

Validación española de la Escala de Evaluación de la Competencia Docente en Educación Física de secundaria

Spanish version of the Evaluation of Teaching Competencies Scale in Physical Education of secondary school

Validação em espanhol da Evaluation of Teaching Competencies Scale em Educação Física no ensino secundário

Antonio Baena-Extremera¹, Antonio Granero-Gallegos² y Marina Martínez-Molina¹

¹Facultad de Ciencias del Deporte, Universidad de Murcia, España y ²Centro del Profesorado Cuevas-Olula, Consejería de Educación, Cultura y Deporte, Junta de Andalucía, España.

Resumen: El objetivo de esta investigación fue analizar las propiedades psicométricas del *Evaluation of Teaching competencies Scale* (ETCS) en una muestra española de Educación Física en educación secundaria. Se administró a un total de 758 estudiantes de secundaria de 12 a 18 años, analizándose la estructura del instrumento mediante procedimientos confirmatorios. Los análisis apoyan el modelo teórico hipotetizado de origen, presentando los ocho ítems medidas adecuadas de homogeneidad y la escala una aceptable consistencia interna (*fiabilidad compuesta* =,93; *varianza media extraída* =,61; *alfa de Cronbach* =,90) y estabilidad temporal. Los datos de ajuste del modelo fueron excelentes: $\chi^2=27,68$, $gl=20$, $p=,117$, $\chi^2/gl=1,38$, $GFI=,99$, $NFI=,98$, $NNFI=,99$, $CFI=,99$, $RMSEA=,023$; $RMSR=,029$; además, el modelo examinado cumple los requisitos de validez convergente (cargas factoriales >,60 y *t-value*>1,96). La validez externa se exploró examinando la relación predictiva de la escala estudiada con la *Sport Satisfaction Instrument* adaptada a la Educación Física. La ETCS-EF presenta una potente relación de predicción con la *satisfacción/diversión* del alumnado en clase, mientras que es negativa con el *aburrimiento*. Así, a mayor competencia del docente, menor probabilidad de que los estudiantes se aburran en las clases. Se concluye, por tanto, que la escala ETCS-EF supone un instrumento contrastado y que sirve para evaluar la competencia docente del profesorado actual de EF.

Palabras clave: Educación Física, profesor, competencias, propiedades psicométricas, ETCS.

Abstract: The aim of this research was to analyze the psychometric properties of *Evaluation of Teaching Competencies Scale* (ETCS) in a Spanish sample of Physical Education in secondary education. It was administered to a total of 758 high school students from 12 to 18 years, analyzing the structure of the instrument using confirmatory procedures. The analyzes support the hypothesized theoretical model of origin, the eight items presented adequate measures and scale homogeneity acceptable internal consistency (composite reliability =,93; average variance extracted =,61; Cronbach's alpha =,90) and temporal stability. The data were excellent model fit: $\chi^2=27.68$, $df=20$, $p=.117$, $\chi^2/df=1.38$, $GFI=.99$, $NFI=.98$, $NNFI=.99$, $CFI=.99$, $RMSEA=.023$; $RMSR=.029$; in addition,

the examined model meets the requirements of convergent validity (factor loadings > 60 and *t-value* > 1.96). External validity was explored by examining the predictive relationship of the scale studied and the Sport Satisfaction Instrument adapted to Physical Education. The ETCS-EF has a powerful predictive relationship satisfaction / fun of students in class while it is negative with boredom. Thus, the higher teacher competence, the less likely that students get bored in class. We conclude, therefore, that the ETCS-EF scale is a proven instrument used to assess teacher competence of the current PE teachers.

Keywords: Physical Education, teacher, competencies, psychometric properties, ETCS.

