



Feedback formativo y afectivo en Educación Física en secundaria: validación del cuestionario FFAPEF

Perceived formative and affective feedback in secondary Physical Education: validation of the PFAPEF

Autoras

Ana Casares Polo ¹
 Alicia M^a Alonso Martínez ²
 M^a Rosario Romero Martín ³
 Nerea Estrada Marcén ⁴
 Sonia Asún Dieste ⁵

^{1,2} Universidad Pública de Navarra (España)

^{3,4,5} Universidad de Zaragoza (España)

Autora de correspondencia:
 Ana Casares Polo
anacasarespolo71@gmail.com

Recibido: 28-03-26
 Aceptado: 05-05-26

Cómo citar en APA

Casares Polo, A., Alonso Martínez, A. M., Romero Martín, M. R., Estrada Marcén, N., & Asún Dieste, S. (2026). Perceived formative and affective feedback in secondary Physical Education: validation of the PFAPEF. *Retos*, 79, 916-929. <https://doi.org/10.47197/retos.v60.119121>

Resumen

Introducción: el feedback es clave en la evaluación formativa, tanto por la información que aporta como por su impacto emocional. En Educación Física, recientes trabajos resaltan su relación con la motivación y la afectividad; sin embargo, se aprecia escasez de trabajos sobre instrumentos y recursos para su estudio y su puesta en práctica.

Objetivo: se persigue validar el cuestionario "Feedback Formativo y Afectivo Percibido en Educación Física" (FFAPEF) para alumnado de Enseñanza Secundaria Obligatoria (ESO), analizando la relación entre el constructo, y las variables género, nivel educativo y tipo de centro. **Metodología:** se realizaron entrevistas a docentes y alumnado para afinar el cuestionario, que se administró a 259 estudiantes de 1.º a 4.º ESO (12-16 años). Se realizó análisis factorial exploratorio y confirmatorio.

Resultados: se mostraron índices de ajuste adecuados (CFI = .982; TLI = .977; RMSEA = 0.097) y una alta fiabilidad (ω McDonald = 0.922, α Cronbach = 0.905). En relación con el género y el tipo de centro, no se observaron diferencias significativas. En cuanto al nivel, el alumnado de 3.º ESO obtuvo puntuaciones significativamente más altas que el de 1.º y 2.º en el 90 % de los ítems.

Discusión: los análisis factoriales indican una estructura unidimensional, entendiendo el feedback como un proceso donde los aspectos afectivos y técnicos se integran en una misma experiencia de aprendizaje.

Conclusiones: se aporta un cuestionario válido y fiable para evaluar la percepción del feedback formativo y afectivo, siendo una herramienta útil para investigación y práctica docente.

Palabras clave

Afectividad; educación física; educación secundaria; evaluación formativa; feedback formativo.

Abstract

Introduction: Feedback is a key component of formative assessment, both, for the information it provides and for its emotional impact. In Physical Education, recent studies highlight its important relationship to motivation and affectivity; however, there is a lack of research on instruments and resources for its study and implementation.

Objective: This study aims to validate the Perceived Formative and Affective Feedback in Physical Education (PFAPEF) questionnaire for secondary education students, analyzing the relationship between the construct and variables such as gender, educational level, and type of school.

Methodology: Interviews with teachers and students were conducted to refine the questionnaire, which was administered to 259 students from 1st to 4th grade of secondary education (ages 12-16) Exploratory and confirmatory factor analyses were performed.

Results: The results showed adequate fit indices (CFI = .982; TLI = .977; RMSEA = 0.097) and high reliability (McDonald's ω = 0.922; Cronbach's α = 0.905). No significant differences were found in relation to gender or type of school. Regarding educational level, 3rd-grade students scored significantly higher than those in 1st and 2nd grade on 90% of the items.

Discussion: Factor analyses indicated a unidimensional structure which conceptualizes feedback as a process in which affective and technical aspects are integrated into a single learning experience.

Conclusions: This study provides a valid and reliable questionnaire to assess perceptions of formative and affective feedback, serving as a useful tool for both, research and teaching practice.

Keywords

Affectivity; formative assessment; formative feedback; physical education; secondary education.

Introducción

La evaluación formativa se consolida en educación como un enfoque que favorece aprendizajes profundos mediante la recopilación sistemática de información durante el proceso de enseñanza-aprendizaje. La evidencia encontrada consolida a nivel global la importancia de esta evaluación para el aprendizaje (Webb & Galamba, 2026). Asimismo, esta evaluación se reconoce como formativa cuando el alumnado utiliza el feedback para mejorar su rendimiento y el profesorado usa las evidencias para adaptarse a las necesidades de quienes aprenden (Black, 2007; Winstone et al. 2022). Así pues, dentro de ella, el feedback resulta esencial para orientar al alumnado hacia la mejora, el ajuste a objetivos y un mayor compromiso con la tarea (Hattie & Timperley, 2007; Lipnevich & Panadero 2021). Además, este feedback entendido desde el constructivismo como un proceso, y no como meros comentarios puntuales del docente, genera un importante apoyo al aprendizaje autorregulado (Molloy & Boud, 2013). Se destaca también la importancia de una dimensión afectiva en la práctica de la evaluación formativa (Xiang, Yum & Lian, 2023) dado que, a partir de la seguridad y la confianza con el profesorado, se pueden generar entornos positivos que favorezcan la motivación y el aprendizaje en el alumnado.

Centrándonos en la dimensión afectiva del feedback, esta se define como el grado de seguridad, confianza y motivación que el mensaje genera en quien aprende; y se relaciona con el lenguaje, la comunicación dialogada, la escucha activa y el respeto. Desde un estado emocional de seguridad y confianza, el alumnado desarrolla estrategias cognitivas para avanzar hacia los objetivos, toma conciencia de cómo aprende y reajusta sus acciones. Los estados emocionales positivos facilitan el avance en el aprendizaje. Sirvan de referencia las palabras de Mora (2017, p. 71) “sin emoción no hay aprendizaje”. Diversos estudios han puesto de manifiesto que el feedback posee esta dimensión afectiva, que influye significativamente en variables como la motivación (Hattie y Timperley, 2007), la percepción de competencia y el bienestar emocional (Lipnevich y Smith, 2022). En la misma línea y según Nicaise et al. (2007), el feedback percibido como controlador puede generar respuestas emocionales negativas, mientras que el feedback de apoyo se asocia con experiencias de aprendizaje más adaptativas y autorreguladas. La eficacia de este feedback, al que llamamos formativo y afectivo, por hallarse relacionado con la evaluación formativa, no sólo depende de la calidad, precisión y grado de afectividad de la información proporcionada, sino de su percepción por las y los estudiantes. Esta percepción actúa como mediadora entre la intención docente y los efectos reales en el aprendizaje, la motivación y la implicación (Carless & Boud, 2018; Shute, 2008).

