

## Diseño y validación de un cuestionario mediante Método Delphi para valorar las relaciones entre metodologías activas y evaluación formativa en Educación Física en Primaria

### Design and validation of an instrument to assess the relationship between active methodologies and formative assessment in Physical Education, using Delphi Method

\*Alba Arijia-Mediavilla, \*María Luisa Santos-Pastor, \*Luis Fernando Martínez-Muñoz, \*\*Pedro Jesús Ruiz-Montero  
\*Universidad Autónoma de Madrid (España), \*\*Universidad de Granada (España)

**Resumen.** La literatura científica y la legislación educativa actual ponen de manifiesto la necesidad de llevar a cabo un cambio metodológico que tome como punto de partida los procesos de evaluación compartidos y formativos. El auge de las metodologías activas es una constante en las clases de Educación Física en la etapa de Primaria. Sin embargo, son escasas las referencias a la evaluación y limitados los instrumentos para valorar la coherencia entre ambos procesos. Este artículo presenta la construcción y validación de un instrumento con el fin de detectar la relación entre los procesos de evaluación (calificación) y la metodología implementada por el profesorado. Para ello se ha empleado la consulta a personas expertas con método Delphi y su posterior pilotaje. El resultado final obtenido ha sido un cuestionario denominado MAEV-EFP. Las limitaciones encontradas se vinculan con la falta de claridad terminológica, así como los escasos conocimientos acerca de la metodología y evaluación (calificación) de los procesos de enseñanza-aprendizaje en Educación Física en Primaria.

**Palabras clave:** Metodologías Activas, Evaluación, Educación Física, Método Delphi, cuestionario

**Abstract.** Scientific research and current educational legislation supports methodological change. This change is based on shared and formative assessment process. Although active methodologies are present in Physical Education in Primary school, there are few references to evaluation and limited instruments to assess the coherence between the two processes. This paper shows the development and validation of an instrument designed to detect the relationship between the assessment process and the methodology implemented by teachers. The instrument has been validated by Delphi method and a pilot test. MAEV-EFP was the resulting questionnaire. The lack of terminological clarity and knowledge about methodology and assessment of teaching-learning process in Physical Education in Primary school, were some of the limitations found.

**Key words:** active methodologies, assessment, physical education, Delphi method, instrument

---

Fecha recepción: 23-08-23. Fecha de aceptación: 21-11-23

Alba Arijia-Mediavilla  
alba.arija@estudiante.uam.es

### Introducción

La finalidad última del proceso de enseñanza-aprendizaje en un contexto educativo es garantizar una formación integral que contribuya al pleno desarrollo de la personalidad del alumnado. En este marco, son fundamentales los procesos reflexivos que favorezcan el meta-aprendizaje y que promuevan la participación del alumnado y el desarrollo de su autonomía (Artículo 6. Real Decreto 157/2022). La búsqueda de una Educación Física (en adelante EF) que garantice el desarrollo de esta finalidad, requiere de la utilización de metodologías más abiertas que promuevan la participación del alumnado y le involucren en el proceso (León-Díaz, 2020).

La revisión de la bibliografía en el ámbito de conocimiento, pone de manifiesto un incremento de los estudios relacionados con las metodologías activas en EF, así como su implementación en los centros educativos (García, 2017). Si bien, la falta de formación percibida por el profesorado (León-Díaz et al., 2020) abre nuevas vías a la investigación orientadas a analizar si la implementación de estas metodologías se está realizando de acuerdo con los principios pedagógicos de las mismas (Pérez-Pueyo & Hortigüela-Alcalá, 2020). En este sentido, para que el cambio metodológico sea coherente, se ha de articular con procesos de evaluación acordes a los principios de participación y responsabilidad del alumnado (Fernández-Río et al., 2016). Así, para que se puedan implementar correc-

tamente las propuestas metodológicas activas, es necesario incluir sistemas de evaluación formativos y compartidos, que involucren al alumnado en su aprendizaje y lo hagan verdaderamente significativo (Barrientos-Hernán et al., 2019; López-Pastor & Pérez-Pueyo, 2017; Pérez-Pueyo et al., 2020). Esta afirmación nos lleva a preguntarnos ¿qué está ocurriendo realmente en las aulas de EF?, ¿existe coherencia entre las metodologías y los procesos de evaluación en EF?

La evaluación formativa en EF se ha consolidado como un enfoque favorecedor de la toma de decisiones del alumnado (Chng & Lund, 2018; González-Palacio et al., 2021) y del profesorado en el proceso de enseñanza-aprendizaje. Además, facilita y mejora la intervención del profesorado incidiendo en un aprendizaje auténtico del alumnado y una evaluación orientada al aprendizaje (Barrientos-Hernán et al., 2020). A partir de la experiencia de investigación en evaluación formativa y compartida, algunos autores (López-Pastor, et al., 2007) detallan algunas ventajas de esta evaluación en el proceso de enseñanza-aprendizaje, entre las que resaltamos: (1) el valor formativo en todas las etapas educativas, al permitir al alumnado tomar conciencia sobre el proceso de aprendizaje e implicarle en su progreso; (2) la estrecha relación con el desarrollo de la autonomía del alumnado, ya que se siente más partícipe del proceso; y (3) mayor profundidad de la información que aporta sobre el proceso y el incremento de la motivación del alumnado hacia la práctica y, en última instancia, hacia la asignatura.

Junto a estas ventajas, la investigación desarrollada a lo largo de los años recoge también una serie de inconvenientes a tener en cuenta a la hora de implementar sistemas de evaluación formativa en el aula tales como: (1) la resistencia al cambio, dado el peso de las costumbres y rutinas en la organización de las sesiones y los procesos de evaluación; (2) la duración del proceso de aprendizaje y adaptación a las nuevas propuestas, especialmente en los primeros cursos; (3) la viabilidad, dada la excesiva carga de trabajo que puede suponer instaurar un proceso de evaluación formativa; y (4) la importancia de recordar la esencia de la evaluación formativa como proceso cualitativo en la fase de calificación y la calificación dialogada (López-Pastor et al., 2013).

Por otro lado, las investigaciones sobre evaluación formativa en EF, abordadas desde enfoques prácticos, se han llevado a cabo mayoritariamente en Educación Secundaria y en Educación Superior (Barrientos & López-Pastor, 2017; León-Díaz et al., 2023). Sin embargo, son pocos los estudios que se centran en Educación Primaria (García-Herranz & López-Pastor, 2019; Jihuallanca et al., 2023), más concretamente, en la asignatura de EF (Ní Chróinín & Cosgrave, 2013; Herranz, 2014; López-Pastor et al., 2005; Torres & López, 2007). De forma más concreta, respecto a la vinculación entre el uso de metodologías activas y procesos de evaluación formativos y compartidos, pese a que la bibliografía recoge la relación inherente entre ambos elementos (Hortigüela-Alcalá, et al., 2020), no encontramos estudios que aporten datos sobre lo que está ocurriendo en las aulas. Estudios previos analizan, por separado, la implementación de sistemas de evaluación formativa a corto plazo en la etapa de Educación Primaria (Cañas, et al., 2019; Fernández-Gracimartín et al., 2019; Ní Chróinín & Cosgrave, 2013), o el uso de metodologías activas como estrategia metodológica (León-Díaz et al., 2020).

