FORMATIVE AND SHARED ASSESSMENT IN PHYSICAL EDUCATION IN EARLY CHILDHOOD AND PRIMARY EDUCATION: A SYSTEMATIC REVIEW

LA EVALUACIÓN FORMATIVA Y COMPARTIDA EN EDUCACIÓN FÍSICA EN LAS ETAPAS DE INFANTIL Y PRIMARIA: REVISIÓN SISTEMÁTICA

***Raúl Martínez-Benito*[1], *Daniel Caballero-Juliá*[1], *Antonio Sánchez-Martín*[1], *Galo Sánchez-Sánchez*[1]**

[1] Department of Didactics of Musical, Artistic, and Body Expression, Faculty of Education Sciences (Zamora) / University of Salamanca, España

Autor para la correspondencia:
Raúl Martínez Benito, raul.martinez@usal.es

Título Abreviado:
Formative and Shared Assessment in School Children: Systematic review

How to cite this article:
Martínez-Benito, R., Caballero-Juliá, D., Sánchez-Martín, A. & Sánchez-Sánchez, G. (2025). Formative and shared assessment in physical education in early childhood and primary education: a systematic review. *Cultura, Ciencia y Deporte, 20*(66), 2357. https://doi.org/10.12800/ccd.v20i66.2357

Received: 18 November 2024 / **Accepted:** 30 April 2025

Abstract

This systematic review examines research conducted over the past decade on Formative and Shared Assessment (FS&A) in Physical Education (PE) at early childhood and primary education levels. The main objectives are to analyze the primary research trends on FS&A, assess its impact on students' cognitive, motor, and socio-emotional development, and evaluate its influence on teaching and learning processes. The review followed the PRISMA methodology, initially identifying 245 articles from databases such as Scopus, Web of Science, and EBSCOHost. Selected articles were evaluated for methodological quality, resulting in the inclusion of 12 articles in the review. The findings indicate that most research has been conducted in Spain, primarily focusing on teachers' perceptions, highlighting the need for deeper exploration of F&SA's direct impact on students. The lack of studies on early primary grades suggests a promising future line of research. F&SA demonstrates multiple benefits in PE, such as fostering student autonomy, motivation, improving classroom climate, and enhancing teaching practice. However, limitations such as geographical concentration, small sample sizes, and methodological diversity underscore the need for future studies with larger and more diverse samples to deepen the understanding and optimization of F&SA implementation in PE. In conclusion, this systematic review provides an understanding of the current state of research on F&SA in PE, offering guidance for future approaches and research directions on this topic.

Keywords: Assessment for learning, educational assessment, formative assessment, physical education, shared assessment.

Resumen

Esta revisión sistemática examina la investigación realizada en la última década sobre la Evaluación Formativa y Compartida (EFyC) en Educación Física (EF) en las etapas de educación infantil y primaria. Los objetivos principales son analizar las principales tendencias en la investigación sobre EFyC, evaluar su impacto en el desarrollo cognitivo, motriz y socioemocional de los estudiantes, y valorar su influencia en los procesos de enseñanza-aprendizaje. La revisión se realizó bajo la metodología PRISMA, identificando inicialmente 245 artículos en bases de datos como Scopus, Web of Science y EBSCOHost. Los artículos seleccionados fueron evaluados según criterios de calidad metodológica, incluyéndose finalmente 12 artículos en la revisión. Los resultados indican que gran parte de la investigación se ha realizado en España, con un enfoque en las percepciones docentes, lo que resalta la necesidad de explorar más a fondo el impacto directo de la EFyC en el alumnado. La escasez de estudios en los primeros cursos de primaria sugiere una línea de investigación futura relevante. La EFyC demuestra múltiples beneficios en EF, como el fomento de la autonomía, la motivación del alumnado, la mejora del clima de aula y de la práctica docente. No obstante, limitaciones como la concentración geográfica, el tamaño reducido de las muestras y la diversidad metodológica subrayan la necesidad de estudios futuros con muestras más amplias y variadas para profundizar en la comprensión y optimización de la implementación de la EFyC en EF. En definitiva, esta revisión sistemática permite comprender el estado actual de la investigación sobre la EFyC en EF, aportando orientaciones para futuros enfoques y líneas de estudio en esta temática.

Palabras clave: Evaluación formativa, evaluación compartida, evaluación para el aprendizaje, evaluación educativa, educación física.

Introduction

Assessment plays a crucial role in the teaching–learning process and, consequently, in the achievement of learning outcomes (Fraile et al., 2020). Over the past decades, the concept of assessment has evolved from being merely an informative tool to determine the degree of learning acquisition to being understood as an integral strategy within the teaching–learning

process (Álvarez Méndez, 2011). This learning-centered perspective of assessment is known as formative assessment (López-Pastor, 1999), in which teaching and assessment are developed in tandem, allowing assessment activities to also serve a learning function (Chávez Ruiz & Martínez Rizo, 2018).

Formative and Shared Assessment (FS&A) is based on continuous feedback that students receive throughout their learning process, in which they also take an active role (López-Pastor & Sicilia-Camacho, 2017). Moreover, for these actions to be genuinely formative, they must be free from grading (Hortigüela et al., 2019), marking a clear departure from traditional evaluation methods in which students are accountable to the teacher only at the end of the process (López-Pastor, 2017). Similarly, both teachers and students must have a clear understanding of the intended learning objectives and the assessment criteria that will guide their evaluation (Carless, 2015; Fraile et al., 2020). In line with Black and Wiliam (2009), the implementation of FS&A processes can be conceptualized through a set of key strategies that provide a framework for teachers' practice (Table 1).

Table 1

Key Strategies for the Implementation of FS&A

	Key Strategies
1	Inform students about the learning objectives.
2	Provide students with the assessment references (evaluation/grading criteria).
3	Offer students continuous feedback throughout the learning process, free from grading.
4	Encourage student participation in the assessment processes.

Source: Own elaboration based on Black and Wiliam (2009).

Another feature that distinguishes FS&A models concerns the techniques and instruments employed. In more traditional assessment models, evaluation is typically carried out through knowledge tests on theoretical content, with examinations being the predominant assessment activity. In the field of Physical Education (PE), many teachers still hold a narrow view of the subject, considering it to be exclusively motor-based. Consequently, they continue to use evaluative practices focused on summative grades derived from physical fitness tests or motor skill assessments. In contrast, alternative assessment models promote a greater diversity of techniques and instruments, driven by the active participation of students in evaluative processes (López-Pastor, 2017).

FS&A focuses on providing both teachers and students with information about learning progress, coupled with feedback that allows for necessary adjustments and improvements (Black & Wiliam, 2009). This approach also seeks to enhance teaching processes by enabling teachers to identify educational needs and guide instructional decision-making (Hortigüela et al., 2019).