En la literatura se observan algunos modelos teóricos de feedback (e.g., Carless y Boud, 2018; Hattie y Timperley, 2007; Lipnevich y Panadero, 2021). Hattie y Timperley (2007) resaltan la importancia del feedback para avanzar desde el momento actual “cómo voy”, hasta el futuro “dónde quiero llegar”. Carless y Boud (2018) enfatizan la relación docente-alumnado y el papel del diálogo como aspectos claves del feedback. Lipnevich y Panadero (2021) refuerzan la idea sobre cómo el alumnado procesa el mensaje y autorregula su propio aprendizaje. De estos modelos y la sólida base teórica que los soporta, se desprende la determinación de tres dimensiones relacionadas con el feedback. La primera es la instruccional y hace referencia a la transmisión de información sobre el aprendizaje, sobre los aciertos y los puntos a mejorar. Más allá de correcciones técnicas, el feedback permite dar el salto entre lo que el alumnado hace por sí solo y lo que puede lograr con ayuda, dando así sentido a la zona de desarrollo próximo defendida por Vigotsky (1978), en su teoría del aprendizaje. La segunda dimensión es la interaccional; la afectividad actúa como una llave para que el alumnado integre el mensaje recibido sin defensas. Cuando percibe el apoyo afectivo del feedback, el alumnado baja la guardia y se abre a escuchar (Brookhart, 2008; Shute, 2008). A partir de esta base emocional se puede construir un vínculo pedagógico y un diálogo que dan credibilidad y sentido a la evaluación (Carless & Boud, 2018; Hattie & Timperley, 2007).

Cuando este clima de confianza y seguridad está presente, se puede avanzar hacia la autorregulación, tomando conciencia de cómo consigue sus resultados y ajustando y dirigiendo su propio aprendizaje (Nicol & Macfarlane-Dick, 2006; Panadero, 2017), encontrándonos entonces en la última de las dimensiones: la autorregulatoria. Así, el feedback no sólo informa sobre el desempeño, sino que, como sostienen Carless y Boud (2018), modela la identidad y la capacidad de juicio del alumnado, siendo la vertiente afectiva “la puerta de paso” hacia las tres dimensiones.



En la misma línea, Panadero y Alonso-Tapia (2013) afirman que el feedback no es sólo cognitivo, sino que supone una respuesta afectiva en el estudiante, y esto es clave para la autorregulación del aprendizaje, siendo esta parte afectiva del feedback la que permite —o bloquea— el procesamiento de la información recibida. En este sentido, Carless & Boud (2018) proponen el *feedback literacy* como la capacidad del alumnado de apreciar, interpretar y utilizar el feedback a partir de la gestión de emociones, la realización de juicios constructivos y la toma de acciones concretas. Esta alfabetización (*literacy*) se desarrolla a partir de un feedback dialogado entre docentes y alumnado; esta es la dimensión interaccional. Desde el *feedback literacy*, se puede avanzar hacia la tercera de las dimensiones: la autorregulatoria, tomando conciencia de cómo consigue sus resultados y ajustando y dirigiendo su propio aprendizaje (Nicol & Macfarlane-Dick, 2006; Panadero, 2017).

En la misma línea, Panadero y Alonso-Tapia (2013) afirman que el feedback no es sólo cognitivo, sino que supone una respuesta afectiva en el estudiante, y esto es clave para la autorregulación del aprendizaje.

En el contexto concreto de la Educación Física (EF), el feedback ha evolucionado desde el enfoque conductista de los años 70-80 con énfasis en la frecuencia y precisión de las correcciones del docente, hacia los enfoques cognitivos y socioconstructivistas de los años 90. Los trabajos clásicos de Siedentop et al. (1986) y Pieron (1992) a través de la observación y estudios cualitativos consideran al alumnado como parte activa del aprendizaje. A partir de aquí, y de forma progresiva, se avanza hacia modelos de evaluación formativa y compartida, donde autores como López-Pastor (2009) defienden un feedback orientado a la autonomía y la autorregulación del aprendizaje. Recientemente, Treschman et al. (2024) arrojan que muchos de los estudios focalizados en el feedback con relación a la EF utilizan como base teórica la evaluación para el aprendizaje y el modelo de aprendizaje autorregulado. Igualmente, señalan que tiene algo de presencia el modelo de feedback de Hattie y Timperley (2007), siendo este último el único específico detectado.

A diferencia de otras asignaturas, en EF el aprendizaje es motor y se manifiesta a través de la exposición pública de la persona y de su corporalidad ante docentes y estudiantes. En este contexto, la afectividad del feedback puede ser altamente relevante, ya que mucho más allá de una corrección técnica, el alumnado puede interpretar el mensaje como un juicio sobre su propia competencia y valía personal, acen tuando la vulnerabilidad propia de la etapa adolescente. Si bien el estudio de García-González et al. (2021), demuestra el impacto de las estrategias pedagógicas en EF sobre la motivación del alumnado, no profundizan en la naturaleza específica de la interacción afectiva del feedback. Asimismo, cabe destacar que Treschman et al. (2024) señalan que la mayoría de los estudios en EF se limitan a medir la frecuencia o el tipo de feedback (lo que el profesor hace), descuidando cómo el alumno percibe y procesa emocionalmente esa información. Prueba de ello es el reciente estudio de Bakali Hadyi et al. (2025) en el que concluyen que “los contextos con feedback abundante se asocian con mayores percepciones de progreso en el aprendizaje motor” (p.,109), sin precisar la receptividad emocional de esa retroalimentación.

En el contexto de la EF, un feedback percibido como excesivamente crítico o controlador puede generar una carga emocional negativa que bloquea los procesos de autoevaluación y reflexión (Cervelló & Santos-Rosa, 2000).

En otro orden de cosas, y en relación con la variable género, Nicaise et al. (2007) demuestran que los docentes de EF suelen ofrecer un feedback más técnico a los chicos, mientras que las alumnas reciben mensajes de carácter más organizativo. También, Castillo-Retamal y Sáez-Gallego (2014) señalan que las chicas presentan una mayor sensibilidad al componente afectivo del feedback, siendo este un factor determinante para su percepción de competencia y motivación intrínseca.

Por otra parte, en cuanto a los instrumentos que evalúan la percepción del feedback por parte del alumnado, se constata la existencia del Perceptions of Teacher Feedback Questionnaire (PTF) (Koka & Hein, 2003). Está centrado en tipologías de feedback, pero no incluye explícitamente la dimensión afectiva ni proporciona un patrón de referencia base. Koka y Hein (2005) diseñaron esta herramienta con intención de estudiar la motivación intrínseca en EF, bajo el marco de la Teoría de la Autodeterminación, y evidenciaron cómo el feedback positivo aumentaba la motivación intrínseca del alumnado en las clases de EF.

Si bien el PFT ha contribuido al estudio del feedback y ha abierto una interesante línea de investigación, no se aprecia relación con los niveles del feedback (tarea, proceso, autorregulación y persona), que proponen en su modelo Hattie y Timperley (2007), ni se hace referencia al componente cíclico de la autorregulación (ajuste de metas y estrategias), propuesto por Panadero y Lipnevich (2022).

A pesar de que la importancia del feedback en el aprendizaje motor se expresa de forma sólida en la literatura, se revela un vacío en relación con la escasez de instrumentos que analicen cómo el alumnado percibe el feedback que recibe en las clases de EF, en un contexto de evaluación formativa. No encontramos en la actualidad instrumentos que nos permitan comprobar cómo el alumnado está percibiendo el feedback. Diferentes estudios mencionan la necesidad del profesorado de contar con recursos para valorar, calibrar y reajustar su práctica docente (López-Pastor & Pérez-Pueyo, 2017). Sin duda hay una necesidad de investigación en acometer no sólo la conducta docente, sino también, y en su justa medida, la percepción de quienes reciben este feedback.

