

MENTOR

Revista de Investigación Educativa y Deportiva

Volume 5

Special
Issue 4

2026

Director: Ph.D. Richar Posso Pacheco

Email: rjposso@revistamentor.ec

Website: <https://revistamentor.ec/>

Editor-in-Chief: Ph.D. Susana Paz Viteri

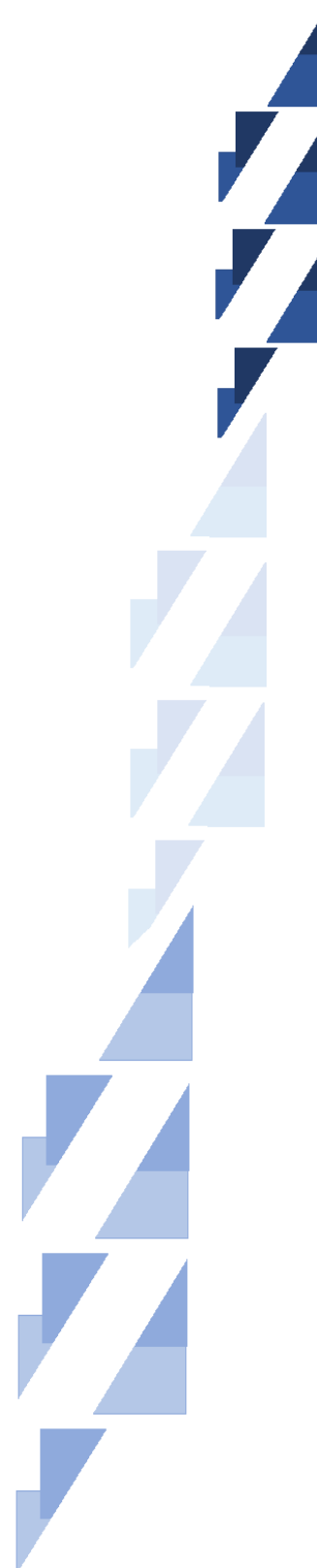
Editorial Coordinator: Ph.D. (c) Josue Marcillo Ñacato

Scientific Committee Coordinator: Ph.D. Laura Barba Miranda

Editorial Supervisor: Ph.D. Isidro Lapuente Álvarez

Editors' Committee Coordinator: Msc. María Gladys Cóndor Chicaiza

Reviewers' Board Coordinator: PhD. Javier Fernández-Rio



Articles

Scale of perceived teaching competence in Physical Education: structural validation and factorial invariance

Escala de competencia docente percibida en Educación Física: validez estructural e invarianza factorial

Perla Jannet Jurado-García ¹

ORCID: <https://orcid.org/0000-0002-0917-0028>

Susana Ivonne Aguirre Vásquez ¹

ORCID: <https://orcid.org/0000-0001-8878-2120>

Víctor Hugo López Trujillo ²

ORCID: <https://orcid.org/0000-0002-3644-1823>

Ramón Alfonso González-Rivas ¹

ORCID: <https://orcid.org/0000-0002-4919-6556>

Universidad Autónoma de Chihuahua, Facultad de Ciencias de la Cultura
Física. Chihuahua, Chihuahua México ¹

Secretaría de Educación Pública, Educación Física. Chihuahua,
Chihuahua México ²

Autor de correspondencia

rgrivas@uach.mx

Received: 13-03-2026

Accepted: 17-06-2026

Available online: 15-07-2026

Abstract

This study aimed to develop and examine the structural validity and measurement invariance of a scale designed to assess perceived teaching competence in Physical Education among Mexican secondary school students. The instrument was developed through the adaptation and contextual expansion of the Evaluation of Teaching Competencies Scale (ETCS) (Catano & Harvey, 2011). Two independent studies were conducted. Study 1 included 320 students aged 12–15 years ($M = 13.22$, $SD = 0.969$), whereas Study 2 involved 598 students aged 12–15 years ($M = 13.27$, $SD = 0.915$). An exploratory factor analysis using Maximum Likelihood extraction and Varimax rotation was performed, followed by a confirmatory factor analysis. The analyses resulted in a 10-item unidimensional model with acceptable fit indices ($\chi^2/df = 2.510$, $CFI = .972$, $RMSEA = .069$) and strong internal consistency ($\alpha = .92$, $\omega = .92$). The unidimensional structure was confirmed in Study 2 ($\chi^2/df = 4.590$, $CFI = .964$, $RMSEA = .078$), maintaining adequate reliability ($\alpha = .92$, $\omega = .92$). Measurement invariance across sex and grade level was supported. The findings indicate that the scale provides a psychometrically robust measure of perceived teaching competence in Physical Education within the Mexican educational context.

Keywords: teaching competence, Physical Education, psychometrics, measurement invariance, adolescents.

Resumen

El objetivo de este estudio fue desarrollar y aportar evidencias de validez estructural e invarianza factorial de una escala para evaluar la competencia docente percibida del profesorado de Educación Física en adolescentes mexicanos de educación secundaria; a partir de la adaptación y ampliación contextual de la *Evaluation of Teaching Competences Scale* (ETCS) (Catano y Harvey, 2011). El proceso fue en dos estudios con muestras independientes. Estudio uno, $N=320$; entre 12 y 15 años $M=13.22$; $DE=.969$; 179 mujeres y 141 hombres, estudio 2, $N=598$; entre 12 y 15 años $M=13.27$; $DE=.915$; 328 mujeres y 270 hombres. Se realizó el análisis factorial exploratorio con extracción por Máxima Verosimilitud y rotación Varimax y un análisis factorial confirmatorio. El proceso condujo a un modelo unifactorial de 10 ítems, con aceptables índices de ajuste ($\chi^2/gl=2.510$; $CFI=.972$; $RMSEA=.069$) y consistencia interna ($\alpha=.92$ y $\Omega=.92$). En la muestra 2 se confirmó la estructura unifactorial ($\chi^2/gl=4.590$; $CFI=.964$; $RMSEA=.078$), manteniendo la consistencia interna ($\alpha=.92$ y $\Omega=.92$). Se demostró la invarianza por sexo y grado escolar. La escala constituye un instrumento psicométricamente adecuado para la evaluación de competencias docentes percibidas en Educación Física en el contexto mexicano.

Palabras clave: competencia docente, Educación Física, psicometría, invarianza factorial, adolescentes.

Introducción

De acuerdo con la Organización de las Naciones Unidas para la Educación la Ciencia y la Cultura (2022), el docente posee un rol esencial en la transformación educativa, para ello es deseable el cumplimiento de un perfil profesional que incluye el desarrollo de competencias. Se reconoce el esfuerzo del gobierno de México para impulsar proyectos de formación continua que fortalecen las competencias de los profesores (Secretaría de Educación Pública, 2024). Sin embargo, los resultados del *Programme for International Student Assessment* (PISA) indican que la educación de dicho país tiene oportunidades de mejora (Organización para la Cooperación y el Desarrollo Económicos, 2022), lo que sugiere la necesidad fortalecer la educación.

