo, modelo comprensivo y educación deportiva. *Revista Española de Educación Física y Deportes, 428*, 47-66. https://doi.org/10.55166/reefd.vi428.881

Pérez-Pueyo, A., Hortigüela-Alcalá, D., & Fernández-Río, J. (2021). *Los Modelos Pedagógicos en Educación Física. Qué, cómo, por qué y para qué*. Universidad de León.

Pérez-Pueyo, A., Hortigüela-Alcalá, D., Casado-Berrocal, O. M., Heras-Bernardino, C., & Herrán-Álvarez, I. (2022). Análisis y reflexión sobre el nuevo currículo de Educación Física. *Revista Española de Educación Física y Deportes, 436*(3), 41-58. https://doi.org/10.55166/reefd.vi463(3).1073

Pérez-Pueyo, A., Hortigüela-Alcalá, D., Gutiérrez-García, C., & Barba-Martín, R. A. (2024). *Evaluación Formativa y Compartida en Educación Física: Fundamentos y experiencias prácticas en todas las etapas educativas.* Universidad de León. https://hdl.handle.net/10612/22962

Preacher, K. J. & Hayes, A. F. (2008). Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models. *Behavior Research Methods, 40*(3), 879-891. https://doi.org/10.3758/brm.40.3.879

Santos-Pastor, M. L, Martínez-Muñoz, L. F., & Cañadas, L. (2019). La evaluación formativa en el aprendizaje-servicio. Una experiencia en actividades físicas en el medio natural. *Revista de Innovación y Buenas Prácticas Docentes, 8*(1), 110-118. https://doi.org/10.21071/ripadoc.v8i1.12000

Santos-Pastor, M. L., Martínez-Muñoz, L. F., Cuenca-Soto, N., Chiva-Bartoll, O., & Ruiz-Montero, P. J. (2024). Evaluación formativa en el modelo pedagógico de aprendizaje-servicio universitario. En A. Pérez-Pueyo, D. Hortigüela-Alcalá, C. Gutiérrez-García, & R. A. Barba-Martín (Coords.), *Evaluación formativa y compartida en educación física: Fundamentos y experiencias prácticas en todas las etapas educativas* (pp. 867-898). Universidad de León.

Shen, Y., & Shao, W. (2022). Influence of hybrid pedagogical models on learning outcomes in physical education: A systematic literature review. *International Journal of Environmental Research and Public Health, 19*(15), 1-16. https://doi.org/10.3390/ijerph19159673

UNESCO (2021). *Reimagining our futures together. A new social contract for education. Report from the International Commission on the Futures of Education.* https://unesdoc.unesco.org/ark:/48223/pf0000379381

Willis, J., Arnold, J., & DeLuca C. (2023). Accessibility in assessment for learning: Sharing criteria for success*. Frontiers in Education. 8*, 1-15. https://doi.org/10.3389/feduc.2023.1170454