Consecuentemente, el objetivo principal de este trabajo es validar el cuestionario FFAPEF mediante análisis factorial exploratorio (AFE) y confirmatorio (AFC). Al respecto, y de acuerdo con el marco teórico expuesto, se plantea como hipótesis principal que el cuestionario presentará una estructura compuesta por tres dimensiones (instruccional, interaccional y autorregulatoria), reflejando así la complejidad del feedback en el contexto motriz de las clases de EF. Asimismo, se hipotetiza que dicho constructo mostrará diferencias significativas en función del género, el curso y el tipo de centro.

Como objetivo secundario se pretende conocer cómo el alumnado interpreta y experimenta el feedback recibido, planteando la hipótesis de que dicho constructo presentará diferencias en función del género, el curso y el tipo de centro en el que se encuentre. En definitiva, el estudio, en su conjunto, intenta arrojar luz sobre la complejidad de la recepción del feedback en contextos motrices y en etapas de gran vulnerabilidad; así como sentar bases para una futura herramienta de reflexión y autoevaluación sobre el propio diseño e impartición del feedback docente.

Método

Participantes

Se utilizó un muestreo no probabilístico de carácter incidental, compuesto por 259 estudiantes de 1.º a 4.º de ESO de dos centros del norte de España (Tabla 1). A pesar de la naturaleza de la muestra, la selección se realizó buscando una proximidad a los datos de género (53.6% chicos; 46.3% chicas) y tipo de centro (58.9% público; 41.3% concertado) de la población de Navarra (Departamento de Educación, 2024). Siguiendo las recomendaciones de la literatura para el desarrollo de escalas (Costello & Osborne, 2005; DeVellis, 2003; Hair et al., 2009; Stevens, 2002), se mantuvo una ratio de 20 participantes por ítem. Este tamaño muestral asegura la robustez y estabilidad de las soluciones factoriales en la representación de la población objetivo (Fabrigar et al., 1999).

Tabla 1. Descriptivo de la Muestra por Centro, Género y Nivel Educativo (n=259).

Variable	Categoría	% muestra	Nº de estudiantes	% chicos	% chicas
Tipo de centro	Público	58,9	152	50	50
	Concertado	41,3	107	59	41
Género	Masculino	53,6	139	53,6	—
	Femenino	46,3	120	—	46,3
Grupo	1.º ESO	33,2	86	63	37
	2.º ESO	22,0	57	47	53
	3.º ESO	21,6	56	59	41
	4.º ESO	23,1	69	42	58
Total	—	100	259	—	—

Procedimiento

El estudio obtuvo la aprobación del Comité de Ética del de la Universidad Pública de Navarra (PI-011/24), en corresponsabilidad con la Universidad de Zaragoza (RAT 2024-186).



En un primer momento se contactó con los equipos directivos y profesorado de cinco centros educativos de la zona para realizar entrevistas sobre la idoneidad del cuestionario para el alumnado de la ESO (12-16 años). Se contó con la participación de siete docentes y ocho estudiantes.

Tras los resultados de estas entrevistas se obtuvo el cuestionario definitivo, y se contactó con dos centros educativos para pasarlo a la muestra.

Los dos centros en los que se tomaron datos, enviaron consentimiento informado a progenitores/tutores, detallando objetivos, voluntariedad y derecho a retirada del estudio en cualquier momento, quedando un total de 53 consentimientos sin entregar y por tanto fuera de la investigación. El profesorado organizó tareas alternativas para quienes no accedieron a participar.

El cuestionario FFAPEF digitalizado, se administró por los tutores y/o profesores de EF durante horario lectivo (8-10 minutos), garantizando las condiciones de tranquilidad y anonimato. Justo antes de comenzar se proyectó un vídeo sobre el significado del feedback con un ejemplo real de las clases de EF, para facilitar la comprensión y asegurar que todo el alumnado recibiese una información inicial similar. Los cursos del mismo centro no tenían un único docente de EF. La recogida de datos transcurrió entre octubre y diciembre de 2025.

Instrumento

El cuestionario propuesto en este estudio (FFAPEF) evalúa la percepción del feedback formativo y afectivo mediante 10 ítems (escala Likert 1-5: Totalmente en desacuerdo a Totalmente de acuerdo), inicialmente distribuidos en tres dimensiones teóricas (ver Tabla 2): instruccional, interaccional y autorregulatoria.

El cuestionario se elaboró en cuatro fases. En la primera fase se realizó una búsqueda en la literatura para confeccionar una serie de ítems representativos del contenido del constructo.

En la segunda fase se validó el contenido de los mencionados ítems mediante el método Delphi con 14 jueces universitarios, expertos en EF (Red de Evaluación Formativa; 43% mujeres) procedentes de cuatro comunidades autónomas. La competencia experta (K) se garantizó obteniendo niveles altos (57%, $K > 0.80$) y medios (43%, K entre 0.50 y 0.80). Tras dos rondas de consulta, se aplicaron criterios de estabilidad y consenso: media 4.2, coeficiente de variación (CV) 30% y W de Kendall ≥ 0.50 ($p < .05$). Los resultados confirmaron un acuerdo moderado-alto en estos seis atributos finales: (1) "el feedback ha de adaptarse al alumnado"; (2) "aportar información sobre el aprendizaje expresando aciertos y errores"; (3) "fomentar el diálogo, la escucha activa, el apoyo mutuo, la cordialidad y el respeto"; (4) "integrarse en estrategias de coevaluación entre iguales"; (5) "incluirse en acciones de autoevaluación", y (6); "orientarse hacia la mejora personal, promoviendo la autoconfianza y el compromiso. No se requirió una tercera ronda, lo que dio lugar a un primer cuestionario de 12 ítems.

En la tercera fase el objetivo fue lograr la validez de contenido y la inteligibilidad de este cuestionario para ser pasado a alumnado de EF de secundaria. Para ello se realizaron entrevistas a siete docentes de ESO en activo ($n=7$), de los cuales cinco eran de EF y dos de Lengua Castellana; (57.1% hombres, edad media 54.2). En la entrevista, los docentes valoraron en una escala 1-4 en qué medida los ítems eran claros y comprensibles para el alumnado de 1º-4º ESO y se asociaban con sus experiencias de clase. Se trataba de adecuar lo más posible el cuestionario a la etapa de vida del alumnado (12-16 años) y a la realidad de las sesiones de EF. También se les preguntó en qué medida ellos como profesionales de la EF valoraban la pertinencia y relevancia de cada ítem con el atributo del feedback correspondiente. Se pasaron entrevistas también a ocho estudiantes de 1º a 4º ESO ($n=8$), dos por cada curso (50% chicas) realizando así un muestreo intencional estratificado por curso y género. En la cuarta fase se realizó la validación de constructo a través de un análisis factorial exploratorio (AFE) y confirmatorio (AFC).

Análisis de datos

Para asegurar la validez estructural, la muestra ($n = 259$) se dividió aleatoriamente en dos submuestras independientes. La primera se destinó al Análisis Factorial Exploratorio ($n = 183$), superando el criterio de 10-15 sujetos por ítem (Henson & Roberts, 2006). La segunda ($n = 76$) se empleó para el Análisis Factorial Confirmatorio.

Aunque este tamaño es ajustado, se considera admisible para modelos de baja complejidad y pocos parámetros (10 ítems), priorizando la independencia de las muestras para evitar la circularidad estadística (Lloret-Segura et al., 2014).