En este sentido, analizar el desempeño docente adquiere relevancia. La Unidad del Sistema para la Carrera de las Maestras y Maestros (USICAMM) es la encargada de realizar la evaluación administrativa para la admisión, promoción y acceso a estímulos laborales (Secretaría de Educación Pública & Unidad del Sistema para la Carrera de las Maestras y los Maestros, 2023). Por otro lado, el Modelo de Evaluación Diagnóstica, Formativa e Integral (MEDFI) cuenta con un enfoque orientado hacia la mejora continua del profesorado (Comisión Nacional para la Mejora Continua de la Educación, 2022). De acuerdo con Rosas y Velázquez (2026) las normativas de evaluación docente requieren nuevas consideraciones, por lo que se contempla un nuevo modelo de evaluación. Esto resalta la existencia de inconsistencias que limitan identificar el desempeño docente y justifica contar con instrumentos de evaluación que respondan a esta demanda educativa.

En esta misma línea de ideas, los últimos datos publicados indicaron que cerca del 50 % de los docentes tiene un nivel de desempeño entre insuficiente y suficiente, mientras que el 7.9 % alcanzaron un desempeño destacado (Secretaría de Educación Pública, 2016). La evaluación del debe considerarse como un referente que posibilite el diseño de programas de actualización pedagógica, aportando al sistema educativo (Rosas y Velázquez, 2026).

Las competencias se refieren al complejo conjunto de desempeños, habilidades, conocimientos y actitudes de una persona y que le son de utilidad para resolver situaciones problemáticas en un contexto específico (Tobón, 2007). Es deseable que el profesorado de Educación Física que imparte clases en nivel básico (preescolar, primaria y secundaria) tenga altos niveles de competencias referentes al conocimiento específico de la asignatura, comunicación asertiva, liderazgo (McKown et al., 2024), para la valoración de habilidades motrices, didácticas y de evaluación de los aprendizajes de los estudiantes (Cañadas et al., 2020; Lounsbury y Coker, 2008), para favorecer inclusión para estudiantes con necesidades educativas especiales (Demchenko et al., 2021), para la aplicación de actividades lúdicas, de igual forma resalta la importancia de habilidades para establecer un juicio crítico y reflexivo en referencia a la asignatura (Asún et al., 2020), entre las competencias más destacadas.

La literatura científica indica la relación existente entre las clases de Educación Física que incluyen actividades motivantes, que estimulan la autonomía e interacción social con la satisfacción de los estudiantes, a su vez la satisfacción está vinculada con la predisposición de los escolares para realizar actividad física en su tiempo libre (Baena-Extremera et al., 2016; Cuevas-

Campos et al., 2020; Leyton-Roman et al., 2020). Este planteamiento resalta la trascendencia de que el profesorado de Educación Física sea un profesional competente; sin embargo, en el caso de México, desde la perspectiva de las autoridades educativas es necesario la realización de estrategias para mejorar el desempeño del profesorado (González-Rivas et al., 2022). Se trata de una problemática que no es exclusiva de dicho país, debido a que es recurrente en diversos contextos geográficos, sociales y económicos (Cox et al., 2023; Lander et al., 2017). Esto manifiesta la relevancia de conocer el nivel de dominio de las competencias del profesorado, para la posible implementación y evaluación de programas de formación continua, dicha valoración debe realizarse por medio de instrumentos que den certeza de las competencias con las que cuenta el profesorado (Cox et al., 2023; Yin et al., 2024).

En este sentido, existen diversos instrumentos para identificar el nivel de dominio de las competencias docentes. Tal es el caso del *Contemporary Teacher Classroom Performance Observation Instruments* (Lavelly et al., 1994) o el *Teachers' General Pedagogical/Psychological Knowledge* (Voss et al., 2011). El interés de conocer sobre el desarrollo de competencias en el contexto hispanoparlante es evidente, con la adaptación de cuestionarios al castellano; por ejemplo, el *Teachers' Sense of Efficacy Scale* (TSES) en Chile (Gálvez-Nieto et al., 2023) o el *Questionnaire DigCompEdu Check-In* en España (Cabero-Almenara et al., 2020). Para el presente estudio se tomó como referencia el *Evaluation of Teaching Competencies Scale* (ETCS) (Catano y Harvey, 2011), pero en la versión propuesta por Baena-Extremera et al. (2015): la Escala de Evaluación de la Competencia Docente en Educación Física de Secundaria, la cual representa un referente teórico sólido para la evaluación de la competencia docente a partir de ocho dimensiones. Se optó por este instrumento debido a que está validado y adaptado para medir específicamente las competencias del profesorado de Educación Física, diferenciándolo de los otros instrumentos orientados a la autoeficacia y el conocimiento pedagógico. Además, la adaptación previa al español representa una base lingüística adecuada para su aplicación con estudiantes mexicanos.

Estudios previos han demostrado una adecuada estructura psicométrica de esta versión adaptada a la Educación Física con estudiantes mexicanos (Baños et al., 2022). Sin embargo, la evidencia psicométrica continúa siendo limitada a un único contexto socioeducativo, resaltando la importancia de probar el comportamiento del instrumento en diferentes grupos poblacionales, así como su equivalencia en diferentes grupos. Esto adquiere relevancia, debido a que México es un país caracterizado por su diversidad social, económica y educativa. Además, este cuestionario cuenta con ítems que parecen integrar más de un contenido conceptual, lo que podría confundir a la persona respondiente. Ante esto, autores como Worthington y Whittaker (2006) recomiendan que cada ítem represente claramente un sólo concepto. Por este motivo, surge el interés científico de desarrollar una escala para valorar el nivel de competencias docentes en el profesorado de Educación Física, a partir del ETCS y procurando precisión conceptual y precisión en la redacción de los ítems.

En síntesis, la mejora de la calidad académica está condicionada por diversos factores, siendo el docente una figura que adquiere mayor relevancia, dado que su desempeño profesional influye directamente con la adquisición de aprendizajes de los estudiantes. Por esto, es necesario

identificar el nivel de dominio de competencias docentes por parte del profesorado de EF. Contar con un instrumento psicométrico, válido, confiable y adaptado al contexto socioeducativo de México, representa un recurso relevante para este fin. El objetivo de este estudio fue desarrollar y obtener evidencias de validez estructural e invarianza factorial de una escala para la evaluación de la competencia docente percibida del profesorado de Educación Física en adolescentes de secundaria, a partir de la adaptación y ampliación de la ETCS (Catano y Harvey, 2011) y la adaptación al español de Baena-Extremera et al. (2015).

Metodología

El presente estudio se enmarca en un diseño instrumental el cual está orientado al desarrollo y validación de un instrumento de medida, siguiendo las recomendaciones para estudios psicométricos en el ámbito de las ciencias sociales (Ato et al., 2013).