De acuerdo con la literatura, los términos metodología, evaluación y calificación están interrelacionados entre sí, conformando la base conceptual para comprender y mejorar los procesos educativos, contribuyendo al desarrollo de prácticas pedagógicas más efectivas. Dada la amplitud y ambigüedad de los términos, consideramos necesario concretar teóricamente qué entendemos por metodología, evaluación y calificación. Así, la metodología podría definirse el conjunto de estrategias, técnicas y enfoques pedagógicos que utilizan los docentes para desarrollar el proceso de enseñanza-aprendizaje. Estaríamos hablando del cómo enseñar (Pérez-Pueyo et al., 2008). La evaluación se constituye como el eje vertebrador del proceso de enseñanza-aprendizaje, siendo un proceso sistemático y continuo de recopilación, análisis e interpretación de información, para la elaboración de un juicio de valor que permita tomar decisiones sobre el proceso (López-Pastor, 2006). La calificación por su parte, constituye un proceso independiente a la evaluación, representando la asignación de valores numéricos o descriptivos al nivel de logro alcanzado en el proceso.

Desde este marco de antecedentes expuestos y teniendo presente qué se entiende por cada uno de los términos, surge la necesidad de conocer los procesos de evaluación y calificación que lleva a cabo el profesorado de EF de Primaria y su relación con la metodología utilizada. Para ello, es necesaria la elaboración de un instrumento que permita explorar dichos procesos en contextos reales de práctica. En este sentido, este trabajo tiene como objetivo presentar la construcción y validación de un instrumento que permita conocer la relación entre los procesos de evaluación (calificación) y la metodología en EF implementada por el profesorado.

## Método

Con el fin de responder al objetivo del estudio, se ha construido un instrumento de evaluación denominado cuestionario MAEV-EFP. La construcción del instrumento se ha llevado a cabo en 5 fases: (1) elaboración inicial de su estructura y componentes a partir de la revisión de la bibliografía y la validación interna de su contenido en el grupo de investigación; (2) consulta a un grupo de expertos (Método Delphi); (3) aplicación de una prueba piloto para comprobar su adecuación en la realidad; (4) aplicación de una prueba de validación de caso único; y (5) análisis y debate entre el equipo de investigación para ajustar los componentes del instrumento a partir de la información recogida en su aplicación. Para el análisis de los datos se utilizó el programa estadístico SPSS v.28.

### *Fase 1. Elaboración inicial de la estructura del cuestionario*

El proceso de construcción del cuestionario comenzó con una revisión de la literatura. Para ello se procedió a hacer una búsqueda en las bases de datos indexadas (WOS y Scopus) empleando los términos (tanto en español como en inglés) “Metodologías Activas y Educación Física”, “Metodología y evaluación”, “Metodologías activas y evaluación formativa”, “Evaluación formativa y Educación Física”, abarcando los años 2010 a 2022, debido a su actual interés. Para el procesamiento de estas referencias se generó un proyecto en la base de datos Mendeley y, posteriormente, se analizó a través de N-Vivo v.18. Así, a partir de la selección, clasificación y análisis de las referencias vinculadas con los procesos de evaluación (calificación) y las metodologías activas, se obtuvo una primera estructura del cuestionario, organizada en dimensiones y variables. De este modo, tras un proceso de análisis y debate en el equipo investigador, se propusieron tres dimensiones para el cuestionario. Estas se concretaron en: (1) Datos personales; (2) Metodología- tipo de metodología empleada, uso de las metodologías activas y coherencia; y (3) Evaluación y Calificación – finalidad (por qué, para qué evalúas), técnicas, pruebas, instrumentos (finalidad de los instrumentos), corrección del alumnado, momentos, participación del alumnado, calificación (obtención de la calificación y tipo de calificación) (ver Figura 1).

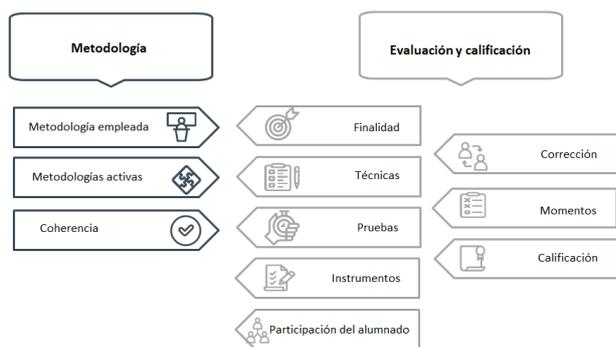


Figura 1. Dimensiones y variables iniciales del cuestionario

Cada indicador se describió de manera precisa, dando prioridad a la explicación detallada del mismo. De este modo, se obtuvo una primera versión de la herramienta, que fue debatida y ajustada por el grupo de investigación a partir de la base de datos creada en Mendeley. En esta fase participó un grupo de investigación formado por 3 personas, dos de ellas pertenecientes a la Red de Evaluación Formativa y Compartida en Educación. La duración de este trabajo fue de tres meses (agosto-octubre de 2021).

### Fase 2. Consulta a grupo de personas expertas. Método Delphi

A partir de esta primera versión del cuestionario, se procedió a realizar la validación con personas expertas mediante el Método Delphi. Este método indica que es necesario crear un grupo de personas expertas sobre un tema o ámbito para que revisen el instrumento, de manera anónima, en sucesivas rondas y aporten su valoración sobre la adecuación de todos los ítems propuestos (dimensiones, variables e ítems), fruto de la interacción y circulación de información (López-Vidal & Llunch, 2019). Para la selección de personas expertas se concretaron unos criterios de inclusión (López-Gómez, 2018): voluntad e interés para participar en el estudio, disponibilidad, nivel de experiencia sobre el objeto de estudio y nivel de conocimiento sobre el tema. Además, se tuvo en cuenta que hubiera una representación de profesorado universitario, profesorado de EF en Educación Primaria y profesorado de EF en Educación Primaria que imparte docencia en la universidad como profesorado asociado. Tras el periodo de captación, finalmente participaron 11 personas de 35 contactadas (Ver tabla 1).