In primary education, PE aims to foster students' holistic development by promoting competencies across cognitive, motor, and socio-emotional domains, rather than focusing solely on the acquisition of physical skills (Otero-Saborido et al., 2023). FS&A contributes positively to cognitive development by enhancing understanding of the learning process; to the motor domain by reinforcing physical skills through reflective practice; and to the socio-emotional domain by fostering a supportive and collaborative peer environment (Wiliam & Thompson, 2007).

Initial teacher training is crucial to ensuring the effective implementation of FS&A (López-Pastor et al., 2020). However, such training is sometimes insufficient to provide teachers with a deep theoretical and practical understanding of this evaluative approach (Barrientos Hernán et al., 2023). This situation, combined with the challenge of breaking away from entrenched traditional practices in many schools, slows the expansion of FS&A in classrooms (López-Pastor, 2017). Many

teachers still associate assessment with grading (Hortigüela et al., 2019), using both terms interchangeably, or focus their evaluative activities solely on measuring learning outcomes (Martínez-Benito, 2022). This represents a limitation that may hinder the full realization of the benefits that FS&A can bring to classrooms and, consequently, to students' holistic development.

According to several authors, true methodological change begins with a transformation in how assessment is understood and applied in the classroom (Pérez-Pueyo et al., 2020). Despite the challenges mentioned, the last two decades have witnessed significant growth in publications related to FS&A. For instance, the work of the *Red de Evaluación Formativa y Compartida en Educación* (REFYCE), created in 2005 in Spain, along with the 16 editions of the *International Congress on Formative and Shared Assessment* held annually since 2006, demonstrates a growing interest in educational change in the field of assessment.

Objective

Based on the foregoing, the objective of this study is to systematically review the research conducted on the implementation of Formative and Shared Assessment (FS&A) processes in Early Childhood and Primary Education (ages 3 to 12) within the field of Physical Education (PE). The aim is to provide an overview of the current state of research on FS&A in PE, offering guidance for future approaches and lines of inquiry in this area.

More specifically, the study seeks to achieve the following objectives:

1. To examine the research trends on FS&A in Early Childhood and Primary Education in PE over the past decade, analyzing the characteristics and studies that have been carried out.
2. To identify the effects of implementing FS&A systems on students in Early Childhood and Primary Education in PE, particularly regarding the different dimensions of human development (cognitive, motor, and socio-emotional).
3. To identify the effects of applying FS&A systems in PE on the improvement of the teaching–learning process and teaching practice.

Research Questions

- What are the current research trends on FS&A in PE at the Early Childhood and Primary Education stages?
- In what ways does FS&A impact the holistic development of students at these educational stages?
- How does FS&A influence the improvement of the teaching–learning process and teaching practice?

Materials and Methods

Search Strategy

The article selection process for this study was carried out following a methodology based on the international PRISMA statement (Moher et al., 2009). Initially, searches were conducted in the following databases: SCOPUS, Web of Science, and EBSCOhost.

To perform the search, several keywords were used in both Spanish and English. In Spanish, the terms *"Evaluación Formativa"*, *"Evaluación Compartida"*, and *"Educación Física"* were employed, while in English the keywords *"Assessment for Learning"*, *"Formative Assessment"*, *"Shared Assessment"*, and *"Physical Education"* were used. These keywords were combined using the Boolean operators OR and AND to obtain more specific results:

- "*Evaluación Formativa OR Evaluación Compartida AND Educación Física*"
- "*Assessment for Learning OR Formative Assessment OR Shared Assessment AND Physical Education*"

The last search was conducted on October 17, 2024, and a time filter was applied to include studies published between 2014 and 2024. The following table presents the search strategies applied across the different databases:

Table 2

Search Strategy

Database	Search Strategy
Scopus	**Article title, Abstract, Keywords** ("Formative assessment" OR "Shared assessment" OR "Assessment for learning" OR "Evaluación formativa" OR "Evaluación compartida" AND "Physical education" OR "Educación física") AND (FILTERS–YEAR–RANGE–FROM–2014–TO–2024).
Web of Science	**Title** ("Evaluación formativa" OR "Evaluación compartida" OR "Formative assessment" OR "Assessment for learning" OR "Shared assessment" AND "Educación física" OR "Physical education") OR **Abstract** ("Evaluación formativa" OR "Evaluación compartida" OR "Formative assessment" OR "Assessment for learning" OR "Shared assessment" AND "Educación física" OR "Physical education") AND (**QUICK FILTERS – PUBLICATION YEARS** – 2024, 2023, 2022, 2021, 2020, 2019, 2018, 2017, 2016, 2015, 2014)
EBSCOHost	**TI** ("Formative assessment" OR "Shared assessment" OR "Assessment for learning" OR "Evaluación formativa" OR "Evaluación compartida" AND "Physical education" OR "Educación física") OR **AB** ("Formative assessment" OR "Shared assessment" OR "Assessment for learning" OR "Evaluación formativa" OR "Evaluación compartida" AND "Physical education" OR "Educación física") AND (**LIMIT-TO PUBLICATION-DATE** FROM-2014-TO-2024)

Inclusion and Exclusion Criteria

Only those studies that met the following criteria were selected: a) scientific articles based on empirical research; b) documents written in English or Spanish; c) studies conducted at the Early Childhood or Primary Education stages; and d) studies published between 2014 and 2024, in order to include the most recent literature from the last decade.

On the other hand, the following were excluded: a) scientific articles focused on the validation of research instruments; b) studies conducted at the Secondary or Higher Education levels, or those in which the age of the students was not specified; c) studies that did not specify the FS&A strategies applied.