La selección de la muestra fue de tipo no probabilístico y por conveniencia, considerando el acceso a los centros educativos y la disponibilidad de los participantes. Todos ellos provenientes de diez escuelas de nivel secundaria inscritos en la clase de Educación Física del noroeste de México, de las cuales cinco eran estatales, cuatro federales y una privada. Se llevaron a cabo dos estudios: El estudio 1, fue realizado con una muestra de desarrollo (N=320), la cual tuvo como objetivo la generación, depuración y validación inicial del instrumento, a través del Análisis Factorial Exploratorio (AFE) y el Análisis Factorial Confirmatorio (AFC) iterativo. El estudio 2, se realizó con una muestra de validación independiente (N=598), para confirmar la estructura factorial obtenida en el estudio 1 y examinar la invarianza de medición de acuerdo con el sexo y grado.

En el estudio 1 participaron 320 adolescentes, con edades comprendidas entre 12 y 15 años (M=13.22; DE=.96). Del total de participantes, 179 fueron mujeres (55.93%) y 141 hombres (44.07%), distribuidos entre los tres grados de secundaria: 1° (n=126, 39.4%), 2° (n=99, 30.9%), y 3° (n=95, 29.7%). En el estudio 2 participaron en el estudio 598 adolescentes con edades comprendidas entre 12 y 15 años (M=13.27; DE=.91). Del total de participantes, 328 fueron mujeres (54.8%) y 270 hombres (45.2%), distribuidos entre 1° (n=214, 35.8%), 2° (n=198, 33.1%) y 3° (n=186, 31.1). Las características sociodemográficas se presentan en la tabla 1.

Tabla 1

Características sociodemográficas de las muestras por estudio

Variable		Muestra 1 (N=320)		Muestra 2 (N=598)	
		n	%	n	%
Sexo	Mujer	179	55.93	328	54.8
	Hombre	141	44.07	270	45.2
Grado	1°	126	39.4	214	35.8
	2°	99	30.9	198	33.1
	3°	95	29.7	186	31.1
Edad (M±DE)		12 y 15 años 13.22±0.96		13.27±0.91	

Nota: M=Media; DE=Desviación estándar

Como instrumento de referencia se empleó la ETCS desarrollada por Catano y Harvey (2011), en su versión adaptada al español en el contexto de educación secundaria en la clase de Educación Física por Baena-Extremera et al. (2015). La escala original evalúa la percepción del alumnado sobre la competencia docente a través de ocho ítems: comunicación, conciencia de trabajo, creatividad, *feedback*, consideración individual al alumno, profesionalidad, resolución de problemas y conciencia social. En el instrumento se solicita que indiquen el grado de acuerdo con los ítems en una escala politómica de 1 a 7, en ocho constructos: Los autores reportaron una consistencia interna de .90.

En función del objetivo del estudio, orientado a obtener evidencias de validez estructural en población adolescente en el contexto específico de la Educación Física. Se procedió a la generación de una batería ampliada de 18 ítems unitarios a partir de los descriptores conductuales del ETCS, con el propósito de garantizar una adecuada cobertura y operacionalización del constructo. El formato de respuesta se reformuló hacia una escala de frecuencia anteponiendo la frase en cada ítem: “Con qué frecuencia el profesor de Educación Física...” con 11 opciones de respuesta que oscilaron entre 0 (nunca), 1-3 (casi nunca), 4-6 (a veces), (7-9) casi siempre y 10 (siempre). Los ítems iniciales se distribuyeron en los ocho constructos teóricos propuestos por la ETCS. Después de los análisis psicométricos realizados, se obtuvo una versión final de 10 ítems con una estructura unidimensional (modelo CODEF-10), la cual demostró indicadores de validez y fiabilidad adecuados para la evaluación de la competencia docente percibida. Con el propósito de conservar la coherencia teórica de la escala, se organizaron los ítems finales de acuerdo con los constructos originales de la ETCS con el fin de evidenciar la correspondencia conceptual (Tabla 2). A pesar de esta organización teórica, los análisis factoriales evidenciaron una estructura unidimensional,

Tabla 2

Validez de contenido de los ítems finales del Cuestionario de competencia docente percibida en EF (CODEF-10).

Dimensión teórica de la ETCS	Ítem	Número de ítem y contenido
F1 Comunicación	3	Indica de manera clara las instrucciones, las explicaciones y la asignación de tareas.
	4	Escucha a los estudiantes.
F2 Conciencia de trabajo	5	Imparte los contenidos que se indican en el programa de la clase.
	8	Proporciona el programa del curso.
F3 Creatividad	9	Utiliza múltiples formas de enseñar las clases.
F4 Feed-Back	10	Me proporciona información detallada durante las actividades para la mejor realización de las tareas.
F5 Consideración individual al alumnado	12	Suele estar disponible para ayudarme.
F6 Profesionalidad	14	Aclara los criterios de evaluación.
F7 Resolución de problemas	15	Explica cuando la actividad no está clara.
F8 Conciencia social	17	Consigue mantener el orden en clase.

Se realizó una reunión informativa con las autoridades educativas y directores de escuelas secundarias, se explicó el alcance, beneficios, requerimientos y riesgos de este proyecto, finalmente se obtuvo la aprobación para llevarlo a cabo. Con el apoyo de los directores de los centros educativos que aceptaron participar, se envió a los padres de familia el consentimiento informado. Se realizó la recolección de datos en el horario escolar. El grupo de investigadores acudió a las secundarias, con los estudiantes que entregaron el consentimiento firmado por los padres de familia o tutores legales, se compartió un código QR para que tuvieran acceso al cuestionario de forma digital por medio de Google Formularios. Es necesario mencionar que en la primera pantalla aparece la solicitud de asentimiento por parte del estudiante, en donde se menciona que la participación es voluntaria, anónima y que sus respuestas no tendrían ningún tipo de repercusión académica; también se hizo énfasis en la relevancia de la honestidad. Este proceso tuvo un tiempo aproximado de 20 minutos y en todo momento estuvo presente un investigador para cualquier duda o aclaración.

El análisis de datos se llevó a cabo en varias fases con el objetivo de examinar las propiedades psicométricas del instrumento, utilizando los softwares IBM SPSS Statistics 21, IBM AMOS 24 y JASP. En el estudio 1, el primer paso del análisis consistió en la descripción de los ítems a través de las medias, desviaciones estándar y los supuestos de normalidad, a través de la asimetría y la curtosis. Se verificó la adecuación de la muestra mediante el índice KMO y la prueba de esfericidad de Barlett.

Se realizó el AFE en la muestra de desarrollo ($n=320$), sobre los 18 ítems iniciales, mediante extracción por Máxima Verosimilitud (ML) y rotación Varimax. La retención de factores se determinó por el criterio de autovalores >1 y el análisis paralelo (Hair et al., 2019). La depuración de los ítems se realizó considerando comunales inferiores a .30 y criterios de coherencia teórica, lo que redujo la escala a 13 ítems.

Después se realizó un AFC sobre la estructura preliminar CODEF-13 utilizando el estimador ML. El refinamiento del modelo fue realizado de manera iterativa considerando los índices de modificación ($IM > 10$) en coherencia con el análisis de contenido previo de los ítems y buscando el modelo más parsimonioso. Este proceso dio lugar a la versión final del instrumento compuesta por 10 ítems (CODEF-10).