Tabla 1.  
Características de los participantes Método Delphi

Participantes	Ámbito					
	Universitario		Escolar		Universitario y Escolar	
	n	%	n	%	n	%
Hombres	3	27.3	0	-	3	27.3
Mujeres	3	27.3	2	18.2	0	-

Fuente: Elaboración propia. El porcentaje está calculado respecto al total de los participantes.

El número de consultas estuvo marcado por el consenso obtenido de acuerdo a los criterios de selección. Se realizaron tres rondas de consulta (R1, R2 y R3) y todo el

procedimiento se recogió en una guía que se facilitó a las personas participantes al inicio del proceso (Ver Figura 2).

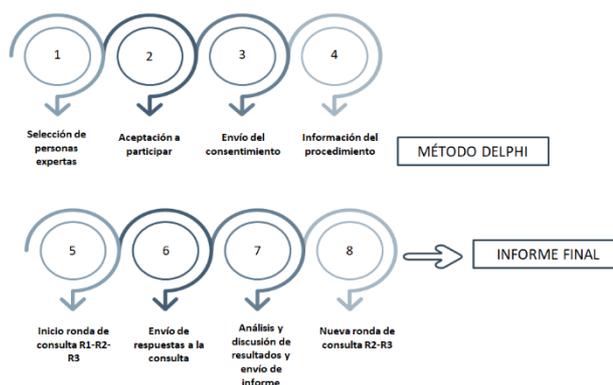


Figura 2. Proceso de consulta a través del método Delphi

Para cada ronda se determinó un tiempo de respuesta de 15-20 días. El proceso de respuesta consistió en valorar, con una escala Likert (1 al 4), la relevancia y claridad de cada ítem propuesto en el cuestionario. Además, se ofreció un espacio abierto para realizar observaciones y reflexiones. Posteriormente, el grupo de investigación analizó (cuantitativa y cualitativamente) los resultados recogidos en un informe final con el que volver a comenzar la siguiente ronda de consulta (Ver Figura 3). Para valorar la validez del contenido, se analizaron los datos obtenidos a partir del juicio de expertos y se calcularon las medias y desviaciones típicas para obtener el valor relativo a necesidad, adecuación y claridad de los ítems (Moliner-Miravet et al., 2017). Como grupo coordinador, de acuerdo George & Trujillo (2018), decidimos modificar o eliminar aquellos ítems que obtuvieran una puntuación inferior al 75% de la media, de acuerdo a valoraciones del grupo de personas expertas. En la escala de respuestas del 1 al 4 este porcentaje equivalía a un índice de 3. Sin embargo, al obtener en todos los ítems valores superiores a 3, se optó por revisar y modificar los ítems con una puntuación inferior al 85%, lo que en la escala del 1 a 4 equivalía a un índice de 3.4. Tras este análisis estadístico y teniendo en cuenta las conclusiones extraídas del análisis cualitativo de las respuestas, se señalaron en verde los ítems validados, en gris los eliminados y en blanco aquellos que debían someterse a la siguiente ronda de validación. Tras la tercera y última ronda de consulta, se reelaboró el cuestionario MAEV-EFP, iniciando así la siguiente fase en el proceso de validación. La duración de este trabajo fue de tres meses (octubre-diciembre de 2021).

### Fase 3. Aplicación de la prueba piloto

El propósito de esta fase fue valorar la calidad y pertinencia de los ítems del cuestionario en un contexto de aplicación real. Participaron 16 maestros y maestras de EF en activo (ver Tabla 2) y evaluaron la adecuación de los ítems en función de tres parámetros: (1) se comprende la pregunta y su finalidad; (2) la redacción es correcta y clara; y (3) se puede responder de manera sencilla. La duración de

este trabajo fue de un mes (enero 2022). Durante el proceso, las personas participantes de la prueba piloto rellenaron el cuestionario valorando la pertinencia de cada ítem y realizando las preguntas, observaciones y comentarios que consideraban pertinentes. Del análisis cualitativo de las respuestas y reflexiones obtenidas se consideró la necesidad de aclarar y detallar algunos términos específicos que generaban controversia en el cuestionario (especialmente en la dimensión de evaluación), así como modificar la tipología de respuesta para convertirla en escala Likert para poder atender a todas las necesidades de respuesta y obtener resultados de mayor calidad. Con la información obtenida, se elaboró una nueva versión del cuestionario MAEV-EFP.

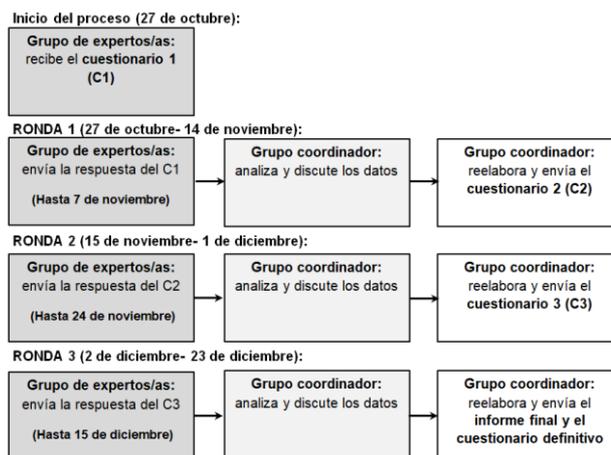


Figura 3. Rondas de consulta Método Delphi

Tabla 2.

Características de los participantes Prueba piloto

	Sexo		Edad					Experiencia						
	H	M	<30	31-40	41-50	51-60	>60	<1	1-3	4-6	7-12	13-18	>18	
N	11	5	2	12	2	0	0	0	2	5	5	3	1	
%	68.75	31.25	12.5	75	12.5	-	-	-	12.5	31.25	31.25	18.75	6.25	

Fuente: Elaboración propia

El porcentaje está calculado respecto al total de los participantes.

#### Fase 4. Aplicación de la prueba de validación de caso único

Dentro del proceso de pilotaje, se sometió el cuestionario a un proceso de validación de caso único, dada la modificación de la tipología de respuestas realizada tras la aplicación de la prueba piloto. El proceso se realizó siguiendo los mismos parámetros e instrumentos de la prueba piloto, seleccionando al participante en función de tres parámetros: (1) cumplimiento de los requisitos de la población a la que va destinado el cuestionario; (2) conocimiento del objeto de estudio (metodologías activas y evaluación formativa); y (3) experiencia en la implementación de metodologías activas y evaluación formativa en el aula. La aplicación de la validación de caso único se realizó en marzo de 2022.