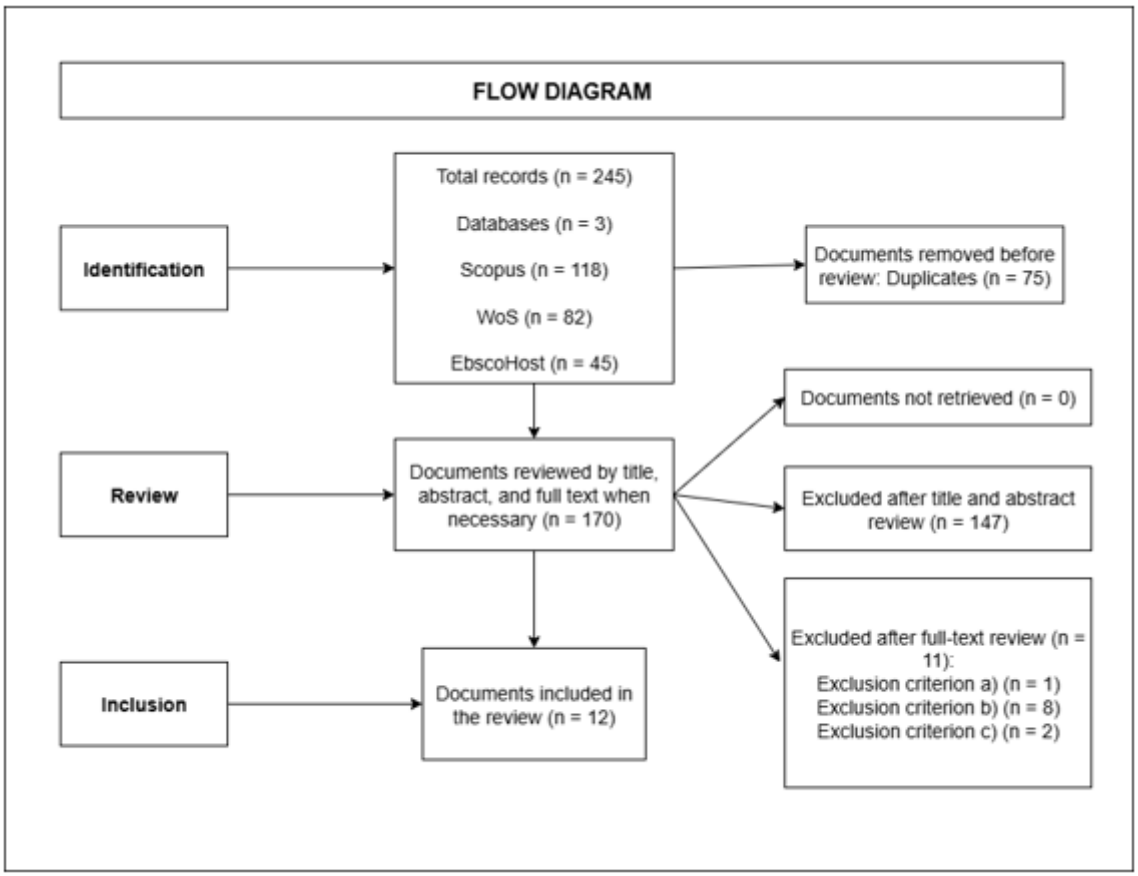
Study Selection Process

The systematic search process (Figure 1) identified a total of 245 studies, of which 75 were removed due to duplication. In the first screening phase, titles and abstracts were examined to exclude those articles that did not match the objective of the study. A total of 147 articles were excluded at this stage.

In a second phase, the 23 selected articles were reviewed in full text to ensure their relevance and alignment with the study objectives. Eleven studies were excluded because they focused on higher or secondary education, did not specify the educational stage, did not explain how FS&A was applied in the study, or were conducted for the design and validation of research instruments. In total, 12 articles were included in the study.

Figure 1

PRISMA Flow Diagram



Methodological Quality Assessment

For the assessment of the quality of the included articles, an independent evaluation was carried out by two reviewers, and any discrepancies were discussed until a consensus was reached.

For the evaluation of quantitative studies (Table 3), the Methodological Checklist for Studies Using Questionnaires (Boynton & Greenhalgh, 2004) was employed. Each item in the checklist was coded as "yes" (1) if it was explicitly presented

and described in the study, "no" (0) if it was absent, or "unclear" (0) if it was not clearly defined. A value of 1 was assigned to "yes" and a value of 0 to both "no" and "unclear," resulting in a total score ranging from 0 to 9. Study quality was assessed according to the following scale: 0–3 = low quality; 4–6 = moderate quality; and 7–9 = high quality.

Table 3

Methodological Quality Assessment of Quantitative Studies

	Items									
Studies	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**TS**
Chan et al. (2016)	1	0	1	1	1	1	1	1	1	8
Chng & Lund (2018)	1	0	1	1	1	1	1	1	1	8
Molina Soria et al. (2020)	0	1	1	0	1	1	1	1	1	7
Carrillo López et al. (2022)	1	0	1	1	1	0	1	1	1	7
Carrillo López & Hortigüela-Alcalá (2022)	1	0	1	1	1	0	1	1	1	7
Carrillo López (2023)	1	0	1	1	1	0	1	1	1	7

1 = Use of a validated instrument; 2 = Example of question provided; 3 = Appropriate sample recruited; 4 = Adequate demographic data reported; 5 = Clear distribution and administration of the questionnaire; 6 = Adequate response rate; 7 = Appropriate statistical analysis; 8 = Data relevant to addressing the research question; 9 = Clear connection between data and conclusion; TS = Total Score.

The quality of qualitative studies (Table 4) was assessed using a checklist developed in accordance with the methodological guidelines for systematic reviews established by the Joanna Briggs Institute (JBI) (Lockwood et al., 2015). The item coding followed the same procedure described in the previous paragraph.

Table 4

Methodological Quality Assessment of Qualitative Studies

	Items										
Studies	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**TS**
Barrientos Hernán et al. (2019)	1	1	1	1	1	0	0	1	1	1	8
López-Benavente et al. (2019)	1	1	1	1	1	0	0	1	1	1	8
Herrero-González et al. (2020)	1	1	1	1	1	0	0	1	1	1	8
Molina Soria et al. (2020)	1	1	1	1	1	0	0	1	1	1	8
Cañadas & Santos-Pastor (2021)	1	1	1	1	0	0	0	1	1	1	8
Barrientos Hernán et al. (2023)	1	1	1	1	1	0	0	1	1	1	8
Palacios-Gómez et al. (2023)	1	1	1	1	1	0	0	1	1	1	8

1 = Is there congruence between the stated philosophical perspective and the research methodology?; 2 = Is there congruence between the research methodology and the research question or objectives?; 3 = Is there congruence between the research methodology and the methods used to collect data?; 4 = Is there congruence between the research methodology and the representation and analysis of data?; 5 = Is there congruence between the research methodology and the interpretation of results?; 6 = Is there a statement locating the researcher culturally or theoretically?; 7 = Is the influence of the researcher on the research, and vice versa, addressed?; 8 = Are participants and their voices adequately represented?; 9 = Is the research ethical according to current criteria, or, for recent studies, is there evidence of ethical approval by an appropriate body?; 10 = Are the conclusions drawn from the research report derived from the analysis or interpretation of the data?; TS = Total Score.

To ensure that the systematic review included studies applying FS&A processes, the assessment systems described in each study were analyzed (Table 5). For this purpose, an ad hoc six-item checklist was designed based on the work of Black and Wiliam (2009). Each criterion was coded as (1) when met and (0) when not met. Studies that did not meet any of the established criteria (exclusion criterion c) were excluded from the review. The items included in the checklist were as follows:

1. **Continuous assessment:** Specifies how teachers and students obtain information throughout the teaching–learning process.
2. **Feedback:** Specifies that students receive feedback regarding their level of performance (cognitive, motor, and/or socio-emotional) or their learning progress.
3. **Ungraded assessment:** Specifies that assessment processes carried out throughout the teaching–learning process are free from grading.
4. **Learning objectives:** Specifies that students are aware of the learning objectives.
5. **Assessment criteria:** Specifies that students are aware of the assessment criteria or reference standards by which they will be evaluated.
6. **Shared assessment:** Specifies how students participate in the assessment processes that are carried out.