Dada la naturaleza ordinal de los ítems, se utilizaron matrices de correlación policórica y el método de mínimos cuadrados ponderados (WLSMV), el cual es robusto y eficiente en muestras pequeñas con variables asimétricas (Brown, 2015; Muthén & Muthén, 2017). La fiabilidad se analizó a partir del coeficiente ω de McDonald y el alfa de Cronbach (0.905). Debido a la naturaleza ordinal de los datos y la ausencia de normalidad multivariante, se emplearon las pruebas no paramétricas U de Mann Whitney para comparar las variables género y tipo de centro y Kruskal-Wallis (H) para la variable grupo, siguiendo las recomendaciones de Siegel y Castellan (1988).

Los datos fueron analizados mediante el programa estadístico Jasp (versión 0.95.4.0).

Resultados

Validez de contenido, comprensión e inteligibilidad

Los resultados cuantitativos de la entrevista a docentes (Tabla 2) mostraron un alto consenso y representatividad del contenido, con puntuaciones medias elevadas ($\bar{x} \geq 2.93$; mayoritariamente $\bar{x} > 3.50$), desviaciones estándar reducidas (0.13–0.42), y baja dispersión (todos los CV ≤ 0.12).

Tabla 2. Resultados cuantitativos de la entrevista a docentes (n=7), versión previa del cuestionario.

Ítem	Media	Mediana	DV	CV
1. Mi profesor/a me da indicaciones que se ajustan a mi nivel y a lo que necesito.	3.60	4	0.29	0.08
2. Mi profesor/a tiene en cuenta mis dificultades al darme feedback.	3.77	4	0.31	0.08
3. Mi profesor/a me explica con claridad qué hago bien y qué debo mejorar.	3.83	4	0.16	0.04
4. Mi profesor/a me explica y corrige con un lenguaje positivo y fácil de entender.	3.83	4	0.23	0.06
5. Mi profesor/a escucha mis dudas y me responde con respeto.	3.86	4	0.15	0.04
6. Mi profesor/a escucha mis opiniones, las valora y podemos dialogar sobre las actividades de clase.	3.52	4	0.42	0.12
7. Mi profesor/a organiza actividades para que nos demos feedback entre compañeros/as.	3.81	4	0.25	0.07
8. El feedback de mis compañeros/as me ayuda a mejorar.	3.71	4	0.24	0.06
9. Mi profesor/a propone actividades para que valore mi propio aprendizaje y reflexione sobre ello.	2.93	4	0.15	0.05
10. Cuando hacemos autoevaluación me doy cuenta por mí mismo/a de mis aciertos y de lo que debo mejorar.	3.81	4	0.15	0.04
11. El feedback que recibo de mi profesor/a me motiva a seguir esforzándome y practicando.	3.90	4	0.13	0.03
12. El feedback que recibo de mi profesor/a me da seguridad y confianza en mí mismo/a para intentarlo de nuevo.	3.90	4	0.13	0.03

Nota. DV = desviación estándar; CV = coeficiente de variación. Escala de respuestas: 1 = Muy en desacuerdo, 2 = En desacuerdo, 3 = Ni de acuerdo ni en desacuerdo, 4 = De acuerdo, 5 = Muy de acuerdo.

De acuerdo con la literatura sobre validez de contenido, la combinación de puntuaciones medias elevadas y baja dispersión es indicativa de un alto nivel de consenso entre expertos y respalda la adecuación de los ítems propuestos (DeVellis, 2017; Polit & Beck, 2006). En conjunto, estos resultados ponen de manifiesto un elevado grado de consenso y representatividad del contenido del cuestionario, por lo que no se consideró necesaria la eliminación de ítems por falta de relevancia o comprensión.

No obstante, el análisis cualitativo de las entrevistas a los docentes sugirió que algunos ítems eran similares semánticamente y siguiendo su criterio profesional se decidió su unificación. Los ítems 3 y 4 relacionados con la claridad y positividad del mensaje y ambos con $\bar{x} = 3.83$ se fusionaron ya que los expertos consideraron estas dimensiones como inseparables en la comunicación docente. De igual modo, los ítems 11 y 12 relacionados con la motivación y la seguridad y ambos con $\bar{x} = 3.90$ ya que los docentes consideraron que ambas respuestas emocionales podían evaluarse conjuntamente.

Estos ajustes dieron lugar a la versión definitiva de 10 ítems (Tabla 3).

Respecto a las entrevistas al alumnado (1.º-4.º de ESO), se registraron altos niveles de comprensión general. Sin embargo, se apreciaron dudas sobre el significado del término 'feedback' lo que motivó la creación del recurso audiovisual descrito en el apartado de Procedimiento.

Se aseguró así la comprensión del concepto por parte del alumnado y una mayor validez de las respuestas. Este procedimiento es coherente con las recomendaciones metodológicas que destacan la utilidad de las entrevistas y del análisis cualitativo para optimizar la claridad y adecuación contextual de los ítems (Boateng et al., 2018).

Tabla 3. Ítems del Cuestionario FFAPEF y relación con atributos y dimensiones, versión definitiva.

Ítem	Atributos	Dimensión predominante
1. Mi profesor/a me da indicaciones que puedo entender y que necesito para aprender y mejorar.	1. Adaptado al alumnado	Instruccional
2. Mi profesor/a tiene en cuenta mis características personales al darme feedback.	1. Adaptado al alumnado	Instruccional / Interaccional
3. Mi profesor/a me explica qué hago bien y qué debo mejorar con un lenguaje respetuoso y fácil de entender.	2. Información sobre aciertos y errores	Instruccional
4. Mi profesor/a escucha mis dudas y peticiones, y me responde con respeto.	3. Diálogo, escucha activa y respeto	Interaccional
5. Mi profesor/a escucha mis opiniones, las tiene en cuenta y hay diálogo.	3. Diálogo, escucha activa y respeto	Interaccional
6. Mi profesor/a organiza actividades para que nos demos feedback entre nosotros/as, para que nos ayudemos dando nuestra opinión al compañero/a.	4. Estrategias de coevaluación	Interaccional / Autorregulatoria
7. Los comentarios y la información que me dan mis compañeros/as me ayudan a mejorar.	4. Estrategias de coevaluación	Autorregulatoria
8. Mi profesor/a propone actividades para que valore mi propio aprendizaje y reflexione sobre cómo lo hago.	5. Acciones de autoevaluación	Autorregulatoria
9. Cuando hago autoevaluación me doy cuenta de mis aciertos y de lo que debo mejorar.	5. Acciones de autoevaluación	Autorregulatoria
10. El feedback que recibo de mi profesor/a me da seguridad y confianza, motivándome a esforzarme y seguir practicando.	6. Mejora personal y compromiso	Interaccional / Autorregulatoria

Validez de constructo y fiabilidad

El análisis factorial exploratorio (AFE) (70 % de la muestra seleccionada de manera aleatoria) con correlaciones policóricas y extracción por mínimos cuadrados ponderados reveló una estructura unifactorial que explicó el 61% de la varianza total, con cargas factoriales entre altas y excelente (.65-.92), todas muy por encima del criterio mínimo (.40) (Tabla 4). Esto indica que todos los ítems contribuyen a un mismo factor y no se identificó ningún ítem problemático. La adecuación de la muestra se evaluó mediante el índice Kaiser-Meyer-Olkin (KMO = .927) y la prueba de esfericidad de Bartlett ($\chi^2(45) = 1392.002, p < .001$), confirmando la idoneidad de los datos para la reducción factorial. Todas las cargas son muy altas (entre .647 y .916) y se sitúan claramente por encima del criterio mínimo recomendado (.40). Esto indica que todos los ítems contribuyen de forma sólida al mismo factor y no se identifica ningún ítem débil ni problemático.