#### Fase 5. Puesta en común y elaboración del instrumento definitivo

Finalmente, tras el proceso de pilotaje, se realizó un debate amplio entre el equipo investigador para valorar conjuntamente los resultados recogidos tras el proceso de validación del cuestionario. Los resultados obtenidos tras esta última fase han dado lugar al cuestionario definitivo MAEV-EFP, el cual está compuesto por cuatro dimensiones y 12 variables. Estas se concretaron en: (1) Datos Personales; (2) Metodología (metodología empleada y metodologías activas); (3) Evaluación (modelo de evaluación, coherencia, dificultades, motivación, finalidad, técnicas e instrumentos, agentes encargados, momento y feedback); y (4) Calificación (modelo de calificación) (ver Figura 4). En esta fase participaron tres personas. La duración de este trabajo fue de 1 mes (abril 2022).

#### Criterios de rigor de la investigación

Para finalizar, hay que señalar que en las diferentes fa-

ses de este trabajo se han considerado los siguientes criterios de rigor y calidad de la investigación (Smith & Sparkes, 2019): credibilidad de los resultados, transferibilidad del estudio, consistencia y confiabilidad del instrumento. Asimismo, esta investigación ha tenido en cuenta los procedimientos éticos y de tratamiento de datos aprobados, tras su revisión, por el Comité Ético de la Universidad Autónoma de Madrid (CEI-131-2718).

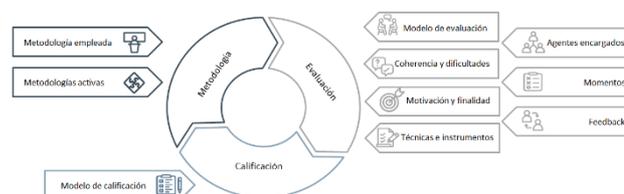


Figura 4. Dimensiones y variables del cuestionario definitivo

## Resultados y discusión

Estructuramos la presentación de los resultados en dos apartados: resultados derivados del proceso de validación por medio del método Delphi y resultados obtenidos tras el proceso de pilotaje.

#### Resultados método Delphi

En este apartado se muestran los resultados cuantitativos y cualitativos obtenidos durante las diversas rondas de consulta a las personas expertas. Los resultados cuantitativos obtenidos se han estructurado en función de las rondas de validación realizadas. En total se realizaron tres rondas de consulta, en las que paulatinamente se fue reduciendo y clarificando el contenido del cuestionario. En cuanto a los resultados cualitativos, se han estructurado en función de las dimensiones y variables constituidas en el cuestionario,

de manera que agrupan y sintetizan los resultados obtenidos de las tres rondas.

### Resultados cuantitativos

#### Resultados 1ª ronda

Para la validación del instrumento creado, se preguntó al grupo de personas expertas sobre la relevancia y claridad de las 3 dimensiones y 23 ítems del cuestionario inicial. De acuerdo a lo especificado en el apartado de método, como grupo coordinador decidimos modificar o eliminar aquellos ítems que obtuvieran una puntuación inferior al 75% de la media, de acuerdo a valoraciones del grupo de personas expertas. En la escala de respuestas del 1 al 4 (siendo 1 nada adecuado y 4 muy adecuado) este porcentaje equi-

valía a un índice de 3. Sin embargo, al obtener en todos los ítems valores superiores a 3, se optó por revisar y modificar los ítems con una puntuación inferior al 85%, lo que en la escala del 1 a 4 equivalía a un índice de 3.4. Tras este análisis estadístico y teniendo en cuenta las conclusiones extraídas del análisis cualitativo de las respuestas, se presentaron los ítems validados, eliminados y aquellos que debían someterse a la siguiente ronda de validación (Tabla 3). Mencionar que, para la eliminación o modificación de los ítems en las tres rondas de consulta, además de los datos estadísticos, se tuvieron en cuenta las aportaciones del grupo de personas expertas (análisis cualitativo de los datos).

Tabla 3.  
Media, Desviación Típica y decisión adoptada tras la ronda 1

Ítem	Media	Desviación Típica	Decisión
<b>Datos personales</b>			
1. Sexo	3.89	.333	Validado
2. Edad	3.44	.527	Modificado el rango de edad
3. Años de experiencia docente	3.78	.667	Modificado el rango de años
4. Cursos en los que imparte docencia	4.00	.000	Validado
5. Titularidad del centro educativo	3.78	.441	Validado
6. Comunidad Autónoma	3.78	.441	Validado
<b>Metodología</b>			
7. ¿Qué tipo de metodología predomina en tus clases?	3.54	.726	Modificado. Se suprimen los términos
8. ¿Qué metodologías y/o estrategias didácticas predominan en tus clases?	3.67	.500	Modificado el enunciado y añadido modelo de educación deportiva
9. La propuesta de evaluación que planteas, ¿está condicionada por la propuesta metodológica que realizas?	3.33	1.118	Modificado el enunciado y tipología de respuesta
<b>Evaluación</b>			
10. ¿Qué finalidad cumple para ti la evaluación?	3.78	.441	Modificado el enunciado y se añaden respuestas
11. ¿Por qué evalúas?	3.56	.527	Modificado. Cambio de orden con la anterior
12. ¿Para qué evalúas?	3.56	.527	Eliminado. Se combina con la 10
13. ¿Qué técnica o técnicas de evaluación llevas a cabo?	3.56	.726	Modificado
14. ¿Qué pruebas o evidencias utilizas para evaluar?	3.22	.833	Eliminado. Se combina con la 13
15. ¿Qué instrumentos de evaluación utilizas?	3.67	.707	Validado
16. ¿Qué finalidad cumplen los instrumentos de evaluación que utilizas?	3.22	.972	Eliminado
17. ¿En qué momentos del proceso de enseñanza-aprendizaje utilizas las técnicas e instrumentos de evaluación?	4.00	.000	Validado
18. ¿Cómo obtienes la calificación del alumnado?	3.44	1.130	Modificado. Se cambia el enunciado y la tipología de respuesta
19. Cuando corrijo al alumnado durante la práctica...	4.00	.000	Modificado. Se añaden dos respuestas más. Se coloca debajo de la 17
20. En tus clases de EF ¿Participa el alumnado en los procesos de evaluación?	3.65	.707	Eliminado. Se integra dentro del ítem 21
21. ¿De qué forma participa?	3.44	.882	Modificado. Se modifica el enunciado y las respuestas. Pasa a ser el ítem 8 en el apartado de Evaluación
22. En tus clases de EF ¿participa el alumnado en la toma de decisiones sobre la calificación?	3.33	1.118	Eliminado
23. ¿Qué tipo de calificación llevas a cabo?	3.44	.882	Eliminado

Fuente: Elaboración propia

Como puede observarse en la tabla 3, 6 de los ítems fueron eliminados o integrados en otros, 6 validados y 11 modificados, propuestos para una nueva ronda de validación. Mencionar además que, fruto del análisis cualitativo de las respuestas, se incorpora una nueva dimensión, *calificación*, y se añaden nuevos ítems en la dimensión de *datos personales*: universidad y localidad en la que cursa los estudios académicos y situación administrativa.