Table 5

Evaluation of the Assessment Process in the Studies

	Authors Criteria						
	1 Continuous	**2 Feedback**	**3 Grading**	**4 Learning Objectives**	**5 Criteria**	**6 Participation**	**TS**
Chan et al. (2016)	1	1	1	1	1	1	6
Chng & Lund (2018)	1	0	1	0	0	1	3
Barrientos Hernán et al. (2019)	1	1	1	1	1	1	6
López-Benavente et al. (2019)	1	0	1	0	0	1	3
Herrero-González et al. (2020)	1	1	0	0	0	1	3
Molina Soria et al. (2020)	1	1	1	1	1	1	6
Cañadas & Santos-Pastor (2021)	1	1	0	1	0	1	4
Carrillo López et al. (2022)	1	1	0	1	0	0	3
Carrillo López & Hortigüela-Alcalá (2022)	1	1	0	1	0	0	3
Barrientos Hernán et al. (2023)	1	1	0	1	1	1	5
Carrillo López (2023)	1	1	0	1	0	0	3
Palacios-Gómez et al. (2023)	0	1	1	0	0	1	3

1 = Does it specify that both the teacher and students receive information throughout the teaching–learning process?; 2 = Does it specify that students receive feedback regarding their level of performance (cognitive, motor, and/or socio-emotional) or their learning progress?; 3 = Does it specify that assessment processes carried out throughout the teaching–learning process are free from grading?; 4 = Does it specify that students are aware of the learning objectives?; 5 = Does it specify that students are aware of the assessment criteria or reference standards by which they will be evaluated?; 6 = Does it specify how students participate in the assessment processes that are carried out?; TS = Total Score.

Data Extraction

For data extraction, a summary table was used to compile the following information: author, year and country where the research was conducted, type of study, participants, study variable, instrument, main results, and limitations.

Results

The following table presents the studies included in the systematic review:

Table 6

Reviewed Articles

Nº	Authors and year	Title
1	Chan et al. (2016)	Improving fundamental movement skills in Hong Kong students through an assessment for learning intervention that emphasizes fun, mastery, and support: the A + FMS randomized controlled trial study protocol
2	Chng & Lund (2018)	The Impact of Formative Assessment on the Success and Response Rate in a 6th Grade Physical Education Badminton Unit
3	Barrientos Hernán et al. (2019)	Why do I do Formative and Share Assessment and/or Assessment For Learning in Physical Education? The influence of Teacher Education
4	López-Benavente et al. (2019)	Evaluación formativa y compartida para la inclusión de descansos activos en infantil
5	Herrero-González et al. (2020)	Formative and shared assessment in cooperative learning contexts in physical education in primary
6	Molina Soria et al. (2020)	El uso de sistemas de evaluación formativa y compartida en las aulas de educación fisica en educación primaria
7	Cañadas & Santos-Pastor (2021)	Formative Assessment in Primary and Secondary Physical Education Classes From Novel Teachers' Perspective
8	Carrillo López et al. (2022)	Formative assessment in Physical Education and its relation to the level of attention of primary school children
9	Carrillo López & Hortigüela-Alcalá (2022)	Impacto de la evaluación formativa en el estado de peso y calidad de la dieta en escolares de primaria
10	Barrientos Hernán et al. (2023)	Challenges with using formative and authentic assessment in physical education teaching from experienced teachers' perspectives
11	Carrillo López (2023)	Formative Assessment of Physical Education Teachers and Self-Concept Level of Primary School Children
12	Palacios-Gómez et al. (2023)	Miedos y percepciones de la aplicación de la evaluación formativa y compartida por un docente novel de Educación Física

The following table summarizes the most significant data from the reviewed studies.

Table 7

Results

N	Author, Year, and Country	Type of Study	Participants	Study Variable	Instrument	Main Results	Limitations
1	Chan et al. (2016) Hong Kong.	Experimental study with control and experimental groups. Pretest and posttest design (quantitative).	3rd-grade students (n = 282).	Competence in motor skills. Perception of physical competence level, motor skills, enjoyment, and teacher support.	TGMD-3 (Test of Gross Motor Development, 3ra edición). The Athletic Competence subscale of the Self-Perception Profile for Children (Harter 1985b). The pictorial scale of perceived movement skill competence for young children (Barnett et al. 2015). PE Enjoyment Rating Scale (Prochaska et al. 2003). Harter's Social Support Scale for Children (Harter 1985a). Self-designed questionnaire.	Formative and shared assessment processes were effective in improving motor skills. Students in the experimental group achieved significantly greater improvements than those in the control group. Students in the experimental group also obtained higher results regarding their perception of physical competence, enjoyment, and teacher support.	The limitations include the duration of the interventions concerning the implementation of formative and shared assessment strategies (550 minutes). A geographical limitation is also noted, as the study was conducted with students from a single academic level (3rd grade).
2	Chng & Lund (2018) United States.	Experimental study (quantitative).	6th-grade students (n = 16), 5 males and 9 females.	Impact of formative assessment on performance (number of successful hits) and practice participation time in badminton. Knowledge of badminton rules. .	Observation guide. Knowledge test.	The experimental group achieved a higher success rate than the control group, although the difference was not significant. Practice time was greater in the experimental group, but the difference was not significant. The experimental group improved its results compared to the control group in the knowledge test.	Students did not receive feedback on motor performance. The study experienced unforeseen interruptions in classroom management, which affected practice time in one of the groups (experimental). The type of game (singles or doubles) was not considered when evaluating the response rate.

Table 7 (cont.)

Results

N	Author, Year, and Country	Type of Study	Participants	Study Variable	Instrument	Main Results	Limitations
3	Barrientos Hernán et al. (2019) Spain.	Multiple case study (qualitative).	PE teachers (n = 4), 3 from primary education and 1 from secondary education. Mean age: 49.5 years.	Implementation of FS&A in PE. Influence of initial or ongoing teacher training on the implementation of FS&A processes.	Interview.	The teachers participating in the study implemented formative and shared assessment processes, highlighting several advantages: improvement of the classroom climate, increased student awareness and participation regarding learning objectives, and regulation of the educational process through self-assessment and peer assessment. Initial teacher training of the participants did not influence the implementation of formative and shared assessment processes. Ongoing professional development, driven by teachers' educational beliefs and convictions, had a strong influence on their progression toward implementing formative and shared assessment processes.	NS

Table 7 (cont.)