El ajuste del modelo se evaluó mediante: χ^2/gl (<3 estricto, <5 aceptable), GFI, AGFI, CFI, TLI ($\geq .90$), SRMR y RMSEA ($\leq .08$), AVE ($\geq .50$) y CR ($\geq .70$) (Hair et al., 2019; Sathyanarayana y Mohanasundaram, 2024). De igual forma se consideró que las cargas factoriales estandarizadas $\geq .60$ son aceptables (Hair et al., 2019). El AVE fue calculada a partir del promedio de las cargas factoriales estandarizadas elevadas al cuadrado de los ítems del factor latente. De forma similar, utilizando el cuadrado de las cargas factoriales y la suma de las varianzas de error asociadas, fue calculado el CR (Fornell & Larcker, 1981).

En el estudio 2, la estructura final CODEF-10 obtenida en la muestra de desarrollo fue sometida a un AFC en una muestra independiente de validación ($n=598$), con el objetivo de

contrastar su estabilidad factorial y aportar evidencia adicional de validez estructural. Sobre esta muestra también se estimaron los índices de fiabilidad (α , Ω , AVE y CR).

Posteriormente se examinó la invarianza factorial de acuerdo con el sexo mediante comparación de modelos anidados en tres niveles: configural, métrico y escalar. Los criterios de decisión fueron mediante el $\Delta CFI \leq .010$ y $\Delta RMSEA \leq .015$ Cheung y Rensvold (2002), si los cambios no superan estos umbrales al ir añadiendo restricciones, la invarianza se considera sostenida. Dado que se comprobó la invarianza entre los grupos, se realizó también una comparación de medias latentes entre hombres y mujeres, fijando las medias de los factores a cero en el grupo de hombres y estimándolas libremente en las mujeres. La consistencia interna se estimó mediante los coeficientes alfa de Cronbach (α) y omega de McDonald (Ω), reportándose por cada muestra.

Resultados

El Análisis Factorial Exploratorio (Estudio 1). Primero se realizó un AFE sobre los 18 ítems iniciales en la muestra de desarrollo ($n=320$), donde los índices de adecuación muestral mostraron valores satisfactorios ($KMO=.928$) y la prueba de esfericidad de Barlett fue significativa ($p<.001$), lo que indica la idoneidad de los datos para el análisis factorial.

El criterio de autovalores mayores a uno y el análisis paralelo sugirieron inicialmente una solución de dos factores. El primer factor presentó un autovalor de 7.965, explicando el 44.25% de la varianza total, mientras que el segundo factor obtuvo un autovalor de 2.031. Además, los ítems 6, 16 y 18 presentaron comunales menores a .30, por lo que fueron eliminados. Asimismo, la inspección de la matriz factorial mostró que el segundo factor estaba constituido exclusivamente por los ítems 1 y 2, ambos redactados en sentido inverso. En consecuencia, dichos ítems fueron eliminados y se continuó el proceso de validación bajo una estructura unidimensional de 13 ítems. Después de este proceso de depuración, todos los reactivos restantes presentaron cargas factoriales adecuadas y coherencia conceptual con el constructo evaluado.

Posteriormente, se llevó a cabo la fase de refinamiento a través del AFC en la misma muestra ($n=320$). El modelo inicial de 13 ítems (CODEF-13) mostró un ajuste no aceptable, por lo que se procedió a la depuración iterativa basada en los índices de modificación ($IM>10$), la redundancia conceptual entre ítems y la coherencia teórica del constructo. Esto permitió mejorar la parsimonia del modelo y obtener una estructura más estable, eliminando progresivamente los ítems con covarianzas residuales elevadas o un solapamiento conceptual para llegar al modelo unidimensional de 10 ítems (CODEF-10). Los índices de ajuste del modelo demostraron una estructura aceptable: $\chi^2/gl=2.510$; $CFI=.972$; $TLI=.964$; $GFI=.947$; $AGFI=.917$; $RMSEA=.069$; y $SRMR=.0315$, explicando el 60.62% de la varianza total (Tabla 4). Todas las cargas factoriales fueron superiores a .70, excepto en el ítem 3 (.63) (Tabla 5).

El Análisis Factorial Confirmatorio (Estudio 2). Se calcularon los estadísticos descriptivos de la versión final del cuestionario (CODEF-10), los ítems mostraron que las medias oscilaron

entre 6.93 y 8.00, la desviación estándar entre 2.62 y 3.14, para los valores de asimetría ± 1.34 y curtosis ± 0.92 , por lo que se infiere que las variables se ajustan adecuadamente a una distribución normal (Tabla 3).

Tabla 3

Estadísticos descriptivos de los ítems de la CODEF-10 con la muestra n=598

Ítem	M	DE	CCIT-c	α sin ítem	AS	CU
3	7.25	3.14	.58	.92	-1.10	-.17
4	7.76	2.79	.72	.91	-1.17	.41
5	7.28	2.85	.74	.91	-.88	-.18
8	7.04	2.91	.71	.91	-.86	-.16
9	7.32	2.98	.75	.91	-1.04	.08
10	7.05	3.00	.76	.91	-.88	-.21
12	6.93	3.08	.70	.91	-.83	-.38
14	7.90	2.77	.69	.91	-1.34	.91
15	8.00	2.62	.71	.91	-1.32	.92
17	7.58	2.75	.66	.91	-1.09	.30

Nota: M=Media; DE=Desviación Estándar; CCIT-c=Correlación ítem total corregida; α =Coeficiente alfa de Cronbach; AS=Asimetría; CU=Curtosis

Se evaluó el modelo resultante (CODEF-10) en una muestra independiente (n=598), en el cual los índices de ajuste $\chi^2/\text{gl}=4.590$; CFI=.964; TLI=.953; GFI=.950; AGFI=.918; RMSEA=.078; y SRMR=.0323, este modelo unidimensional explica el 59.29% de la varianza total (Tabla 4).

Todas las cargas factoriales fueron superiores a .70, excepto en el ítem 3 (.60) e ítem 17 (.69) el cual estuvo muy cercano (Tabla 5). Durante el proceso de refinamiento se agregó una covarianza residual entre los ítems 5 y 8, ésta se justifica por la proximidad conceptual entre los indicadores, donde la redacción se refiere a la planificación y la organización de los contenidos del curso. Esta modificación en el modelo permitió la mejora significativa del ajuste del modelo final.

Tabla 4

Índices absolutos, incrementales y de parsimonia para los modelos generados de la CODEF

Modelo	Índices absolutos				Índices incrementales			Índices de parsimonia	
	χ^2	GFI	RMSEA	SRMR	AGFI	TLI	CFI	χ^2/gl	AIC
CODEF-10 N=320	87.834	.947	.069	.0315	.917	.964	.972	2.510	127.834
CODEF-10 N=598 sin covarianza	187.871*	.937	.086	.0349	.901	.942	.955	5.368	227.871
CODEF-10 N=598 con covarianza	156.053*	.950	.078	.0323	.918	.953	.964	4.590	198.053

Nota: * $p < .05$; GFI = índice de bondad de ajuste; RMSEA = error cuadrático medio de aproximación; AGFI = índice de bondad ajustado; TLI = índice de Tucker-Lewis; CFI = índice de ajuste comparativo; χ^2/gl = razón de Chi-cuadrado sobre los grados de libertad; AIC = criterio de información de Akaike

La confiabilidad de la escala fue excelente en ambas muestras con valores en α y Ω superiores a .90 (Tabla 5) de acuerdo a los criterios de George y Mallery (2024). Asimismo, los valores de la confiabilidad compuesta ($CR > .90$) y la varianza de los ítems explicada por el constructo latente indicó una adecuada validez convergente ($AVE \geq .50$) (Hair et al., 2019).