#### Resultados 2ª ronda

Siguiendo la estructura de la ronda de consulta anterior, para la validación del instrumento creado se preguntó

al grupo de personas expertas por la relevancia y claridad de los ítems reformulados y propuestos para la revisión. Así, en esta segunda ronda de consulta se revisaron un total de 13 ítems, correspondientes a cuatro dimensiones. En la Tabla 4 se muestran los resultados obtenidos tras la realización del análisis descriptivo de las respuestas cuantitativas, teniendo en cuenta las respuestas cualitativas del grupo de personas expertas.

Como puede observarse en la tabla 4, de los 13 ítems propuestos para su validación, 1 fue eliminado de manera consensuada, 7 validados y 5 modificados, propuestos para una nueva ronda de validación. La necesidad de profundi-

zar en mayor medida en las respuestas relativas a la dimensión de Evaluación, a partir de las reflexiones proporcionadas por el grupo de personas expertas, se incorporó un nuevo ítem a esta dimensión. De acuerdo con esto, quedaron 6 ítems pendientes de revisión tras su modificación,

para la 3ª y última ronda de consulta.

#### Resultados 3ª ronda

Finalmente, en la 3ª y última ronda de consulta, se revisó la pertinencia y calidad de 6 ítems, quedando 4 de ellos validados y 2 aceptados con modificaciones (Tabla 5).

Tabla 4.

Media, Desviación Típica y decisión adoptada tras la ronda 2

Ítem	Media	Desviación Típica	Decisión
Datos personales			
1. Edad	3.43	.787	Modificado. Pregunta abierta
2. Años de experiencia docente	3.57	.787	Modificado. Se añade la opción de 1-3 años por ser un momento determinante en la carrera profesional del docente
3. Universidad y localidad en la que cursó sus estudios universitarios	3.57	.787	Validado
4. Situación administrativa en la que se encuentra	3.29	1.254	Eliminado
Metodología			
5. ¿Qué tipo de metodología predomina en tus clases?	3.57	.535	Validado
6. ¿Qué metodologías, estrategias didácticas y/o modelos pedagógicos predominan en tus clases?	4.00	.000	Validado
7. ¿Realizas una evaluación acorde a los principios de la metodología que utilizas? ¿Por qué?	3.50	.837	Modificado. Se omiten las explicaciones para evitar condicionar la respuesta
Evaluación			
8. ¿Por qué evalúas?	3.50	.837	Modificado. Se añaden dos opciones de respuesta más y se modifica la segunda
9. En función de la finalidad que tiene para ti la evaluación, ¿Para qué evalúas?	3.67	.816	Validado
10. ¿Qué técnica o técnicas de evaluación llevas a cabo?	4.00	.000	Validado
11. En función de quienes son los encargados de evaluar en tus clases de EF ¿qué tipo de evaluación llevas a cabo?	3.67	.816	Validado
12. Cuando corrijo al alumnado durante la práctica...	3.83	.408	Modificado. Se añade una nueva opción de respuesta y se cambia el orden de respuesta de Nunca a Siempre
Calificación			
13. ¿Cómo obtienes la calificación del alumnado?	3.83	.408	Validado

Fuente: Elaboración propia

Tabla 5.

Media, Desviación Típica y decisión adoptada tras la ronda 3

Ítem	Media	Desviación Típica	Decisión
Datos personales			
1. Edad	3.89	.333	Validado
2. Años de experiencia docente	4.00	.000	Validado
Metodología			
3. ¿Cómo crees que debería ser el proceso de evaluación asociado a las metodologías/estrategias y/o modelos pedagógicos que utilizas mayoritariamente en tus clases de EF?	3.33	.707	Modificado e incorporado en la variable evaluación
4. Es el modelo de evaluación que llevas a cabo en tus clases de EF? ¿Por qué?	3.44	.527	Modificado e incorporado en la variable evaluación
Evaluación			
5. ¿Por qué evalúas?	3.89	.333	Validado
6. Cuando corrijo al alumnado durante la práctica...	3.78	.441	Validado

Fuente: Elaboración propia

Una vez realizadas las modificaciones, se preguntó nuevamente al grupo de personas expertas si las consideraban pertinentes, obteniendo su consenso. Así, tras las 3 rondas de validación y la reflexión del equipo investigador, las dimensiones e ítems del cuestionario fueron los siguientes (Tabla 6).

#### Resultados cualitativos

A partir de los resultados cuantitativos, en cada una de las rondas se fueron analizando los datos cualitativos extraídos de las reflexiones y observaciones realizadas por el grupo de personas expertas, con la finalidad de obtener una mayor comprensión del proceso y dotar de una mayor coherencia al instrumento. A continuación, se recoge el debate generado en torno a la concreción de las distintas variables e ítems. Para su análisis, se han tomado como

referencia las dimensiones y variables del cuestionario una vez finalizado el proceso de validación por el método Delphi.

#### Dimensión Datos personales

Inicialmente, el ítem edad correspondía a una pregunta con opción de respuesta cerrada. Con la finalidad de obtener una mayor precisión en el análisis estadístico, a partir de la segunda ronda de consulta se modifica la categoría de respuesta, constituyéndose como una pregunta de respuesta abierta. Respecto a los años de experiencia docente, al considerarse los primeros años de docencia como determinantes por el grupo de personas expertas, se modifican los rangos de edad establecidos en las opciones de respuesta, reduciendo el intervalo en los primeros años. Se considera que la opción 1 a 6 años es

muy amplia, por lo que se establecen dos periodos, de 1 a 3 años y de 4 a 6 años de experiencia docente. El ítem relativo a la universidad y localidad en la que se cursaron los estudios universitarios generó controversia, al considerarse una pregunta que podía tener una finalidad clasificadora. Si bien, se determinó mantenerla, ya que podía aportar información relevante sobre la relación entre la metodología y los procesos de evaluación utilizados y la

formación inicial recibida. Finalmente, en la primera ronda de consulta se consideró añadir el ítem Situación administrativa, eliminándolo del cuestionario en la segunda ronda de consulta ya que, siguiendo las aportaciones de tres de las personas expertas consultadas, se determinó que la situación en la que se encontrase el profesorado, no era un factor condicionante.