Tabla 4. Factores de la carga (AFE).

	Factor 1	Uniqueness
10.- Seguridad, confianza y motivación	0.916	0.162
4.- Escucha del docente, respuesta con respeto.	0.836	0.301
8.- Actividades de valoración del propio aprendizaje	0.831	0.310
5.- Escucha del docente y diálogo sobre actividades	0.826	0.318
3.- Lenguaje respetuoso y fácil de entender".	0.816	0.334
1.- Recibo indicaciones	0.796	0.366
6.- Actividades feedback entre iguales	0.726	0.473
2. Características personales	0.715	0.489
9.- Toma de conciencia de aciertos y errores a través de la autoevaluación	0.657	0.568
7.- Ayuda del feedback de mis compañeros/as	0.647	0.581

Nota: el método de rotación aplicado es oblimin.

El análisis factorial confirmatorio (AFC) se realizó mediante el estimador WLSMV (Weight Least Squares Mean and Variance adjusted), dada la naturaleza ordinal de los ítems (escala Likert de 5 puntos) y la ausencia de normalidad multivariante. Este método es el más robusto para variables categóricas ya que emplea matrices de correlaciones policóricas, evitando los sesgos que produciría el uso de estimadores basados en la normalidad (Finney & DiStefano, 2013; Flora & Curran, 2004). A partir de aquí, el AFC corroboró esta estructura: ($\chi^2(35) = 85.32, \chi^2/df = 2.44$), con ajustes incrementales satisfactorios (todos

superiores a .90): CFI = .982; TLI = .977; IFI = .953), cargas factoriales significativas (.677-.884, $p < .001$), y validez convergente adecuada (AVE = .559).

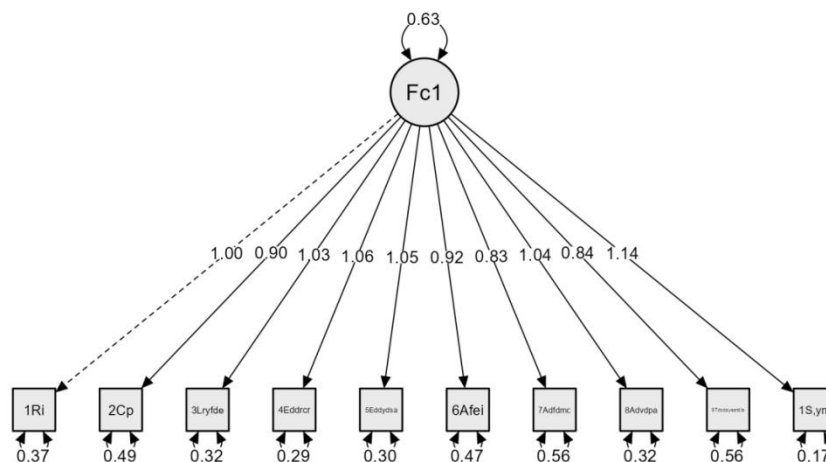
Aunque el valor del RMSEA fue .097, situándose por encima del umbral de .08, este índice tiende a sobrestimar el error en modelos con pocos grados de libertad (Kenny, Kaniskan & McCoach, 2015). Por ello, el reporte se completó con el índice SRMR, que obtuvo un valor de .048, indicando un ajuste óptimo ($< .08$). La interpretación conjunta de estos indicadores, sumada a la fuerza y significación de las cargas factoriales (.677-.884), confirma la validez y robustez de la estructura factorial propuesta (Hu & Bentler, 1999).

En relación con las cargas factoriales no estandarizadas que se muestran en el diagrama de senderos del modelo (Figura 1), se aprecian valores entre 0.83 y 1.14, lo que indica una relación sólida y positiva entre el constructo latente (feedback formativo afectivo) y sus indicadores observados (10 ítems de cuestionario).

La varianza del factor se estimó en .63, confirmando una dispersión adecuada del rasgo medido en la muestra. En cuanto a las varianzas residuales (errores), estas se mantuvieron en niveles aceptables. El ítem 1 (recibo indicaciones), destacó como el indicador más preciso con el error más bajo del modelo (0.17). Por el contrario, los ítems 7 (feedback de los compañeros/as) y 9 (autoevaluación) presentaron los niveles más altos de ruido o varianza no explicada (0.56), aunque siguen aportando significatividad a la definición del factor global. Asimismo, la escala muestra excelente fiabilidad interna, evidenciada tanto por el coeficiente ω de McDonald (0.922) como por el alfa de Cronbach (0.905), respaldando su uso en alumnado de ESO similar.

En conjunto, los resultados respaldan la validez factorial y la consistencia interna del instrumento para su uso en población similar a la estudiada.

Figura 1. Diagrama de senderos (AFC).



Análisis descriptivo

A partir de los descriptivos estadísticos de la puntuación total del cuestionario, los resultados mostraron que el alumnado tiene una percepción del feedback formativo y afectivo altamente positiva, con medias situadas en el rango superior de la escala (3.5 a 4.2). La mínima distancia observada entre la media y el primer cuartil (< 1 punto) reflejó que la gran mayoría de la muestra ($> 75\%$) se situó en los niveles más altos del constructo (Tabla 5).

Según Howell (2016), esta proximidad en el extremo superior de la distribución refleja un fuerte consenso grupal y una baja variabilidad en la zona de insatisfacción, confirmando que los 259 estudiantes percibieron en gran medida y de forma positiva este feedback formativo y afectivo.

Tabla 5. Estadísticos Descriptivos puntuaciones totales de cada ítem del cuestionario (1 a 10)

	1	2	3	4	5	6	7	8	9	10
Válido	259	259	259	259	259	259	259	259	259	259
Ausente	0	0	0	0	0	0	0	0	0	0
Moda	5.000	4.000	4.997	5.000	4.003	4.000	4.000	4.001	4.000	4.989
Media	4.127	3.761	4.085	4.212	3.826	3.548	3.591	3.761	4.108	3.892
Desviación típica	0.958	1.040	1.027	0.955	1.112	1.152	1.122	1.084	0.934	1.112
Coefficiente de variación	0.232	0.277	0.251	0.227	0.291	0.325	0.312	0.288	0.227	0.286
Primer cuartil	4.000	3.000	3.500	4.000	3.000	3.000	3.000	3.000	4.000	3.000
Segundo cuartil (mediana)	4.000	4.000	4.000	4.000	4.000	4.000	4.000	4.000	4.000	4.000
Tercer cuartil	5.000	4.500	5.000	5.000	5.000	4.000	4.000	5.000	5.000	5.000
Suma	1069	974.0	1058	1091	991.0	919.0	930.0	974.0	1064	1008

A continuación, se optó por analizar la relación entre las variables género, grupo y tipo de centro con la puntuación total del cuestionario (máximo 50 puntos), y con cada uno de los 10 ítems, para valorar la posible existencia de relaciones significativas. Se aplicó la U de Mann-Whitney para comparar la puntuación total y de cada ítem con el género y con el tipo de centro.