Tabla 5

Soluciones estandarizadas del AFC para el modelo por muestras.

Item	CODEF-10 N=320	CODEF-10 N=598
Pesos Factoriales		
3	.63	.60
4	.76	.76
5	.72	.76
8	.71	.72
9	.76	.79
10	.81	.80
12	.76	.74
14	.77	.73
15	.79	.74
17	.71	.69
Consistencia interna	$\alpha = .927$	$\alpha = .922$
	$\Omega = .928$	$\Omega = .923$
	AVE = .54	AVE = .54
	CR = .92	CR = .92

Nota: α =Coeficiente alfa de Cronbach; Ω =Coeficiente Omega de McDonald;
 AVE=Varianza media extraída; CR=Confiabilidad compuesta

Se examinó la invarianza en función de las variables sexo y grado escolar, a través del análisis multigrupo. Los resultados mostraron un ajuste adecuado ($CFI = .953$; $RMSEA = .063$) del modelo CODEF-10 (N=598) en los niveles configural, métrico, escalar y estricto evaluados, ya que los cambios en los índices de ajuste fueron mínimos ($\Delta CFI \leq .01$ y $\Delta RMSEA \leq .015$), por lo que se comprueba la equivalencia de la estructura entre hombres y mujeres ($CFI = .944$; $RMSEA = .068$), métrica ($\Delta CFI \leq .002$), escalar y estricta ($\Delta CFI \leq .010$), además se respalda la invarianza de medición entre estudiantes de distintos grados escolares ($CFI = .942$; $RMSEA = .058$), métrica ($\Delta CFI \leq .002$), escalar y estricta ($\Delta CFI \leq .010$) (Tabla 6).

Tabla 6

Índices de bondad de ajuste de cada uno de los modelos puestos a prueba en la invarianza factorial por sexo y grado escolar

Invarianza por sexo							
Modelo	Índice de Ajuste						
	χ^2	gl	GFI	NFI	CFI	RMSEA	AIC
Configural	228.738	68	.929	.935	.953	.063	312.738
Métrica	245.471	77	.925	.931	.951	.061	311.471
Escalar	249.391	78	.924	.929	.950	.061	3313.391
Invarianza por grado escolar							
Modelo	Índice de Ajuste						
	χ^2	gl	GFI	NFI	CFI	RMSEA	AIC
Configural	306.862	102	.909	.916	.942	.058	432.862
Métrica	330.360	120	.902	.910	.940	.054	420.360
Escalar	330.655	122	.902	.910	.941	.054	416.655

Nota: * $p < .05$; GFI = índice de bondad de ajuste; NFI = índice de ajuste normado; CFI = índice de ajuste comparativo; RMSEA = error cuadrático medio de aproximación; AIC = criterio de Información de Akaike

Una vez comprobada la invarianza escalar por sexo y grado escolar, se compararon las medias latentes. En el caso de la variable sexo, las mujeres mostraron una puntuación ligeramente inferior en la competencia percibida docente con respecto a los hombres ($\Delta M = -.306$, $SE = .206$, $CR = -1.870$, $p = .061$; $IC95\% [-.710, .098]$). Mientras que en la comparación de las medias latentes de acuerdo con el grado escolar tomando como referencia a los estudiantes de primer grado, no se encontraron diferencias significativas entre primero y segundo grado ($\Delta M = -.193$, $SE = .248$, $CR = -.778$, $p = .436$; $IC95\% [-.679, .293]$), ni entre primero y tercero ($\Delta M = -.440$, $SE = .258$, $CR = -1.706$, $p = .88$; $IC95\% [-.946, .066]$). En todos los casos, los intervalos de confianza incluyeron el valor cero, sugiriendo que las diferencias observadas fueron pequeñas y compatibles con la ausencia de diferencias sustantivas en la competencia docente percibida entre los grupos comparados.

Discusión

El objetivo del presente estudio fue desarrollar y obtener evidencias de validez estructural de una escala para evaluar la competencia docente percibida de los profesores de Educación Física en adolescentes de secundaria, partiendo de la adaptación y ampliación de la ETCS. De manera global, los resultados respaldan la validez y fiabilidad de la versión final propuesta con un modelo unidimensional de 10 ítems (CODEF-10), la cual aporta evidencia empírica a este constructo en este contexto específico. Los resultados del AFE evidenciaron una solución predominantemente unidimensional, donde el primer factor explicó una proporción sustancial de la varianza total.

Aunque surgió un segundo factor en la fase inicial, estuvo compuesto por dos ítems que estaban redactados en sentido inverso, lo que sugiere la presencia de varianza de método, más que una dimensión teórica diferenciada. Esto es consistente con la literatura psicométrica, donde se advierte que este tipo de ítems pueden presentar artefactos factoriales (Tomás et al., 2013), especialmente en poblaciones de adolescentes, estudios sobre medidas aplicadas a adolescentes muestran que los modelos que separan la varianza de método por ítems invertidos mejoran considerablemente el ajuste, además se han identificado patrones de respuesta inconsistentes asociados a ítems invertidos que afectan la fiabilidad y validez en jóvenes estudiantes (Antoniou y Alghamdi, 2024; Sirlopú et al., 2019). Por lo que, la eliminación de estos ítems contribuyó a mejorar la claridad en la estructura de la escala. Este proceso de depuración a través del AFC permitió llegar a una versión ampliada de 10 ítems, conservando aquellos indicadores con mayor consistencia empírica y coherencia teórica.

Debido a que el presente instrumento fue desarrollado tomando como referencia la ETCS de Catano y Harvey (2011), en su versión adaptada al español por Baena-Extremera et al. (2015), es relevante comparar los hallazgos psicométricos entre estos estudios. Se identificó una serie de consistencias parciales: valores aceptables del *CCIT-c* (Nunnally y Bernstein, 1995), cargas factoriales en su mayoría aceptables (Hair et al., 2019), una consistencia interna similar y adecuada (Revelle y Zinbarg, 2009), e índices de ajuste menos robustos pero aceptables (Byrne, 2013). Estas similitudes aportan evidencia favorable de la estructura factorial del instrumento respecto a la medición del constructo de competencia docente medida en estos cuestionarios.