Tabla 6.  
Dimensiones variables e ítems tras la validación método Delphi

Dimensiones	Variables e ítems	Respuesta
Datos personales	Sexo	Cerrada
	Edad	Abierta
	Años de experiencia docente	Cerrada
	Cursos en los que impartes docencia	Cerrada
	Titularidad del centro educativo	Cerrada
	Comunidad Autónoma en la que impartes docencia	Abierta
	Universidad y localidad en la que cursaste los estudios universitarios	Abierta
Metodología	Metodología empleada	
	1. ¿Qué tipo de metodología predomina en tus clases de EF?	Cerrada
	Metodologías activas	
	2. ¿Qué metodologías activas predominan en tus clases de EF?	Cerrada respuesta múltiple
	Coherencia	
	3. ¿Qué tipo de evaluación crees más coherente utilizar al implementar metodologías activas en las clases de EF?	Abierta
	4. Teniendo en cuenta la respuesta anterior, ¿qué modelo de evaluación utilizas habitualmente en tus clases de EF? ¿Por qué?	Abierta
	Finalidad	
Evaluación	5. ¿Por qué evalúas?	Cerrada respuesta múltiple
	6. En función de la finalidad que tiene para ti la evaluación ¿Para qué evalúas?	Cerrada respuesta múltiple
	Técnicas e instrumentos	
	7. ¿Qué técnica o técnicas de evaluación llevas a cabo?	Cerrada respuesta múltiple
	8. ¿Qué instrumentos de evaluación utilizas?	Cerrada respuesta múltiple
	Agentes encargados	
	9. En función de quienes son los encargados de evaluar en tus clases de EF ¿Qué tipo de evaluación llevas a cabo?	Cerrada respuesta múltiple
Calificación	Momentos	
	10. ¿En qué momento o momentos del proceso de enseñanza-aprendizaje utilizas las técnicas e instrumentos de evaluación?	Cerrada respuesta múltiple
	Feedback	
	11. Cuando doy feedback o retroalimentación al alumnado durante la práctica...	Escala de Likert
	Modelo de calificación	
	12. ¿Cómo obtienes la calificación del alumnado?	Cerrada respuesta múltiple

Fuente: Elaboración propia

### Dimensión Metodología

A partir de la primera ronda de consulta se modificaron las opciones de respuesta de la variable metodología empleada, incorporando terminología que no condicionara la respuesta de las personas participantes. Así, los términos tradicional y alternativa fueron sustituidos por: metodología basada en estilos de enseñanza más directivos, donde la toma de decisiones recae mayoritariamente sobre el profesorado y metodología basada en propuestas activas, que promueven un proceso de enseñanza-aprendizaje centrado en el alumnado. Además, se incorporó la opción de especificar el uso de otra metodología, si así se considerase. Respecto a la variable *metodologías activas*, se optó por utilizar este término aglutinando los modelos pedagógicos propios de la EF, con la finalidad de simplificar la comprensión y asociación de los ítems relacionados. Los ítems relativos a variable *coherencia* experimentaron modificaciones a lo largo del proceso de validación del cuestionario derivando, tras las 3 rondas de consulta, en dos preguntas López-Pastor (2006).

que posteriormente (tras la prueba piloto) incorporaríamos en la dimensión de evaluación del cuestionario. La finalidad de estos ítems fue determinar qué tipo de evaluación consideraban coherente utilizar al implementar metodologías activas y qué modelo de evaluación utilizaba habitualmente el profesorado en sus clases de EF.

### Dimensión Evaluación

Las opciones de respuesta e ítems de cada una de las variables se fueron modificando y ajustando a lo largo de las tres rondas de validación con la finalidad de proporcionar a las personas participantes opciones de respuesta relacionadas con la evaluación, acordes a la realidad de su práctica docente. Así, a partir de la primera ronda, se organizaron los aspectos relativos a la finalidad, el *para qué* y el *por qué* de la evaluación del profesorado en dos cuestiones: (1) ¿por qué evalúan?, tratando de identificar la motivación de los docentes, y (2) ¿para qué evalúan?, buscando la finalidad que cumple la evaluación en los procesos de enseñanza-aprendizaje, coincidiendo con lo recogido por

Tras la primera ronda de validación, se consideró que

las pruebas o evidencias que se utilizan para evaluar no tenía sentido al margen de los instrumentos que se utilizan para llevar a cabo dicha evaluación, por lo que quedó incorporada en la dimensión de instrumentos de evaluación (González-Palacio et al., 2021). De igual manera, se consideró que preguntar sobre la finalidad de los instrumentos de evaluación era incidir de nuevo sobre la finalidad que cumple la evaluación para el profesorado, por lo que se optó por eliminar la pregunta. Asimismo, la participación del alumnado en los procesos de evaluación estaba recogida en las opciones de respuesta de tres cuestiones: técnicas de evaluación, participación del alumnado en la evaluación y forma en la que se materializa dicha participación. De acuerdo con la literatura (Chng & Lund, 2018) y las observaciones del grupo de personas expertas, se consideró crear, a partir de la segunda ronda de validación, una nueva variable con una pregunta específica para determinar el tipo de evaluación que se lleva a cabo en función de quienes son las personas que evalúan.

Desde un punto de vista terminológico, en estas variables se sustituyó la palabra coevaluación por evaluación entre iguales y la palabra rúbrica por escala descriptiva. El término corrección del alumnado fue sustituido por feedback o retroalimentación, constituyendo una nueva variable en el cuestionario. Tras las tres rondas de validación se modificó la opción de respuesta, pasando a ser escala de Likert. Por otro lado, se incorporaron como ítems, a los ya validados: dar feedback general a toda la clase y la utilización del alumnado como modelo y ejemplo.

#### *Dimensión Calificación*

Tras la primera ronda de validación, con la finalidad de separar claramente los procesos de evaluación y calificación, se incorporaron los aspectos relativos a la calificación en una nueva dimensión, con la que se buscó determinar cómo obtiene el profesorado la calificación del alumnado. Para tratar de acotar la respuesta del profesorado participante, tras la primera ronda de validación, se determinó convertir el modelo de respuesta en respuesta cerrada e incorporar otra opción que nos permitiera conocer qué, cómo, cuánto y cuándo se califica.

#### *Resultados del proceso de pilotaje*

A partir del cuestionario validado tras la aplicación del método Delphi, con la finalidad de dar respuesta a las demandas y observaciones de las personas participantes de la prueba piloto, se incorporaron al cuestionario una serie de modificaciones que hemos agrupado en tres bloques para facilitar su comprensión.

#### *Escala Likert como opción de respuesta*

La metodología y evaluación en educación no son procesos independientes y fijos en el tiempo, sino que van a depender de variables como la edad del alumnado, el momento del proceso de enseñanza y aprendizaje, los contenidos que se estén trabajando, la unidad de programación

o situación de aprendizaje, el contexto, etc. (Hortigüela-Alcalá et al., 2021). Asimismo, tratar de detectar los modelos, principios o finalidades de estos procesos por medio de preguntas con respuesta única no permite al profesorado reflejar la realidad de su práctica (León-Díaz, 2020). Para atender a esta necesidad manifestada durante el proceso de pilotaje y permitir a las personas participantes ajustar su respuesta a la realidad del aula, se incorporó una escala de tipo Likert con cuatro opciones de respuesta, atendiendo al número óptimo de respuestas recogido en la literatura (Domínguez et al., 2013; Lee & Paek, 2014).