Results

N	Author, Year, and Country	Type of Study	Participants	Study Variable	Instrument	Main Results	Limitations
4	López-Benavente et al. (2019) Spain.	Case study (qualitative).	Preschool students (n = 46), 22 boys and 24 girls, aged 5–6 years, from a public school in Murcia, Spain.	Effectiveness of FS&A in PE.	Diary for assessing motor wedges (Cebrián et al., 2013). Individual evaluation target for motor wedges (modified from González-González et al., 2014).	It is feasible to implement formative and shared assessment systems from early ages. Formative assessment processes allow teachers to obtain relevant information to improve their teaching practice.	NS
5	Herrero-González et al. (2020) Spain.	Multiple case study (qualitative).	Primary PE teachers (n = 3).	Implementation of FS&A in Cooperative Learning in PE. Responsibility assumed by students in the assessment process.	Non-participant direct observation, semi-structured interview, and document analysis.	Teachers employ FS&A, and students assume responsibility in the assessment process. It enhances the teaching–learning process, provides continuous feedback, and promotes students' autonomy and self-regulation.	The research was conducted over a limited period of seven weeks, which may have reduced the depth of the data obtained. Qualitative study, not generalizable.
6	Molina Soria et al. (2020) Spain.	Mixed-methods study.	Primary PE teachers (n = 17) from the province of Valladolid, Spain. Of the 17 teachers, only 3 participated in the qualitative phase of the study (interviews).	Implementation of FS&A in PE. Advantages of applying FS&A in PE.	Self-designed questionnaire. Semi-structured interview. .	FS&A processes promote both student learning and teaching practice. Students develop greater awareness of the learning process. Students' motivation and interest in the teaching process increase. It fosters teacher reflection, allowing for the improvement of classroom practice. The main drawback identified is the workload it entails.	NS

Table 7 (cont.)

Results

N	Author, Year, and Country	Type of Study	Participants	Study Variable	Instrument	Main Results	Limitations
7	Cañadas & Santos-Pastor (2021) Spain.	Case study (qualitative).	Novice PE teachers (n = 4) with less than 3 years of experience. Two teach at the primary level and two at the secondary level.	Implementation of FS&A in PE.	Semi-structured interviews, non-participant observations, and document analysis.	The study shows that teachers do not apply FS&A fully or consistently throughout the teaching process. Only some elements are integrated into assessment practices, attempting to give them a different approach, but without achieving a systematic use of formative assessment.	Small number of participants and short observation period. Results are not generalizable to all educational contexts. .
8	Carrillo López et al. (2022) Spain.	Descriptive and cross-sectional study (quantitative).	Primary school students (n = 172), 93 boys and 79 girls, aged 10–13 years (M = 11.40, SD = 1.68), in the Autonomous Community of the Canary Islands, Spain.	Relationship between students' attention level and their perception of formative assessment practices in PE.	Teacher Performance Questionnaire on Formative Assessment Practices (Cerón et al., 2020). Perception of Similarities and Differences Test to measure attention (Caras-R) (Thurstone & Yela, 2019).	Students with lower attention levels perceived a greater use of formative assessment practices by their PE teachers.	Limited sample size and cross-sectional design that does not allow causal inferences.
9	Carrillo López & Hortigüela-Alcalá (2022) Spain.	Descriptive cross-sectional study (quantitative).	Primary school students (n = 122), 65 boys and 57 girls, aged 10–12 years (M = 10.84, SD = 1.20), in the Autonomous Community of the Canary Islands, Spain.	Relationship between weight and diet quality and the perception of the implementation of formative assessment in PE.	Teacher Performance Questionnaire on Formative Assessment Practices (Cerón et al., 2020). KIDMED Questionnaire (Serra-Majem et al., 2004) for diet quality. Body Mass Index (BMI) assessment.	No significant differences were found in perceived assessment practices in relation to weight status and diet quality, except for proactive formative assessment.	The study is cross-sectional, with a limited sample size and based on student self-reports, which may introduce bias. The results are not generalizable, as only students' perceptions were considered, and the study lacked a control group.

Table 7 (cont.)

Results

N	Author, Year, and Country	Type of Study	Participants	Study Variable	Instrument	Main Results	Limitations
10	Barrientos Hernán et al. (2023) Spain.	Qualitative study.	PE teachers (n = 4), 3 in primary education (2 males and 1 female) and 1 in secondary education (1 female).	Challenges in implementing FS&A in PE.	Semi-structured interview.	When properly implemented, FS&A supports student learning. Teachers may encounter difficulties in collecting the information obtained through observation processes. The importance of planning for the effectiveness of teaching processes and of teacher participation in communities or working groups on the topic was highlighted.	Small sample of teachers who already have experience and competence in FS&A, which limits generalization.
11	Carrillo López (2023) Spain.	Descriptive and cross-sectional study (quantitative).	Primary school students (n = 122), 65 boys and 57 girls, aged 10–12 years (M = 10.84, SD = 1.20), in the Autonomous Community of the Canary Islands, Spain.	Relationship between students' self-concept and their perception of formative assessment practices in PE.	Teacher Performance Questionnaire on Formative Assessment Practices (Cerón et al., 2020). Spanish version of the Piers-Harris Self-Concept Scale (Cardenal & Fierro, 2003).	Teachers' formative assessment practices are associated with higher self-concept in several dimensions (behavioral, intellectual, and social).	Limited sample size and cross-sectional design; does not allow causal inferences, and the sample was selected by convenience.

Table 7 (cont.)

Results

N	Author, Year, and Country	Type of Study	Participants	Study Variable	Instrument	Main Results	Limitations
12	Palacios-Gómez et al. (2023) Spain.	Qualitative study.	Academic year 2019–2020: Primary school students (n = 126) from 4th, 5th, and 6th grades; 60 boys and 66 girls. Academic year 2020–2021: Primary school students (n = 126) from 4th, 5th, and 6th grades; 61 boys and 65 girls. PE teacher (n = 1) with less than 4 years of experience.	Fears and insecurities shown by the PE teacher in the implementation of FS&A. Student participation and involvement in FS&A processes.	Teacher follow-up diary (teacher). Focus group (students).	Initially, the implementation of FS&A generated certain uncertainties among students: concern about grading, doubts about how to complete the assessment instruments, and insecurity when participating in peerassessment processes. Peerassessment processes presented some difficulties: lack of acceptance of feedback received from classmates, discrepancies between the self-perception of the student being assessed and the peer assessor, and conflicts of interest. The teacher's main insecurities were related to the difficulties students experienced in providing effective feedback to their peers and in understanding an assessment process free from grading. Positive effects were observed in students' participation and involvement in the assessment process. It enabled students to understand the progression of their learning and fostered greater awareness of their own learning.	NS

Note. NS = Not Specified.

Discussion

The discussion has been organized based on the research objectives.

Objective 1. To Examine the Line of Research Conducted on FS&A in Physical Education (PE) at the Early Childhood and Primary Education Stages Over the Last Decade, Analyzing the Characteristics and Studies That Have Been Carried out

Most of the reviewed studies were conducted in Spain (*n* = 10). The remaining ones were carried out in the United States (Chng & Lund, 2018) and Hong Kong (Chan et al., 2016). Regarding the type of study, there is a balance among different approaches: qualitative (*n* = 6), quantitative (*n* = 5), and mixed-methods (*n* = 1).

The reviewed studies present a dual perspective in terms of participants. On the one hand, there are studies focused on teachers (*n* = 6), and on the other, those directed toward students (*n* = 6). Only one study offers a more comprehensive view by including both groups (Palacios-Gómez et al., 2023). Concerning teachers, most studies have small samples ranging from 1 to 4 participants (*n* = 5). One study stands out for including a larger sample of 17 participants, providing a broader analysis (Molina Soria et al., 2020). Among these studies, two focus on novice teachers—defined as those with less than three years of experience (Cañadas & Santos-Pastor, 2021) or fewer than four (Palacios-Gómez et al., 2023).