En relación al género, no se apreciaron diferencias significativas con la puntuación total ($p = .691$) ni con cada uno de los ítems ($p > .05$ en todos ellos). En relación al tipo de centro tampoco se detectaron diferencias significativas con la puntuación total ($p = .345$) ni con cada uno de los ítems ($p > .05$ en todos ellos). Por otra parte, en relación con el grupo, y tras la aplicación de Kruskal-Wallis, se apreció la existencia de relaciones altamente significativas en la puntuación total ($p < .001$) entre 3º, 1º y 2º ESO. Las comparaciones post-hoc de Dunn indicaron que el grupo de 3º ESO obtuvo puntuaciones totales significativamente más altas que 1º y que 2º ($p < .05$).

Al analizar la relación entre el curso y cada uno de los ítems, los números 1,3,4,7 y 10 mostraron diferencias significativas ($p < .05$). Las comparaciones post-hoc de Dunn indicaron que el grupo de 3º ESO puntuó significativamente más alto que 1º en los ítems 1,3,4,5,6 y 9 (p Holm $< .05$) y significativamente más alto que 2º ESO en los ítems 3,7,8 y 10 (p Holm $< .05$) (Tabla 7).

Se concluye que el cuestionario funciona de forma estable independientemente del género del estudiante y del tipo de centro. No obstante existe un efecto significativo del curso académico, con el 90% de los ítems mostrando diferencias significativas ($p < .05$) y tamaños del efecto entre pequeños y medianos (0.01-0.08). Específicamente las puntuaciones en 3º ESO cuando son significativamente más altas en comparación con los cursos inferiores (1º y 2º), y en casi todos los ítems.

Tabla 6. Análisis comparativo entre el curso y los ítems del Cuestionario FFAPEF (n=259)

Ítem	H	gl	p	ω^2	Z	PHolm	Cursos
1. Dar indicaciones que el alumnado puede entender y que necesita para aprender y mejorar	17.31	3	<.001	0.06	3.47	.003	1º-3º
2. Tener en cuenta las características personales del alumnado	5.42	3	.102	0.01	1.87	.356	2º-3º
3. Expresarse en un lenguaje respetuoso y fácil de entender para explicar al alumnado lo que hace bien y lo que debe mejorar	14.25	3	.003	0.03	3.47	.003	2º-3º
4. Fomentar el diálogo y el respeto	21.34	3	<.001	0.08	3.87	<.001	1º-3º
4. Fomentar el diálogo y el respeto	21.34	3	<.001	0.08	4.18	<.001	2º-3º
5. Escuchar, tener en cuenta las opiniones del alumnado y dialogar sobre las actividades de clase	9.62	3	.022	0.04	3.03	.015	1º-3º
6. Organizar actividades para que el alumnado se dé feedback entre iguales y se ayuden dándose opiniones	9.08	3	.028	0.02	2.64	.049	1º-3º
7. Los comentarios e información que se dan los alumnos/as entre sí les ayuda a mejorar	15.30	3	.002	0.05	3.79	<.001	1º-3º
8. Proponer actividades para que cada alumno/a valore su aprendizaje y reflexione sobre cómo lo hace	12.31	3	.006	0.03	2.88	.024	2º-3º
9. Plantear actividades de autoevaluación para que el alumnado se dé cuenta de sus aciertos y de sus puntos a mejorar	12.86	3	.005	0.03	3.16	.010	1º-3º
10. Aportar seguridad, confianza y motivación para esforzarse y seguir practicando	16.76	3	<.001	0.05	3.77	<.001	2º-3º

Nota: H = estadístico de Kruskal-Wallis; gl = grados de libertad; p Holm = valor p corregido por Holm; ω^2 = tamaño del efecto.

Discusión



Los resultados evidencian adecuados niveles de validez de contenido, de constructo y fiabilidad interna, respaldando el uso del cuestionario en investigación y práctica docente. La validez de contenido se apoya en un proceso riguroso (estudio Delphi, entrevistas y revisiones) alineado con las necesidades de validación cognitiva y contextual (Boateng et al., 2018; DeVellis, 2017). Las valoraciones elevadas en claridad y pertinencia confirman la adecuación del instrumento al alumnado de 1º a 4º de ESO.

Desde el punto de vista psicométrico, los análisis factoriales (AFE y AFC) sostienen una estructura unidimensional sólida. Aunque el feedback se fundamentó teóricamente en tres dimensiones (instruccional, interaccional y autorregulatoria), los resultados obtenidos sugieren que la estructura interna del constructo opera de forma unidimensional en la práctica educativa, lo que puede obedecer a factores de tipo cognitivo y contextual.

En primer lugar, es probable que el alumnado adolescente perciba el feedback como una experiencia holística e indiferenciada que explicaría la estructura unidimensional obtenida. Los estudiantes no diferencian el qué me dicen vs. cómo me lo dicen, sino que lo interpretan de forma global, y posiblemente emocional. Así, las dimensiones instruccional y autorregulatoria quedan supeditadas al 'paraguas' de la dimensión interaccional, fundiéndose en una única percepción; lo cognitivo y lo afectivo son inseparables en la práctica, y en esta etapa de vulnerabilidad. Estos resultados se alinean con el marco del feedback literacy (Carless y Boud, 2018), que propone una serie de capacidades integradas en el feedback que alumnado y docentes pueden desarrollar. Estas capacidades son la gestión emocional, la emisión de juicios y la propia apreciación del feedback. Desde esta perspectiva las tres dimensiones inicialmente planteadas (instruccional, interaccional y autorregulatoria), se fusionan en una única dimensión, y se interrelacionan, siendo la vertiente afectiva del mensaje, la clave para entender la estructura unidimensional del instrumento en este contexto específico.

En segundo lugar, el contexto de la clase de EF puede ejercer un efecto de homogeneización en las respuestas. A diferencia de otras asignaturas más teóricas, la EF se desarrolla en un entorno de alta interacción social y motriz, donde el feedback suele ser inmediato, público y vinculado a la acción motriz. En este escenario, la instrucción técnica y el mensaje afectivo suelen ocurrir simultáneamente (e.g un gesto de aprobación mientras se corrige una postura), lo que refuerza la percepción de una experiencia de enseñanza-aprendizaje integrada e indivisible.

Se podría concluir que las dimensiones planteadas confluyen en una única. En este sentido, Boekaerts (2011), propone que la activación emocional precede y envuelve los procesos de autorregulación del aprendizaje.

Más allá de sus propiedades psicométricas, una de las fortalezas más destacadas del cuestionario es su viabilidad a la hora de ponerlo en práctica. Al estar compuesto por únicamente 10 ítems, puede ser un recurso de aplicación directa a las clases. La brevedad del cuestionario permite su implementación como un instrumento de *screening* rápido sin interferir en la dinámica natural de las sesiones ni reducir el tiempo de compromiso motor, tan necesario para el alumnado. Al requerir apenas unos minutos para su cumplimentación, el cuestionario también permite al docente obtener un feedback del feedback (Hattie y Timperley, 2007) y ajustar sus estrategias, convirtiéndose en un recurso práctico para mejorar la interacción docente-alumnado.

Respecto a las variables género, tipo de centro y curso, la ausencia de diferencias significativas en género y tipo de centro aporta robustez al instrumento, facilitando su aplicación sin baremaciones diferenciadas. Por el contrario, la existencia de diferencias significativas por curso sugiere que el instrumento es sensible a los cambios madurativos o de aprendizaje propios de la etapa (12-16 años). El curso de 3º ESO (14-15 años) muestra puntuaciones significativamente más altas respecto a 1º y 2º, lo que puede asociarse con el momento de mayor madurez y cambio en la etapa de la adolescencia. La edad parece ser determinante en la percepción de este tipo de feedback lo que parece describir una etapa especialmente sensible (14-15 años) en el alumnado, para identificarlo y poder así después autorregular su aprendizaje.