#### *Aclaración terminológica*

La confusión terminológica y los problemas en la definición de los aspectos relacionados con la metodología y la evaluación sigue siendo una constante en el ámbito educativo (Hortigüela-Alcalá et al., 2019). Este aspecto se ha puesto de manifiesto durante el proceso de pilotaje, viéndonos en la necesidad de realizar aclaraciones terminológicas a lo largo del proceso. Para tratar de hacer accesible el cuestionario al profesorado, con independencia de su grado de conocimiento y formación sobre metodología y evaluación, se explicaron o ejemplificaron los ítems que generaban mayor duda o controversia. Éstos se concentraron en las variables de técnicas, instrumentos y agentes implicados. Así, por ejemplo, al instrumento escala gráfica, se le añadió, a modo de descripción: diana de evaluación, caras con emociones, semáforo; o el término evaluación compartida, en la variable agentes implicados, fue explicado como la puesta en común con el profesorado tras la autoevaluación.

Incorporación de las variables modelo de evaluación, coherencia y dificultad en la dimensión evaluación:

Dentro de la dimensión Metodología se plantearon dos preguntas con las que se buscaba detectar el modelo de evaluación llevado a cabo por el profesorado, así como la coherencia entre la metodología implementada y los procesos de evaluación utilizados.

Inicialmente, se diseñaron con opción de respuesta abierta, tratando de buscar la reflexión y amplitud de respuesta de los docentes. Si bien, cuando se aplicó el proceso de pilotaje, este planteamiento generó inseguridad y confusión en el profesorado que manifestó no tener claro qué es lo que tenían que contestar y reclamó opciones de respuesta o ejemplos que les ayudasen a focalizar su reflexión. Así, partiendo del análisis de las respuestas, los comentarios del profesorado participante y de las observaciones realizadas durante el proceso de pilotaje, se determinó por el grupo de investigación sustituir las dos preguntas que generaban confusión y cuyas respuestas no eran representativas para la investigación en la dimensión Evaluación: (1) evaluación que utilizas habitualmente, con tres opciones de respuesta y escala tipo Likert, con la finalidad de enfocar la respuesta de los docentes; (2) consideras que el tipo de evaluación que planteas es coherente con la metodología utilizada, de respuesta abierta; y (3) dificultades que plantea la evaluación, de respuesta abierta.

### Validación de caso único

Modificado el cuestionario tras el pilotaje, se sometió al tercer y último paso en el proceso de validación, una prueba de validación de caso único, en la que quedaron finalmente validados todos los ítems del cuestionario y tras la que únicamente se añadió la posibilidad de especificar qué otra metodología y modelo de evaluación se utiliza, en el caso de haber marcado esa opción de respuesta en la pregunta anterior. Por otro lado, con la finalidad de dejar clara la diferencia entre evaluación y calificación, se especificó en el encabezado de la sección la diferencia entre ambos términos.

Se puede acceder a la estructura completa del cuestionario MAEV-EFP validado a través del siguiente [enlace](#)

### Conclusiones

La construcción del cuestionario MAEV-EFP ha sido fruto del análisis y debate compartido de las personas que constituyen el grupo de investigación, en colaboración con las personas que han formado parte del proceso de validación del cuestionario en sus tres fases, método Delphi, proceso de pilotaje y validación de caso único. Se constata que el método Delphi es adecuado para la validación de instrumentos (Sánchez-Taraza & Ferrández-Berruero, 2022). Además, a la luz de los resultados, se puede afirmar que el instrumento elaborado en forma de cuestionario queda validado cuantitativa y cualitativamente, coincidiendo las decisiones con la literatura existente (Pérez-Pueyo, et al., 2020; Zubillaga & Cañadas, 2021). De este modo, nos permite indagar sobre los procesos de evaluación y calificación llevados a cabo por el profesorado de EF; analizar la coherencia existente entre el modelo de evaluación y la metodología implementada, respondiendo a la inexistencia de instrumentos para valorar la relación entre ambos procesos. Además, puede ser un instrumento valioso para comprender el estado actual de la EF y emprender procesos de innovación para mejorar su calidad.

La complejidad intrínseca en la definición e implementación de procesos metodológicos y de evaluación y calificación, así como la confusión terminológica y en muchos casos, falta de conocimiento científico de los docentes sobre estos aspectos, se ha constituido como una limitación y a la vez inconveniente a la hora de diseñar y validar el instrumento.

### Referencias

- Barrientos, E., & López-Pastor, V.M (2017). Evaluación orientada al aprendizaje en educación superior: últimas tendencias. *Revista Infancia, Educación y Aprendizaje*, 3 (2), 582-587. <https://doi.org/10.22370/ieya.2017.3.2.784>
- 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*, 36, 37-43. <https://doi.org/10.47197/retos.v36i36.66478>
- Barrientos-Hernán, E.J., López-Pastor, V.M., & Pérez-Brunicardi, D. (2020). Evaluación Auténtica y Evaluación Orientada al Aprendizaje en Educación Superior. Una Revisión en Bases de Datos Internacionales. *Revista Iberoamericana de Evaluación Educativa*, 13(2), 67-83. <https://doi.org/10.15366/rie2020.13.2.004>
- Cañas Encinas, M., García Martín, N., Pinedo González, R., & Caballero San José, C. (2019). Aplicación del sistema de evaluación formativa y compartida en un proyecto gamificado. *Revista Infancia, Educación y Aprendizaje (IEYA)*, 5 (2), 212-219. <https://doi.org/10.22370/ieya.2019.5.2.1649>
- Chng, L.S., & Lund, J. (2018). Assessment for Learning in Physical Education: The what, why and how. *Journal of Physical Education, Recreation & Dance*, 89(8), 29-34. <https://doi.org/10.1080/07303084.2018.1503119>
- Domínguez, S., Villegas, G., Yauri, C., Mattos, E., & Ramírez, F. (2013). Propiedades psicométricas de una escala de autoeficacia para situaciones académicas en estudiantes universitarios peruanos. *Revista de Psicología*, 2(1), 27-39.
- Fernández-Gracimartín, C., López-Pastor, V.M., & Pascual-Arias, C. (2019). Aprendiendo a desarrollar sistemas de evaluación formativa y compartida en las prácticas como maestra de Educación Física en Primaria. *Revista Innovación y Buenas Prácticas Docentes*, 8, 119- 131.
- Fernández-Río J., Calderón, A., Hortigüela-Alcalá, D., Pérez-Pueyo, A., & Aznar Cebamano, 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>
- García, A. (2017). *Otra educación ya es posible: una introducción a las pedagogías alternativas*. Litera.
- García-Herranz, S., & López-Pastor, V.M. (2019). Evaluación formativa y compartida en 1º de Educación Primaria: una experiencia de éxito en el área de Matemáticas. *Revista de Innovación y Buenas Prácticas Docentes*, 8, 48- 57.
- George, C.E., & Trujillo, L. (2018). Aplicación del Método Delphi Modificado para la Validación de un Cuestionario de Incorporación de las TIC en la Práctica Docente. *Revista Iberoamericana de Evaluación Educativa*, 11(1), 113-135.
- González-Palacio, E.V., Chaverra-Fernández, B.E., Bustamante-Castaño, S.A., & Toro-Suaza, C.A. (2021). Diseño y validación de un cuestionario sobre las concepciones y percepción de los estudiantes sobre la evaluación en educación física. *Retos*, 40, 317-325. <https://doi.org/10.47197/retos.v1i40.80914>
- Herranz-Sancho, M. (2014). *Desarrollo de procesos de autoevaluación y evaluación compartida en la etapa de educación primaria. Un estudio de casos en el área de educación física* (Tesis doctoral). Universidad de Valladolid.
- 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.
- 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.
- Hortigüela-Alcalá, D., González-Víllora, S., & Herándo-Garijo, A. (2021). Do we really assess learning in physical educa-