Asimismo, es importante resaltar que, inclusive con las modificaciones realizadas, se mantiene una estructura unifactorial. Esto indica que las variables observables responden congruentemente al factor latente. Sumando a esto, los constructos evaluados en el CODEF-10 corresponden con las dimensiones del instrumento propuesto por Baena-Extremera et al. (2015). Esto sugiere que se conservó la congruencia teórica y operativa, pero se optimizó la redacción de los ítems en favor de un modelo más parsimonioso (Byrne, 2013). Esto implica que las adaptaciones realizadas no comprometieron el constructo analizado y permite mediciones precisas para evaluar el nivel de competencia del profesorado del Educación Física desde la perspectiva de los estudiantes. Por otro lado, la convergencia existente entre estos instrumentos respecto a las dimensiones en un factor, pueden indicar estrechas relaciones en la comprensión del desempeño docente por parte de los estudiantes. Es decir que, para ellos, la comunicación, la resolución de problemas, la creatividad y los otros elementos que componen la escala, se manifiestan de forma simultánea durante las clases de Educación Física (Catano & Harvey, 2011).

Por otro lado, las discrepancias entre estos estudios pudieran deberse, a factores contextuales que enfrenta cada sistema educativo. En el caso de México, se ha señalado la existencia de grupos de profesores que se resisten a la actualización pedagógica y disciplinar; además de inconsistencias administrativas, como profesores que no cumplen con el perfil adecuado para impartir clases de Educación Física (G.-Olvera et al., 2021; González-Rivas et al., 2023). En España, existen distintos desafíos; la falta de congruencia entre los contenidos de los cursos de formación continua y las necesidades del profesorado, así como la carencia de estímulos, han sido

reportadas en estudios recientes (Serrano-Brazo et al., 2023). Estas condiciones contextuales podrían influir en el nivel de dominio de las competencias, lo que contribuye a explicar parcialmente las diferencias entre estos estudios. Aunque se reconoce que se tratan de hipótesis que buscan explicar los resultados obtenidos y que son necesarios estudios que las comprueben.

Por otro lado, otra característica relevante de este estudio fue la modificación en las opciones de la escala de respuesta, que son congruentes con otros estudios realizados en México (Blanco et al., 2021; González-Rivas et al., 2026). De acuerdo con Alzina y Pérez-Escoda (2015), esta modificación favorece la sensibilidad de la respuesta, mejorando la precisión discriminativa entre cada categoría, lo que permite identificar con mayor precisión el nivel de dominio de la competencia docente.

En relación a la invarianza de acuerdo al sexo, el resultado del presente estudio se encuentra en la misma línea que los hallazgos de Martínez-Molina et al. (2020), quienes aplicaron la versión adaptada de Baena-Extremera et al. (2015). La consistencia en estos resultados sugiere que la aplicación de este instrumento ofrece resultados equivalentes, independientemente del sexo de los estudiantes. Adicionalmente, el presente estudio proporciona evidencia de la invarianza según el grado escolar de los estudiantes. Estos hallazgos sugieren que la interpretación de los ítems es similar entre hombres y mujeres de los tres grados escolares de nivel secundaria (Byrne, 2013). En este orden de ideas, es posible realizar comparaciones significativas entre diferentes grupos, con menor posibilidad de sesgos derivados del instrumento; por lo que los resultados pueden ser interpretados con mayor confianza (Putnick & Bornstein, 2016).

Por último, los análisis de medias latentes mostraron que no existen diferencias estadísticamente significativas ni entre hombres y mujeres, a diferencia de los reportados por (Martínez-Molina et al., 2020; Wei, 2025); ni entre los distintos grados escolares. Este hallazgo sugiere que la percepción de la competencia docente en EF es relativamente homogénea entre estos grupos estudiados, algo que se puede interpretar como una consistencia en las experiencias educativas de los estudiantes o en los criterios que se utilizan para evaluar a sus docentes.

Los resultados de esta investigación deben ser interpretados con detenimiento, esto debido a las siguientes limitaciones. La muestra del estudio fue no probabilística y por conveniencia, por lo que no es posible generalizar los hallazgos obtenidos. Al ser un instrumento de tipo autoinforme, es posible la existencia de sesgo de deseabilidad social, haciendo referencia a que el participante responde lo que socialmente se considera correcto, dejando de lado la veracidad de su respuesta (Krumpal, 2013). De igual forma, el estudio se enfocó en la validez del constructo, por lo que no fueron contemplados otros criterios de validación (Indu et al., 2025). Además, la estructura factorial obtenida tuvo una ruta metodológica que incluyó dos estudios, en diferentes tamaños de muestra y la depuración de los ítems y reespecificación del modelo a partir del sustento estadístico y teórico. En este sentido, es posible que algunas soluciones factoriales puedan estar influenciadas por las particularidades de las muestras. Por esto, los resultados deben ser considerados con cautela. Se recomienda que próximas investigaciones, que deseen aportar a la estructura psicométrica de este instrumento contemplen una muestra probabilística. Asimismo, resulta de interés considerar la validación cruzada, temporal y de criterio.

En cuanto a la practicidad de este estudio, es de utilidad para diversos actores involucrados en dar seguimiento al desempeño del profesorado de Educación Física. En este sentido, para los investigadores este instrumento es conveniente en la recolección de datos, debido a su practicidad de aplicación por estar constituido por solamente 10 ítems. Para las autoridades educativas puede utilizarse para conocer el nivel de dominio de las competencias genéricas del profesorado, siendo un referente que sirva de diagnóstico, y contribuya a orientar los contenidos de los programas de formación continua.

Conclusiones

La escala de competencia docente percibida en Educación Física (CODEF-10) evidenció adecuadas propiedades psicométricas para evaluar las competencias del profesorado desde la perspectiva de los estudiantes. El AFE y los AFC respaldan la estructura factorial, la invarianza y el contraste de las medias la equivalencia de medida. Es una escala invariante en cuanto al sexo y grado escolar de los estudiantes, favoreciendo su aplicabilidad de forma equivalente. Debido a su carácter invariante, resulta ser una herramienta útil para realizar estudios comparativos. No obstante, son necesarios estudios instrumentales que continúen aportando evidencia de la estructura factorial de este cuestionario, con muestras más amplias y de distintos contextos socioeducativos, además del interés de comprobar la validez de criterio y estabilidad temporal.