Por otra parte, el cuestionario FFAPEF puede también plantearse como herramienta de autopercepción docente para reflexionar sobre la propia práctica. Esta doble funcionalidad permite al profesorado tomar conciencia no sólo de qué se dice, sino de cómo se dice (tono e intencionalidad), aspectos clave para la educación emocional.

Como líneas futuras, se propone el diseño de acciones formativas para el profesorado y la elaboración de recursos pedagógicos para un feedback más consciente y motivador. A partir de los ítems del cuestionario validado se podría generar un instrumento de autochequeo y guía para los docentes en relación al feedback.

Por otro lado, la irrupción actual de la Inteligencia Artificial (IA) generativa plantea una transformación profunda del feedback educativo, permitiendo una inmediatez y personalización sin precedentes (Meyer et al., 2026).

No obstante, el impacto de estas herramientas depende de su integración con principios pedagógicos que prioricen la dimensión humana. Un estudio reciente arroja que el feedback ofrecido por docentes consigue mayores niveles de motivación intrínseca en practicantes de EF que el feedback automatizado. Además, se posiciona como promotor de relación social, y como de generador de compromiso proactivo; sugiriendo esto que la interacción humana resulta clave en la dimensión socio-colaborativa (Prieto Andreu & Lopes, 2025).

En este escenario, el cuestionario FFAPEF podría ser un instrumento para futuras líneas de investigación en torno a si el feedback automatizado, es capaz de emular la carga afectiva necesaria en EF, o si por el contrario, genera desconexión emocional en los adolescentes. La investigación futura podría combinar el uso del FFAPEF con sistemas para monitorizar en tiempo real si el feedback de la IA está logrando el impacto emocional positivo necesario para el aprendizaje formativo. Se abre así un camino hacia modelos híbridos de feedback, donde la precisión de la IA se combine con la empatía docente, preservando siempre la conexión emocional como eje del aprendizaje motriz.

Conclusiones

El estudio aporta un instrumento válido y fiable fundamentado en un marco conceptual sólido. Los resultados ponen de manifiesto que el alumnado percibe el feedback formativo y afectivo de forma integral y que en los 10 ítems del cuestionario se entrelazan la dimensión más instruccional con la interaccional o afectiva y la autorregulatoria. Se puede decir que la unión del “qué decir”, el “cómo decirlo” y el “para qué” confiere solidez y eficacia a un feedback que busca mejorar el aprendizaje en un ambiente de seguridad, afecto y motivación.

La validación del cuestionario FFAPEF ofrece a los profesionales de la EF una herramienta operativa que permite sin duda indagar sobre el clima del aula desde una perspectiva holística.

El alumnado percibe el feedback como una experiencia unidimensional, y sin duda su aplicación puede plantear una reflexión en el profesorado sobre la unidad entre la dimensión instruccional y la interaccional, dado que parece que ambos procesos ocurren simultáneamente en la percepción del adolescente.

Aunque el instrumento ha sido validado en el ámbito de la EF, podría ser transferible a otras materias e incluso a la iniciación deportiva, ofreciendo a entrenadores un recurso para reflexionar sobre la importancia de este feedback formativo y afectivo en la comunicación con sus deportistas y en la promoción de su autorregulación.

En cuanto a las limitaciones, el uso de un muestreo incidental en una única comunidad autónoma restringe la generalización de los resultados, a pesar de su proximidad a los datos poblacionales. Asimismo, el diseño transversal impide analizar la estabilidad temporal del instrumento. Estas restricciones orientan futuras investigaciones hacia muestras más amplias y diversificadas que permitan confirmar la estabilidad de la estructura factorial del FFAPEF en distintos contextos y asegurar la generalización de los hallazgos.

Además, se sugieren diseños longitudinales para evaluar la sensibilidad del cuestionario en diferentes intervenciones didácticas en clase, y su capacidad para detectar diferencias por género o curso con mayor potencia estadística.

Agradecimientos



Las autoras agradecen la colaboración de los centros educativos, el alumnado y las familias que participaron en este estudio, cuya disposición fue fundamental para la realización de la investigación.

Financiación

El estudio ha estado financiado por el Grupo de investigación de la Universidad de Zaragoza EFYPAF (Educación Física y Promoción de Actividad Física).

Referencias

- Bakali Hadyi, N., Moumen, A. & Harfouf, F. (2025). Estudio sobre los factores clave que influyen en el aprendizaje motor entre futuros docentes de Educación Física en formación universitaria. *Retos*, 73, 1090-1100. <https://doi.org/10.47197/retos.v73.117152>
- Black, P. (2007). Full marks for feedback. *Make the Grade: Journal of the Institute of Educational Assessors*, 2(1), 18-21.
- Boateng, G. O., Neilands, T. B., Frongillo, E. A., Melgar-Quinonez, H. R., & Young, S. L. (2018). Best practices for developing and validating scales for health, social, and behavioral research: A primer. *Frontiers in Public Health*, 6 (149), 1-18. <https://doi.org/10.3389/fpubh.2018.00149>
- Boekaerts, M. (2011). Emotions, emotion regulation, and self-regulation of learning. In B. J. Zimmerman & D. H. Schunk (Eds.), *Handbook of self-regulation of learning and performance* (pp. 408-425). Routledge.
- Brookhart, S. M. (2008). *How to give effective feedback to your students*. ASCD.
- Boud, D., & Molloy, E. (2013). Rethinking models of feedback for learning: The challenge of design. *Assessment & Evaluation in Higher Education*, 38(6), 698-712. <https://doi.org/10.1080/02602938.2012.691462>
- Brown, T. A. (2015). *Confirmatory factor analysis for applied research* (2nd ed.). Guilford Press.
- Carless, D., & Boud, D. (2018). The development of student feedback literacy: Enabling uptake of feedback. *Assessment & Evaluation in Higher Education*, 43(8), 1315-1325. <https://doi.org/10.1080/02602938.2018.1463354>
- Castillo-Retamal, F., & Cordero-Tapia, M. (2014). Género y educación física: Una mirada a las prácticas pedagógicas. *Revista Ciencias de la Actividad Física UCM*, 15(2), 51-60. <https://revista-caf.ucm.cl/article/view/51>
- Costello, A. B., & Osborne, J. W. (2005). Best practices in exploratory factor analysis: Four recommendations for getting the most from your analysis. *Practical Assessment, Research & Evaluation*, 10(7). <https://doi.org/10.7275/jyj1-4868>
- DeVellis, R. F. (2017). *Scale development: Theory and applications* (4th ed.). Sage.
- Fabrigar, L. R., Wegener, D. T., MacCallum, R. C., & Strahan, E. J. (1999). Evaluating the use of exploratory factor analysis in psychological research. *Psychological Methods*, 4(3), 272-299. <https://doi.org/10.1037/1082-989X.4.3.272>
- Finney, S. J., DiStefano, C., & Kopp, J. P. (2016). Overview of estimation methods and preconditions for their application with structural equation modeling. In *Principles and methods of test construction: Standards and recent advances* (pp. 135-165). <https://doi.org/10.1027/00449-000>
- Flora, D. B., & Curran, P. J. (2004). An empirical evaluation of alternative methods of estimation for confirmatory factor analysis with ordinal data. *Psychological Methods*, 9(4), 466-491. <https://doi.org/10.1037/1082-989X.9.4.466>
- Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2010). *Multivariate data analysis* (7th ed.). Pearson.
- Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, 77(1), 81-112. <https://doi.org/10.3102/003465430298487>
- Henson, R. K., & Roberts, J. K. (2006). Use of exploratory factor analysis in published research: Common errors and some comment on improved practice. *Educational and Psychological Measurement*, 66(3), 393-416. <https://doi.org/10.1177/0013164405282485>
- Howell, D. C. (2016). *Fundamental statistics for the behavioral sciences* (9th ed.). Cengage Learning.



- Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, 6(1), 1–55. <https://doi.org/10.1080/10705519909540118>
- Kenny, D. A., Kaniskan, B., & McCoach, D. B. (2015). The performance of RMSEA in models with small degrees of freedom. *Sociological Methods & Research*, 44(3), 486–507. <https://doi.org/10.1177/0049124114543236>
- Koka, A., & Hein, V. (2003). Perceptions of teacher's feedback and learning environment as predictors of intrinsic motivation in physical education. *Psychology of Sport and Exercise*, 4(4), 333–346. [https://doi.org/10.1016/S1469-0292\(02\)00012-2](https://doi.org/10.1016/S1469-0292(02)00012-2)
- Koka, A., & Hein, V. (2005). The effect of perceived teacher feedback on intrinsic motivation in physical education. *International Journal of Sport Psychology*, 36(2), 91–106.
- Lipnevich, A. A., & Panadero, E. (2021). A review of feedback models and theories: Descriptions, definitions, and conclusions. *Frontiers in Education*, 6, 720195. <https://doi.org/10.3389/educ.2021.720195>
- Lipnevich, A. A., & Smith, J. K. (2022). Student–feedback interaction model: Revised. *Studies in Educational Evaluation*, 75, 101208. <https://doi.org/10.1016/j.stueduc.2022.101208>
- López-Pastor, V. M. (2009). *Evaluación formativa y compartida en Educación Superior: Propuestas, técnicas, instrumentos y experiencias* (Vol. 21). Narcea Ediciones.
- López Pastor, V. M., & Pérez Pueyo, Á. (2017). *Evaluación formativa y compartida en educación: experiencias de éxito en todas las etapas educativas*. Universidad de León.
- Lloret-Segura, S., Ferreres-Traver, A., Hernández-Baeza, A., & Tomás-Marco, I. (2014). El análisis factorial exploratorio de los ítems: Una guía práctica, revisada y actualizada. *Anales de Psicología*, 30(3), 1151–1169. <https://doi.org/10.6018/analesps.30.3.199361>
- Meyer, J., Köller, O., Jansen, T., Fleckenstein, J., Asher, M. W., Bichler, S., ... & Winstone, N. (2026). The Future of Feedback: How Can AI Help Transform Feedback to Be More Engaging, Effective, and Scalable? *arXiv preprint arXiv:2603.12463*. <https://doi.org/10.48550/arXiv.2603.12463>
- Mora, F. (2017). *Neuroeducación: Sólo se puede aprender aquello que se ama* (2ª ed.). Alianza Editorial.
- Muthén, L. K., & Muthén, B. O. (2017). *Mplus User's Guide* (8ª ed.). Muthén & Muthén.
- Nicaise, V., Bois, J. E., Fairclough, S. J., Amorose, A. J., & Cogérino, G. (2007). Girls' and boys' perceptions of physical education teachers' feedback: Effects on performance and psychological responses. *Journal of Sports Sciences*, 25(8), 915–926. <https://doi.org/10.1080/02640410600898095>
- Nicol, D. J., & Macfarlane-Dick, D. (2006). Formative assessment and self-regulated learning: A model and seven principles of good feedback practice. *Studies in Higher Education*, 31(2), 199–218. <https://doi.org/10.1080/03075070600572090>
- Panadero, E., & Alonso-Tapia, J. (2013). Autoevaluación: Connotaciones teóricas y prácticas. *Electronic Journal of Research in Educational Psychology*, 11(30), 551–576. <https://doi.org/10.14204/ejrep.30.12202>
- Panadero, E. (2017). A review of self-regulated learning: Six models and four directions for research. *Frontiers in Psychology*, 8, 422. <https://doi.org/10.3389/fpsyg.2017.00422>
- Pieron, M. (1992). La investigación en la enseñanza de las actividades físicas y deportivas. *Apunts. Educación Física y Deportes*, 4(30), 6–19. <https://raco.cat/index.php/ApuntsEFD/article/view/381452>
- Polit, D. F., & Beck, C. T. (2006). The content validity index: Are you sure you know what's being reported? *Research in Nursing & Health*, 29(5), 489–497. <https://doi.org/10.1002/nur.20147>
- Prieto Andreu, J. M., & Lopes, A. (2025). Artificial intelligence as a virtual coach: comparative effectiveness of automated feedback versus traditional methods in Physical Education. *Retos*, 72, 823-834. <https://doi.org/10.47197/retos.v72.117154>
- Siedentop, D., Mand, C. L., & Taggart, A. (1986). *Physical education: Teaching and curriculum strategies for grades 5–12*. Mayfield Pub. Co.
- Siegel, S., & Castellan, N. J. (1988). *Nonparametric statistics for the behavioral sciences* (2nd ed.). McGraw-Hill.
- Shute, V. J. (2008). Focus on formative feedback. *Review of Educational Research*, 78(1), 153–189. <https://doi.org/10.3102/0034654307313795>
- Stevens, J. P. (2002). *Applied multivariate statistics for the social sciences* (4th ed.). Lawrence Erlbaum Associates.



- Treschman, P., Stylianou, M., & Brooks, C. (2024). A scoping review of feedback in physical education: Conceptualisations and the role of teachers and students. *European Physical Education Review*, 30(4), 545–562. <https://doi.org/10.1177/1356336X241230829>
- Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes* (M. Cole, V. John-Steiner, S. Scribner, & E. Souberman, Eds.). Harvard University Press.
- Webb, M., & Galamba, A. (2026). Transforming pedagogy with GenAI-supported formative assessment: Challenges for teacher education. *British Journal of Educational Technology*. <https://doi.org/10.1111/bjet.70051>
- Winstone, N. E., Balloo, K., & Carless, D. (2022). Discipline-specific feedback literacies: A framework for curriculum design. *Higher Education*, 83(1), 57–77. <https://doi.org/10.1007/s10734-020-00632-0>
- Xiang, X., Yum, S., & Lian, R. (2023). Effect of cognitive attitude and affective attitude on formative assessment practice: Intention as a mediator. *Social Behavior and Personality: An International Journal*, 51(9), 1–16. <https://doi.org/10.2224/sbp.12484>

Datos de las autoras y traductora:

Ana Casares Polo	anacasarespolo71@gmail.com	Autora
Alicia M ^a Alonso Martínez	alonso@unizar.es	Autora
M ^a Rosario Romero Martín	rromero@unizar.es	Autora
Nerea Estrada Marcén	nereaes@unizar.es	Autora
Sonia Asún Dieste	sonasun@unizar.es	Autora
Loreto Alonso Martínez	loretoalonsom@gmail.com	Traductora