Resumo: O objetivo desta pesquisa foi analisar as propriedades psicométricas da *Evaluation of Teaching competencies Scale* (ETCS) numa amostra espanhola de educação física no ensino secundário. Foi administrada a um total de 758 estudantes do ensino secundário de 12 a 18 anos. A estrutura do instrumento foi analisado por procedimentos de confirmação. As análises confirmam o modelo teórico na hipótese de origem. Os oito itens apresentaram homogeneidade e medidas aceitáveis de consistência interna na escala (= *fiabilidade composta*, 93; *variância média extraída* = 61; *alfa de Cronbach* = 90) e estabilidade temporal. Os dados de ajuste foram excelentes: $c2 = 27,68$, $df = 20$, $p = , 117$, $c2 / gl = 1,38$, $GFI = 99$, $NFI = , 98$, $NNFI = 99$, $= CFI, 99$, $RMSEA = , 023$; $= RMSR, 029$; Além disso, o modelo examinado satisfaz os requisitos de validade convergente (cargas factoriais > 60 e *t-value* > 1,96). A validade externa foi explorada examinando a relação preditiva da escala de estudo co *Sport Satisfaction Instrument* adaptado para a educação física. O ETCS-EF tem um poderoso preditivo relacionamento satisfação /diversão dos alunos na aula, ao mesmo tempo que é negativo com respeito ao tédio. Assim, a maior competência do professor, menos provável que os alunos ficam chatos nas aulas. Concluímos, portanto, que a escala ETCS-EF constitui um instrumento comprovado para avaliar competência pedagógica dos professores atuais de educação física.

Palavras-chave: professor de educação física, habilidades, propriedades psicométricas, ETCS.

Dirección para correspondencia [Correspondence address]: Antonio Granero-Gallegos. Centro del Profesorado Cuevas-Olula, Consejería de Educación, Cultura y Deporte, Junta de Andalucía, España. Ctra de Águilas s/n – 04610 – Cuevas del Almanzora, Almería (España). E-mail: agranerog@gmail.com

Introducción

La enseñanza por competencias requiere nuevos planteamientos en el proceso de enseñanza-aprendizaje, tanto de innovación metodológica como de contenidos. La figura del docente adquiere nuevo protagonismo y más aún si se tienen en cuenta los resultados alcanzados en informes externos de carácter internacional, como el informe PISA. Como consecuencia de ello, la evaluación del docente y del requerimiento de ciertas competencias en la eficacia de su labor profesional comienzan a ser un interrogante que necesita ser abordado.

Blázquez (2013, p. 4) considera que las competencias docentes tienen que ver con las *características específicas de los desempeños docentes asociados a las funciones y responsabilidades propios del quehacer cotidiano (lo genérico), pero sobre todo articuladas estrechamente a una práctica docente concreta, esto es, asociadas al quehacer específico y auténtico de una práctica docente peculiar*. Autores como Denyer, Furnemont, Poulain y Vanloubbeeck (2007), consideran que desglosar las competencias docentes permitiría evaluar aspectos específicos y observables de la actividad docente, ayudando así a discriminar entre el buen y el mal docente. Además, según el Modelo de Evaluación de Competencias Docentes para la Educación Media y Superior (García-Cabrero, Loredó, Luna, y Rueda, 2008), la evaluación de estos aspectos específicos y observables podría centrarse en tres momentos, considerados como secuenciales y paralelos para el desarrollo correcto de la función docente, como son: la previsión del proceso enseñanza-aprendizaje, la conducción del proceso enseñanza-aprendizaje y la valoración de proceso enseñanza-aprendizaje (Reyes, Serrano, y Caso, 2014). Estos momentos abarcan, a su vez, actuaciones del educador, que sumado a sus facetas particulares, pueden desencadenar una exitosa o fracasada intervención didáctica.

Un ejemplo claro de evaluación de competencias docentes se puede apreciar cada año en los procesos selectivos de maestros y profesores. Para acceder al cuerpo de funcionarios docentes del estado español es necesario superar un proceso de concurso-oposición en el que el candidato debe mostrar todas sus facetas como buen docente: buen comunicador, buen planificador y diseñador didáctico, creativo, capaz de atender y proporcionar información al alumno, poseer cierta capacidad de resolución de problemas, gran conocedor de la materia, profesionalidad, etc. (ver por ejemplo, la Orden de 23 de marzo de 2015, por la que se efectúa convocatoria de procedimiento selectivo para el ingreso en el Cuerpo de Maestros en Andalucía). Estas y otras características son las idóneas que deberían estar presentes durante todos y cada uno de los días de actividad docente de nuestros maestros y profesores. No obstante, muchas de estas particularidades con el tiempo se pierden, o bien, nunca se tuvieron afianzadas; por ello, es necesario conocer continuamente el estado

actual de los docentes, por la repercusión que en el sistema educativo puede tener.