Referencias

- Alzina, R. B., & Pérez-Escoda, N. (2015). Les escales de Likert poden augmentar en sensibilitat? *REIRE Revista d'Innovació i Recerca en Educació*, 8(2), 129–147-129–147. <https://doi.org/10.1344/reire2015.8.2828>
- Antoniou, F., & Alghamdi, M. H. (2024). Confidence in mathematics is confounded by responses to reverse-coded items. *Frontiers in Psychology*, 15, 1489054. <https://doi.org/10.3389/fpsyg.2024.1489054>
- Asún, S., Chivite, M. T., & Romero, M. R. (2020). Perceptions of professional competences in physical education teacher education (PETE). *Sustainability*, 12(9), 3812. <https://doi.org/10.3390/su12093812>
- Ato, M., López-García, J. J., & Benavente, A. (2013). Un sistema de clasificación de los diseños de investigación en psicología. *Anales de Psicología/Annals of Psychology*, 29(3), 1038-1059. <https://doi.org/10.6018/analesps.29.3.178511>
- Baena-Extremera, A., Gómez-López, M., Granero-Gallegos, A., & Martínez-Molina, M. (2016). Modelo de predicción de la satisfacción y diversión en Educación Física a partir de la autonomía y el clima motivacional. *Universitas Psychologica*, 15(2), 39-49. <https://doi.org/10.11144/Javeriana.upsy15-2.mpsd>

- Baena-Extremera, A., Granero-Gallegos, A., & Martínez-Molina, M. (2015). Validación española de la Escala de Evaluación de la Competencia Docente en Educación Física de secundaria. *Cuadernos de psicología del deporte*, 15(3), 113-122. <https://doi.org/10.4321/S1578-84232015000300011>
- Baños, R., Machado-Parra, J. P., Barretos-Ruvalcaba, M., Pérez-Morán, J. C., & Baena-Extremera, A. (2022). Propiedades psicométricas y medición de la invarianza por género de la Escala de Evaluación de Competencias Docentes de Educación Física en el contexto mexicano. *Retos*, 46, 349-357. <https://doi.org/10.47197/retos.v46.93958>
- Blanco, J. R., Ivonne Aguirre, S., Jiménez Lira, C., Mondaca-Fernández, F., & Jurado García, P. J. (2021). Composición e invarianza factorial del cuestionario MIFA en adolescentes deportistas y no deportistas. *Sportis. Scientific Journal of School Sport, Physical Education and Psychomotricity*, 7(1), 131-149. <https://doi.org/10.17979/sportis.2021.7.1.5995>
- Byrne, B. M. (2013). *Structural equation modeling with Mplus: Basic concepts, applications, and programming*. routledge. <https://doi.org/10.4324/9780203807644>
- Cabero-Almenara, J., Gutiérrez-Castillo, J.-J., Palacios-Rodríguez, A., & Barroso-Osuna, J. (2020). Development of the teacher digital competence validation of DigCompEdu check-in questionnaire in the university context of Andalusia (Spain). *Sustainability*, 12(15), 6094. <https://doi.org/10.3390/su12156094>
- Cañadas, L., Santos-Pastor, M. L., & Javier Castejón, F. (2020). Physical Education Teachers' Competencies and Assessment in Professional Practice. *Apunts: Educació Física i Esports*(139), 33-41. [https://doi.org/10.5672/apunts.2014-0983.es.\(2020/1\).139.05](https://doi.org/10.5672/apunts.2014-0983.es.(2020/1).139.05)
- Catano, V. M., & Harvey, S. (2011). Student perception of teaching effectiveness: development and validation of the Evaluation of Teaching Competencies Scale (ETCS). *Assessment & Evaluation in Higher Education*, 36(6), 701-717. <https://doi.org/10.1080/02602938.2010.484879>
- Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural equation modeling*, 9(2), 233-255. https://doi.org/10.1207/S15328007SEM0902_5
- Comisión Nacional para la Mejora Continua de la Educación. (2022). *Modelo de evaluación diagnóstica, formativa e integral. La evaluación al servicio de la mejora continua de la educación*. Gobierno de México. https://guias.usicamm.gob.mx/documentos/PVEB/direccion_educacion_especial/3.pdf
- Cox, A., Noonan, R. J., & Fairclough, S. J. (2023). PE teachers' perceived expertise and professional development requirements in the delivery of muscular fitness activity: PE Teacher EmPOWERment Survey. *European Physical Education Review*, 29(2), 251-267. <https://doi.org/10.1177/1356336X221134067>

- Cuevas-Campos, R., Fernández-Bustos, J. G., González-Cutre, D., & Hernández-Martínez, A. (2020). Need satisfaction and need thwarting in physical education and intention to be physically active. *Sustainability*, 12(18), 7312. <https://doi.org/10.3390/su12187312>
- Demchenko, I., Maksymchuk, B., Bilan, V., Maksymchuk, I., & Kalynovska, I. (2021). Training future physical education teachers for professional activities under the conditions of inclusive education. *BRAIN. Broad Research in Artificial Intelligence and Neuroscience*, 12(3), 191-213. <https://doi.org/10.18662/brain/12.3/227>
- Fornell, C., & Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of marketing research*, 18(1), 39-50. <https://doi.org/10.1177/002224378101800104>
- G.-Olvera, A., Pacheco-Miranda, S., Orozco Núñez, E., Aguilar-Farias, N., López Fernández, I., & Jáuregui, A. (2021). Implementation of physical education policies in elementary and middle schools in Mexico. *Health Promotion International*, 36(2), 460-470. <https://doi.org/10.1093/heapro/daaa068>
- Gálvez-Nieto, J. L., Salvo-Garrido, S., Domínguez-Lara, S., Polanco-Levicán, K., & Mieres-Chacaltana, M. (2023). Psychometric properties of the teachers' sense of efficacy scale in a sample of Chilean public school teachers. *Frontiers in Psychology*, 14, 1272548. <https://doi.org/10.3389/fpsyg.2023.1272548>
- George, D., & Mallery, P. (2024). *IBM SPSS Statistics 29 Step by Step: A Simple Guide and Reference* (18th ed.). Routledge. <https://doi.org/10.4324/9781032622156>
- González-Rivas, R. A., Jurado-García, P. J., López-Trujillo, V. H., & Aguirrez, S. I. (2026). Cuestionario breve de conductas disruptivas en Educación Física: propiedades psicométricas. *Retos*, 78, 126-139. <https://doi.org/10.47197/retos.v78.118469>
- González-Rivas, R. A., Laguna Celia, A., & Nuñez Enriquez, O. (2023). Factores que influyen en la Educación Física en México. *Retos: Nuevas Perspectivas de Educación Física, Deporte y Recreación*, 48. <https://doi.org/10.47197/retos.v48.96752>
- González-Rivas, R. A., Núñez Enríquez, O., Zueck-Enríquez, M. d. C., Gastelum-Cuadras, G., Ramírez-García, A. A., López-Alonzo, S. J., & Guedea-Delgado, J. C. (2022). Analysis of the factors that influence a quality Physical Education in Mexico: school supervision's perspective. *International Journal of Environmental Research and Public Health*, 19(5), 2869. <https://doi.org/10.3390/ijerph19052869>
- Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2019). *Multivariate Data Analysis*. Cengage.
- Indu, P., Vidhukumar, K., Chacko, D., Menon, V., Grover, S., & Gupta, S. (2025). Criterion validity, construct validity, and factor analysis: An introductory overview. *Indian Journal of Psychiatry*, 67(9), 916-921. https://doi.org/10.4103/indianjpsychiatry_911_25