Regarding student-focused studies, only one includes participants from early childhood education (López-Benavente et al., 2019). Studies at the primary level show variability in sample size, ranging from 16 participants (Chng & Lund, 2018) to 282 (Chan et al., 2016). It is worth noting that no studies were found on the early years of primary education, with third grade being the earliest level studied (Chan et al., 2016). Most research focuses on the last three years of primary education (*n* = 4).

With respect to study variables, among teacher-focused research, most analyze the implementation of FS&A in PE (*n* = 4). Of these, Herrero-González et al. (2020) adopt a more specific focus by examining FS&A within cooperative learning. Barrientos Hernán et al. (2019) explore the influence of initial and ongoing teacher training on the implementation of FS&A. Carrillo López et al. (2022) investigate the advantages of applying this type of assessment, while Cañadas and Santos-Pastor (2021) analyze its application among novice teachers. Other studies focus on the challenges faced by teachers when implementing FS&A (Barrientos Hernán et al., 2023) or on the fears and insecurities experienced by novice teachers during its application (Palacios-Gómez et al., 2023). The latter also analyzes student participation and involvement in FS&A processes.

In terms of student-centered research, the studies address various variables, with only two focusing on the motor domain. The study by Chng and Lund (2018) analyzes motor performance—measured by the number of successful hits in badminton—along with practice participation time and theoretical knowledge of the sport's rules, depending on whether FS&A processes were applied. Similarly, Chan et al. (2016) assess the effectiveness of a PE program implementing FS&A to improve students' motor competence in skills such as jumping, throwing, and catching. Additionally, this study examines students' perceived physical competence, enjoyment, and perceived teacher support. Other studies (*n* = 3) explore the relationship between students' perceptions of FS&A in the learning process and other variables such as weight status and diet quality (Carrillo López & Hortigüela-Alcalá, 2022), self-concept (Carrillo López, 2023), and attention level (Carrillo López et al., 2022).

A notable diversity of instruments was observed. In qualitative studies, the interview emerges as the primary data collection method (*n* = 5). In addition to interviews, Herrero-González et al. (2020) and Cañadas and Santos-Pastor (2021) employed non-participant observation and document analysis. In López-Benavente et al. (2019), a teacher diary (Cebrián et al., 2013) was used to record practice-based observations, along with an individual evaluation target for motor wedges (González-González et al., 2014), which uses a visual scale to assess satisfaction with activities among early childhood students. Palacios-Gómez et al. (2023) used focus groups and a teacher follow-up diary, in which the teacher mainly documented personal reflections.

Among quantitative studies, the *Teacher Performance Questionnaire on Formative Assessment Practices* (Cerón et al., 2020) stands out. This instrument consists of 21 items grouped into six subscales, yielding a score that reflects the level of formative assessment practices of teachers. Carrillo López (2023) used it to explore its relationship with students' self-concept, measured through the Spanish version of the *Piers-Harris Self-Concept Scale* (Cardenal & Fierro, 2003). Carrillo López et al. (2022) used the same questionnaire alongside the *Perception of Similarities and Differences Test (Caras-R)* (Thurstone &

Yela, 2019) to examine its relationship with student attention. Finally, Carrillo López and Hortigüela-Alcalá (2022) applied it together with the *KIDMED Questionnaire* (Serra-Majem et al., 2004) to assess its relationship with diet quality.

Regarding studies focusing on motor aspects, Chng & Lund (2018) developed an observation guide to record successful and failed hits, as well as participation rate per minute during badminton games. Chan et al. (2016), on the other hand, applied the *TGMD-3* (*Test of Gross Motor Development*), which evaluates motor skill levels based on six tasks related to jumping and ball handling. Each participant receives an overall score based on specific and observable performance criteria. In the same study, several complementary instruments were used: *Harter's Social Support Scale for Children* (Harter, 1985a) to assess perceived teacher support, the *Athletic Competence Subscale of the Self-Perception Profile for Children* (Harter, 1985b) to measure perceived physical competence, the *Pictorial Scale of Perceived Movement Skill Competence for Young Children* (Barnett et al., 2015) to assess perceived motor skills, and the *PE Enjoyment Rating Scale* (Prochaska et al., 2003) to measure students' enjoyment in PE classes.

Objective 2. To Identify the Effects of Implementing FS&A Systems in Students From Early Childhood and Primary Education Stages in PE With Respect to the Different Dimensions of the Person (Cognitive, Motor, and Socio-Emotional)

The study by Chng and Lund (2018) showed that the group in which FS&A was implemented achieved better results on a knowledge test about badminton rules. Continuing with the cognitive dimension, the study by Carrillo López and Hortigüela-Alcalá (2022) found no significant differences in the improvement of students' diet quality in relation to the application of FS&A during a learning situation focused on nutritional habits.

At the motor level, Chng and Lund (2018) demonstrated that assessment processes promoted greater motor success in badminton practice, as well as increased practice time. Similarly, Chan et al. (2016) showed the effectiveness of FS&A in improving students' motor skills.

In the socio-emotional dimension, Carrillo López (2023) argued that formative assessment practices are associated with higher student self-concept in behavioral, intellectual, and social domains. Moreover, Chan et al. (2016) revealed a relationship between the implementation of FS&A and students' perceptions of their physical competence, enjoyment of PE classes, and perceived teacher support. In turn, Carrillo López et al. (2022) concluded that students with lower attention levels perceived a greater presence of formative assessment practices from their teachers.

Regarding the limitations of the results presented in this section, two main aspects stand out. First, several studies (n = 3) point to the small sample size as a limitation (Carrillo López et al., 2022; Carrillo López & Hortigüela-Alcalá, 2022; Carrillo López, 2023). These same studies also note the inability to establish causality due to their cross-sectional design. Second, Chan et al. (2016) highlight the limited duration of the intervention as an important constraint.

Objective 3. To Identify the Effects of Implementing FS&A Systems in PE on the Improvement of the Teaching–Learning Process and Teaching Practice

The results of the analyzed studies reveal several advantages in the implementation of FS&A. First, a significant improvement was observed in classroom climate, as well as in student participation and awareness regarding learning objectives (Barrientos Hernán et al., 2019; Palacios-Gómez et al., 2023). These processes not only enhance student learning but also foster teacher reflection and improvement of classroom practice (Molina Soria et al., 2020). Moreover, FS&A promotes student autonomy and self-regulation, as learners assume greater responsibility for their own assessment and develop a deeper understanding of their academic progress (Herrero-González et al., 2020; Palacios-Gómez et al., 2023). Another reported benefit is the increase in student motivation and interest, along with greater awareness of their learning process (Molina Soria et al., 2020). The studies also indicate that, although initial teacher education does not significantly influence the adoption of these methods, ongoing professional development—supported by educational beliefs and convictions—is essential for the effective implementation of these processes (Barrientos Hernán et al., 2019).