- tion? Teacher's perceptions at different educational stages. *Retos*, 42, 643-654.
- Jihuallanca-Ruelas, I., Arenas-Figueroa, M., & Jihuallanca-Ruelas, J.A. (2023). La evaluación formativa en estudiantes de educación primaria: una revisión sistemática. *Ciencia Latina Revista Científica Multidisciplinar*, 7(1), 752-770. [https://doi.org/10.37811/cl\\_rcm.v7i1.4428](https://doi.org/10.37811/cl_rcm.v7i1.4428)
- Lee, J., & Paek, I. (2014). In Search of the Optimal Number of Response Categories in a Rating Scale. *Journal of Psychoeducational Assessment*, 32(7), 663-673.
- León-Díaz, O. (2020). *Concepciones teóricas y realidades prácticas de las metodologías activas en Educación Física. Un estudio en Educación Secundaria en la Comunidad de Madrid* (Tesis doctoral). Universidad Autónoma de Madrid.
- León-Díaz, O., Arijá-Mediavilla, A., Martínez-Muñoz, L.F., & Santos-Pastor, M.L. (2020). Las metodologías activas en Educación Física. Una aproximación al estado actual desde la percepción de los docentes en la Comunidad de Madrid. *Retos*, 38, 587-594.
- León-Díaz, O., Martínez-Muñoz, L.F. & Santos-Pastor, M.L. (2023). Metodologías activas en la Educación Física. Una mirada desde la realidad práctica. *Retos*, 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 XXI*, 21(1), 17-40. <https://doi.org/10.5944/educXXI.15536>
- López-Pastor, V. M., González-Pascual, M., & Barba-Martín, J.J. (2005). La participación del alumnado en la evaluación: la autoevaluación, la coevaluación y la evaluación compartida. *Revista Tándem*, 17, 21-37.
- López-Pastor, V. M. (coord.). (2006). *La 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., Barba-Martín, J. J., Monjas-Aguado, R., Manrique-Arribas, J. C., Heras-Bernardino, C., González-Pascual, M., & Gómez García, J. M. (2007). Trece años de Evaluación Compartida en Educación física. *Revista Internacional de Medicina y Ciencias de la Actividad Física y el Deporte*, 7, 69-86.
- López-Pastor, V. M., Kirk, D., Lorente-Catalán, E., McPhail, A., & McDonald, D. (2013). Alternative assesment in physical education: a review of international literature. *Sport, Education and Society*, 18(1), 57-76. DOI: <https://doi.org/10.1080/13573322.2012.713860>
- López-Pastor, V. M., & Pérez-Pueyo, A. (coord.). (2017). *Evaluación formativa y compartida en educación: experiencias de éxito en todas las etapas educativas*. Universidad de León.
- López-Vidal, F. J., & Lluch, A. C. (2019). Diseño y validación mediante Método Delphi de un cuestionario para conocer las características de la actividad física en personas mayores que viven en residencias. *Retos*, 36, 515-520. <https://doi.org/10.47197/retos.v36i36.69773>
- Moliner-Miravet, L., Aguirre-García-Carpintero, A., Domenech-Vidal, A., Vallet-Bellmunt, T., Vallet-Bellmunt, I. & Alegre-Ansuategui, F. (2017). Diseño, validación y análisis factorial exploratorio y confirmatorio de la escala de actitud Cohesiva para la evaluación de la eficiencia de los talleres de habilidades cooperativas. *Estudios Pedagógicos XLIII*, 1, 213-234.
- Ní Chróinín, D., & Cosgrave, C. (2013): Implementing formative assessment in primary physical education: teacher perspectives and experiences. *Physical Education and Sport Pedagogy*, 18(2), 219-233. <https://doi.org/10.1080/17408989.2012.666787>
- Pérez-Pueyo, A., Martínez-Samperio, C. & Garrote-García, J. (2008): "Metodología: un término confuso que todos parecemos tener claro". En L. Ariza & M. Guillén: *IV Congreso internacional y XXV Nacional de Educación Física*. Universidad de Córdoba y Junta de Andalucía. Córdoba.
- Pérez-Pueyo, A., & Hortigüela-Alcalá, D. (2020). ¿Y si toda la innovación no es positiva en Educación Física? Reflexiones y consideraciones prácticas. *Retos*, 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.
- Real Decreto 157/2022, de 1 de marzo, por el que se establecen la ordenación y las enseñanzas mínimas de Educación Primaria.
- Sánchez-Taraza, L., & Ferrández-Berruero, R. (2022). Aplicación del método Delphi en el diseño de un marco para el aprendizaje por competencias. *Revista de Investigación Educativa*, 40(1), 219-235. <http://doi.org/10.6018/rie.463611>
- Smith, B., & Sparkes, A. (2019). *Routledge Handbook of Qualitative Research in Sport and Exercise*. Routledge.
- Torres García, D., & López Pastor, V. (2007). Una experiencia de evaluación formativa en educación física en educación primaria. *Revista La Peonza: Revista de Educación Física para la paz*, 2, 14-24.
- Zubillaga Olague, M., & Cañadas, L. (2021). Diseño y validación del cuestionario "#EvalEF" para conocer el proceso de evaluación desarrollado por los docentes de educación física. *Retos*, 42, 47-55. <https://doi.org/10.47197/retos.v42i0.86627>