- Krumpal, I. (2013). Determinants of social desirability bias in sensitive surveys: a literature review. *Quality & quantity*, 47(4), 2025-2047. <https://doi.org/10.1007/s11135-011-9640-9>
- Lander, N. J., Hanna, L., Brown, H., Telford, A., Morgan, P. J., Salmon, J., & Barnett, L. M. (2017). Physical education teachers' perspectives and experiences when teaching FMS to early adolescent girls. *Journal of Teaching in Physical Education*, 36(1), 113-118. <https://doi.org/10.1123/jtpe.2015-0201>
- Lavelly, C., Berger, N., Blackman, J., Follman, J., & McCarthy, J. (1994). Contemporary teacher classroom performance observation instruments. *Education*, 114(4), 618-625. <https://go.gale.com/ps/i.do?id=GALE%7CA16138719&sid=googleScholar&v=2.1&it=r&linkaccess=abs&issn=00131172&sw=w&p=AONE&userGroupName=anon%7E8a5a4558&aty=open-web-entry>
- Leyton-Roman, M., Nunez, J. L., & Jimenez-Castuera, R. (2020). The importance of supporting student autonomy in physical education classes to improve intention to be physically active. *Sustainability*, 12(10), 4251. <https://doi.org/10.3390/su12104251>
- Lounsbery, M., & Coker, C. (2008). Developing skill-analysis competency in physical education teachers. *Quest*, 60(2), 255-267. <https://doi.org/10.1080/00336297.2008.10483580>
- Martínez-Molina, M., Granero-Gallegos, A., Baena-Extremuera, A., & Baños, R. (2020). Invarianza factorial por sexo del cuestionario para las conductas disruptivas y escala de evaluación de la competencia docente en educación física en estudiantes de secundaria. *Psychology, Society & Education*, 12(2), 125-137. <https://ojs.ual.es/ojs/index.php/psye/article/view/3172>
- McKown, H. B., Merica, C. B., & Egan, C. A. (2024). Preparing teachers for physical activity leadership: Status of comprehensive school physical activity program in physical education Teacher education. *Journal of Teaching in Physical Education*, 44(1), 154-162. <https://doi.org/10.1123/jtpe.2023-0271>
- Nunnally, J., & Bernstein, I. (1995). *Teoría psicométrica*. McGraw-Hill.
- Organización de las Naciones Unidas para la Educación la Ciencia y la Cultura. (2022). *Transformar la seneñanza desde adentro: tendencias actuales en la situación y desarrollo del personal docente*. UNESCO. https://unesdoc.unesco.org/ark:/48223/pf0000383002_spa
- Organización para la Cooperación y el Desarrollo Económicos. (2022). *Resultados PISA 2022 (Volumen I y II) - Notas de país: México*. https://www.oecd.org/en/publications/pisa-2022-results-volume-i-and-ii-country-notes_ed6fbcc5-en/mexico_519eaf88-en.html
- Putnick, D. L., & Bornstein, M. H. (2016). Measurement invariance conventions and reporting: The state of the art and future directions for psychological research. *Developmental review*, 41, 71-90. <https://doi.org/10.1016/j.dr.2016.06.004>

- Revelle, W., & Zinbarg, R. E. (2009). Coefficients alpha, beta, omega, and the glb: Comments on Sijtsma. *Psychometrika*, 74(1), 145-154. <https://doi.org/10.1007/s11336-008-9102-z>
- Rosas, S., & Velázquez, M. A. (2026). La evaluación docente en el sexenio 2018-2024: balance crítico de los modelos USICAMM y MEDFI. *CPU-e, Revista de Investigación Educativa*, (42). 125-154. <https://doi.org/10.25009/cpue.v1i42.2918>
- Sathyanarayana, S., & Mohanasundaram, T. (2024). Fit indices in structural equation modeling and confirmatory factor analysis: reporting guidelines. *Asian Journal of Economics, Business and Accounting*, 24(7), 561-577. <https://doi.org/10.9734/ajeba/2024/v24i71430>
- Secretaría de Educación Pública. (2016). *Resultados de la #EvaluaciónDelDesempeño, Ciclo Escolar 2015-2016*. <https://www.gob.mx/sep/articulos/resultados-de-la-evaluaciondeldesempeno-ciclo-escolar-2015-2016>
- Secretaría de Educación Pública, & Unidad del Sistema para la Carrera de las Maestras y los Maestros. (2023). *Acuerdo que contiene las disposiciones, criterios e indicadores para la realización del proceso de promoción a funciones directivas o de supervisión en educación básica, ciclo escolar 2023-2024*. http://usicamm.sep.gob.mx/usicamm_dsk08/2023-2024/compilacion/EB/Acuerdo_PromocionVertical_EB.pdf
- Secretaría de Educación Pública. (2024). *Perfiles profesionales, criterios e indicadores para el personal docente, técnico docente, de asesoría técnica pedagógica, directivo y de supervisión escolar*. Secretaría de Educación Pública. http://usicamm.sep.gob.mx/usicamm_dsk08/2025-2026/compilacion/EB/Marco_EB.pdf
- Serrano-Brazo, M., Sánchez, A. R. R., & Macías, G. R. (2023). Formación Permanente del Profesorado de Educación Física en Educación Secundaria Obligatoria: características y diseño (In-service training for Physical Education teachers in compulsory secondary education: characteristics and design). *Retos*, 50, 1272-1280. <https://doi.org/10.47197/retos.v50.99307>
- Sirlopú, D., Pérez-Salas, C., Villarroel, V., García, C., Barrera, K., Fierro, E., & Castillo, I. (2019). Adaptación y validación de la versión en español de la Escala de Respeto Incondicional hacia las personas en una muestra de escolares chilenos. *Universitas Psychologica*, 18(1), 1-13. <https://doi.org/10.11144/Javeriana.upsy18-1.avve>
- Tobón, S. T. (2007). El enfoque complejo de las competencias y el diseño curricular por ciclos propedéuticos. *Acción pedagógica*, 16(1), 14-28. <https://dialnet.unirioja.es/revista/13000/V/16?anualidad=2007>
- Tomás, J. M., Galiana, L., Hontangas, P., Oliver, A., & Sancho, P. (2013). Evidencia acumulada sobre los efectos de método asociados a ítems invertidos. *Psicológica*, 34(2), 365-381. <https://www.redalyc.org/pdf/169/16929535013.pdf>

- Voss, T., Kunter, M., & Baumert, J. (2011). Assessing teacher candidates' general pedagogical/psychological knowledge: Test construction and validation. *Journal of educational psychology*, 103(4), 952. <https://doi.org/10.1037/a0025125>
- Wei, Y. (2025). A study on the influence of physical education teachers' teaching beliefs and attributions of teaching success and failure on their effective teaching behaviors. *Plos one*, 20(4), e0321137. <https://doi.org/10.1371/journal.pone.0321137>
- Worthington, R. L., & Whittaker, T. A. (2006). Scale development research: A content analysis and recommendations for best practices. *The counseling psychologist*, 34(6), 806-838. <https://doi.org/10.1177/0011000006288127>
- Yin, H., Omar Dev, R. D., Soh, K. G., Li, F., & Lian, M. (2024). Assessment and development of physical education teachers' physical literacy: A systematic review. *Plos one*, 19(7), e0307505. <https://doi.org/10.1371/journal.pone.0307505>

Financiación

Los autores no recibieron financiación para el desarrollo de la presente investigación

Conflicto de intereses

Los autores declaran no tener conflicto de intereses

Contribución de autoría

Los autores han participado en la construcción del documento.