However, some difficulties and resistances in the implementation of FS&A are also highlighted. One of the main limitations is the additional workload for teachers, especially in managing and analyzing feedback derived from observation processes, which requires careful planning and, in some cases, support from professional learning communities (Molina Soria et al., 2020; Barrientos Hernán et al., 2023). It is also evident that less experienced teachers tend to integrate only certain elements of FS&A without achieving systematic implementation, mainly due to their limited experience with such

processes (Cañadas & Santos-Pastor, 2021). Likewise, FS&A can generate insecurity among both students and teachers. Students often express concern about the absence of grading in assessments, as well as uncertainty when completing evaluation instruments or participating in peer assessments, which may lead to rejection of feedback from classmates due to discrepancies between self-perception and peer judgment (Palacios-Gómez et al., 2023). Teachers, in turn, face the challenge of ensuring that students provide constructive feedback to their peers and understand the value of non-graded assessment (Barrientos Hernán et al., 2023).

The main limitations reported in these studies include the impossibility of generalizing results (Barrientos Hernán et al., 2023; Cañadas & Santos-Pastor, 2021; Chng & Lund, 2018), the short duration of the intervention period (Barrientos Hernán et al., 2023), and the small sample size (Barrientos Hernán et al., 2023; Cañadas & Santos-Pastor, 2021). It should be noted that the studies by Barrientos Hernán et al. (2023) and Barrientos Hernán et al. (2019) correspond to the same research project. Similarly, the works of Carrillo López et al. (2022) and Carrillo López (2023) derive from the same study. This duplication limits the overall perspective of the systematic review.

Conclusions

This systematic review has highlighted key areas and challenges in research on FS&A in Physical Education (PE) at the Early Childhood and Primary Education stages. Although multiple benefits associated with FS&A have been documented—such as the promotion of student autonomy, increased motivation, and improvements in classroom climate and the teaching–learning process—many of the reviewed studies focus primarily on teachers' perceptions, underscoring the need to explore its direct impact on students more deeply.

Research in this area has been conducted mainly within the Spanish context, presenting an opportunity to expand investigations to other countries and to enrich the understanding of FS&A in diverse educational settings. Moreover, there is a noticeable lack of studies focusing on the early years of primary education, which suggests a promising line of research to better understand how FS&A may influence learning and development from the earliest stages. Taken together, these findings highlight the importance of continuing to explore and refine FS&A to optimize its application in PE and to contribute to the holistic development of students during their foundational educational stages.

The main limitations of this review include the concentration of studies conducted in Spain—which limits the generalizability of findings to other cultural contexts—and the frequent use of small samples, particularly in qualitative studies involving teachers, which reduces the representativeness of the results. In addition, the scarcity of studies focused on the early years of primary education prevents an analysis of FS&A's impact during the initial stages. Finally, the methodological diversity among the included studies may have complicated direct comparisons, emphasizing the need for future research conducted in more varied contexts and with larger samples.

In summary, current lines of research on FS&A in PE at the Early Childhood and Primary Education stages highlight its positive impact on students' holistic development and on the improvement of teaching–learning processes. However, this review also underscores the need to further explore under-researched areas—such as Early Childhood Education, the first years of Primary Education, and more diverse geographical contexts—in order to optimize the implementation of FS&A and maximize its benefits in the educational field.

Ethics Committee Statement

This study did not require approval from an ethics committee, as it is a systematic review based exclusively on previously published literature and does not involve the collection of primary data or the direct participation of humans or animals.

Conflict of Interest Statement

The authors declare that they have no economic, professional, or personal conflicts of interest that could have influenced the content of this article.

Funding

Include in this section any internal or external funding the research has received, specifying 'This research was funded by NAME OF FUNDER, grant number XXX' or 'This research received no funding due to…'. You can check the details of the source of funding at https://search.crossref.org/funding.

Authors' Contribution

Conceptualization M-B.R., C-J.D., S-M.A. & S-S.G.; Methodology M-B.R. & C-J.D.; Software M-B.R. & C-J.D.; Validation M-B.R., C-J.D., S-M.A. & S-S.G.; Formal Analysis M-B.R., C-J.D., S-M.A. & S-S.G.; Investigation M-B.R., C-J.D., S-M.A. & S-S.G.; Resources M-B.R. & C-J.D.; D; Data Curation S-M.A. & S-S.G.; Writing – Original Draft M-B.R., C-J.D., S-M.A. & S-S.G; Writing – Review & Editing S-M.A. & S-S.G.; Visualization S-M.A. & S-S.G.; Supervision M-B.R., C-J.D., S-M.A. & S-S.G.; Project Administration M-B.R. & C-J.D. All authors have read and agreed to the published version of the manuscript.'

Data Availability Statement

Data are available on request from the corresponding author raul.martinez@usal.es

References

Álvarez Méndez, J. M. (2011). *Evaluar para conocer. Examinar para excluir.* Morata.

Barnett, L. M., Ridgers, N. D., Zask, A., & Salmon, J. (2015). Face validity and reliability of a pictorial instrument for assessing fundamental movement skill perceived competence in young children. *Journal of Science and Medicine in Sport, 18*(1), 98–102. https://doi.org/10.1016/j.jsams.2013.12.004

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 Tendencias en Educación Física, Deportes y Recreación, 36*, 37–43. https://doi.org/10.47197/retos.v36i36.66478

Barrientos Hernán, E. J., López-Pastor, V. M., Lorente-Catalán, E., & Kirk, D. (2023). Challenges with using formative and authentic assessment in physical education teaching from experienced teachers' perspectives. *Curriculum Studies in Health and Physical Education, 14*(2), 109–126. https://doi.org/10.1080/25742981.2022.2060118

Black, P., & Wiliam, D. (2009). Developing the theory of formative assessment. *Educational Assessment, Evaluation and Accountability, 21*(1), 5–31. https://doi.org/10.1007/s11092-008-9068-5

Boynton, P. M., & Greenhalgh, T. (2004). Selecting, designing, and developing your questionnaire. *BMJ, 328*(7451), 1312–1315. https://doi.org/10.1136/bmj.328.7451.1312

Cañadas, L. & Santos-Pastor, M. L. (2021). La evaluación formativa desde la perspectiva de docentes noveles en las clases de educación física en primaria y secundaria. *Revista Electrónica Educare, 25*(3), 1-20. https://doi.org/10.15359/ree.25-3.25

Cardenal, V. & Fierro, A. (2003). Componentes y correlatos del autoconcepto en la escala de Piers-Harris. Estudios de Psicología, 24(1), 101–111. https://doi.org/10.1174/021093903321329094

Carless, D. (2015). *Excellence in University Assessment. Learning from award-winning practice*. Routledge.

Carrillo López, P. J. & Hortigüela-Alcalá, D. (2022). Impacto de la evaluación formativa en el estado de peso y calidad de la dieta en escolares de primaria: *Sportis. Scientific Journal of School Sport, Physical Education and Psychomotricity, 8*(1), 19-39. https://doi.org/10.17979/sportis.2022.8.1.8671

Carrillo López, P. J. (2023). Formative assessment of physical education teachers and self-concept level of primary school children. *Central European Journal of Sport Sciences and Medicine, 41*(1), 35-47. https://doi.org/10.18276/cej.2023.1-04

Carrillo López, P., Maneiro-Dios, R., & Moral-García, J. (2022). Formative assessment in Physical education and its relation to the level of attention of primary school children. *Specijalna Edukacija i Rehabilitacija, 21*(3), 147-161. https://doi.org/10.5937/specedreh21-36486

Cerón, C., Cossio-Bolaños, M., Pezoa-Fuentes, P. & Gomez-Campos, R. (2020). Diseño y validación de un cuestionario para evaluar desempeño docente asociado a las prácticas evaluativas formativas. *Revista Complutense de Educación, 31*(4), 463–472. https://doi.org/10.5209/rced.65512

Chan, C., Ha, A., & Ng, J. Y. Y. (2016). Improving fundamental movement skills in Hong Kong students through an assessment for learning intervention that emphasizes fun, mastery, and support: The A + FMS randomized controlled trial study protocol. *SpringerPlus, 5*(1), 724. https://doi.org/10.1186/s40064-016-2517-6

Chávez Ruiz, Y. & Martínez Rizo, F. (2018). Evaluar para aprender: Hacer más compleja la tarea a los alumnos. *Educación Matemática, 30*(3), 211–246. https://doi.org/10.24844/EM3003.09

Cebrián, B., Isabel, M. & Miguel, A. (2013). *Cómo trabajar la motricidad en el aula. Cuñas motrices para Infantil y Primaria*. Miño y Dávila.

Chng, L., & Lund, J. (2018). The impact of formative assessment on the success and response rate in a 6th grade physical education badminton unit. *International Journal of Kinesiology in Higher Education, 3*(1), 12-22. https://doi.org/10.1080/24711616.2018.1509680

Fraile, J., Gil-Izquierdo, M., Zamorarno-Sande, D. & Sánchez-Iglesias, I. (2020). Autorregulación del aprendizaje y procesos de evaluación formativa en los trabajos en grupo. *RELIEVE – Revista Electrónica de Investigación y Evaluación Educativa, 26* (1). https://doi.org/10.7203/relieve.26.1.17402

González-González, C. S., Cairós-González, M. & Navarro-Adelantado, V. (2014). Validación de un instrumento para evaluación emocional en niños y niñas: "EMODIANA." *Revistafaz, 7*, 160–174.

Harter, S. (1985a). *Manual for the social support scale for children*. University of Denver.

Harter, S. (1985b). *Self-perception profile for children (revision of the perceived competence scale for children)*. University of Denver.

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

Hortigüela, D., Pérez-Pueyo, Á. & 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). https://doi.org/10.15366/riee2019.12.1.001

Lockwood, C., Munn, Z., & Porritt, K. (2015). Qualitative research synthesis methodological guidance for systematic reviewers utilizing meta-aggregation. *International Journal of Evidence-Based Healthcare, 13*(3), 179–187. https://doi.org/10.1097/XEB.0000000000000062

López-Benavente, A., Ureña-Ortín, N. & Alarcón, F. (2019). Evaluación formativa y compartida para la inclusión de descansos activos en infantil. *Journal of Sport and Health Research, 11*(Supl 1), 143- 154. https://recyt.fecyt.es/index.php/JSHR/article/view/80939

López-Pastor, V. M., Molina Soria, M., Pascual Arias, C. & Manrique Arribas, J. C. (2020). La importancia de utilizar la evaluación formativa y compartida en la formación inicial del profesorado de educación física: Los proyectos de aprendizaje tutorado como ejemplo de buena práctica. *Retos. Nuevas Tendencias en Educación Física, Deporte y Recreación, 37*, 620–627. https://doi.org/10.47197/retos.v37i37.74193

López-Pastor, V. (1999). La evaluación formativa: Una propuesta para la mejora continua de la enseñanza y del aprendizaje. *Revista de Investigación Educativa, 17*(2), 309–333.

López-Pastor, V. M. (2017). Evaluación formativa y compartida: Evaluar para aprender y la implicación del alumnado en los procesos de evaluación y aprendizaje. In V. M. López-Pastor & Á. Pérez-Pueyo (Eds.), *La evaluación formativa y compartida en educación: Experiencias de éxito en todas las etapas educativas* (pp. 34–69). Universidad de León.

López-Pastor, V. & Sicilia-Camacho, A. (2017). Formative and shared assessment in higher education. Lessons learned and challenges for the future. *Assessment & Evaluation in Higher Education, 42*(1), 77–97. https://doi.org/10.1080/02602938.2015.1083535

Martínez-Benito, R. (2022). *El aprendizaje cooperativo en educación física desde la perspectiva docente: Dificultades y estrategias para su correcta aplicación* [Tesis doctoral, Universidad de Salamanca]. https://doi.org/10.14201/gredos.149615

Moher, D., Liberati, A., Tetzlaff, J., & Altman, D. G. (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *PLoS Medicine, 6*(7), e1000097. https://doi.org/10.1371/journal.pmed.1000097

Molina Soria, M., Pascual Arias, C. & López Pastor, V. M. (2020). El uso de sistemas de evaluación formativa y compartida en las aulas de educación física en educación primaria. *Educación Física y Deporte, 39*(1), 47-71. https://doi.org/10.17533/udea.efyd.v39n1a04

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

Palacios-Gómez, J., Hortigüela-Alcalá, D. & Hernando-Garijo, A. (2023). Fears and perceptions of the application of formative and shared assessment by a novel physical education teacher. *Espiral. Cuadernos del profesorado, 16*(34), 28-39. https://doi.org/10.25115/ecp.v16i34.9627

Pérez-Pueyo, Á., 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

Prochaska, J. J., Sallis, J. F., Slymen, D. J., & McKenzie, T. L. (2003). A longitudinal study of children's enjoyment of physical education. *Pediatric Exercise Science, 15*(2), 170–178.

Serra-Majem, L., Ribas, L., Ngo, J., Ortega, R. M., García, A., Pérez-Rodrigo, C., & Aranceta, J. (2004). Food, youth and the Mediterranean diet in Spain. Development of KIDMED, Mediterranean Diet Quality Index in children and adolescents. *Public Health Nutrition, 7*(7), 931–935. https://doi.org/10.1079/PHN2004556

Thurstone, L. L., & Yela, M. (2019). *Caras-R: Test de percepción de diferencias-revisado: Manual [R-Faces: Difference perception test-revised: Manual].* Tea.

Wiliam, D., & Thompson, M. (2007). Integrating assessment with instruction: What will it take to make it work? In C. A. Dwyer (Ed.), *The future of assessment: Shaping teaching and learning* (pp. 53–82). Lawrence Erlbaum Associates.