La idea de un docente eficaz para cada individuo puede ser variable. Algunos estudiantes entienden que un maestro efectivo es a menudo una persona perfeccionista, alentadora, atenta y cariñosa; otras veces ven más una persona inteligente pero, sobre todo, entusiasta, divertida, inteligente, afectiva y comprensiva, abierta y con un estilo relajado, mientras enseña (Moreno, 2009). En síntesis, el docente debe reunir un conjunto de rasgos en su personalidad, y ciertas habilidades y conocimientos específicos relacionados correctamente con la enseñanza en clase (Keeley, Christopher, y Buskist, 2012). En este caso los estudiantes tienen un gran papel en el mantenimiento de la calidad y en la mejora del aprendizaje gracias a su participación en los procesos de evaluación que garantizan la calidad interna de la enseñanza (Lidice y Saglam, 2012). Estos procesos de evaluación de los profesores y de su tarea docente sirven como retroalimentación a los estudiantes, docentes, departamentos universitarios, administradores, políticos, gobiernos e investigadores (Kuzmanovic, Savic, Popovic, y Martic, 2012, 2013), ayudando a ajustar y mejorar la práctica educativa. La mayoría de estas investigaciones se basan en el enfoque tradicional donde los estudiantes evalúan la enseñanza recibida durante un curso realizando dicha evaluación al final del mismo (Marsh, Ginns, Morin, Nagengast, y Martin, 2011).

Una de las principales deficiencias de estas evaluaciones realizadas por los alumnos es que no se sabe si los elementos de las encuestas utilizadas representan adecuadamente los constructos subyacentes para los que fueron desarrollados. De hecho, Catano y Harvey (2011), Marsh (2007), Simpson y Sigua (2000) muestran su preocupación en el uso de cuestionarios para evaluar al docente por los alumnos, debido a la falta de fiabilidad, validez y sesgo potencial.

Recientemente, Luna, Calderón, Caso y Cordero (2012), validaron el Cuestionario de Evaluación de la Competencia Docente. Estos autores distinguieron cuatro factores (entendidos como las competencias docentes) a tener en cuenta en dicho instrumento: planeación y gestión del proceso enseñanza-aprendizaje, interacción didáctica en el aula, evaluación y comunicación del proceso de enseñanza-aprendizaje, y tecnologías de la información y la comunicación.

Por otro lado, Catano y Harvey (2011) validaron con alumnado universitario (18 a 60 años) la *Evaluation of Teaching Competencies Scale* (ETCS) con idea de crear un instrumento capaz de evaluar las competencias del profesor, demostrando en su trabajo la fiabilidad de dicho instrumento. En dicha escala, tras diversos estudios, se establecieron los requisitos que el profesor debe cumplir para ser competente: buena comunicación, conciencia de trabajo, creatividad, feed-back, consideración individual, profesionalidad, resolución de problemas y conciencia social.

A diferencia del campo universitario, donde se busca más la excelencia educativa y existen incluso vicedecanatos de calidad educativa, en educación secundaria el profesorado no es comúnmente evaluado por el alumnado mediante encuestas que garanticen la calidad interna del proceso de enseñanza-aprendizaje y menos aún en Educación Física (EF), donde el proceso educativo es diferente al resto de áreas de aula. Sumado a esto, hay que destacar el papel tan importante que tiene esta área dentro del currículum y en la formación del alumno, sobre todo en la promoción de práctica de actividad física y saludable más allá de las clases, entre los adolescentes (Stratton, Fairclough, y Ridgers, 2008). En este caso, el papel del profesorado es fundamental y cada vez más estudiado y tenido en cuenta, pues hay diversas variables que puede manejar y que tienen gran relevancia en el alumnado para adoptar un estilo de vida físicamente activo fuera de la escuela: feedback, competencia percibida, selección de contenidos apropiados y adaptados, disfrute de sus experiencias (Wallhead y Buckworth, 2004).

Por tanto, hemos de tener en cuenta que la figura del profesor y su proceso de enseñanza constituyen un elemento clave en el sistema educativo, sobre el que se está investigando en menor medida. Por todo ello, se cree necesario realizar la adaptación de la ETCS a la EF para evaluar la competencia del profesorado de secundaria. El objetivo de este trabajo es aportar evidencias sobre la dimensionalidad de la versión española del ETCS adaptada a la EF en una muestra de estudiantes adolescentes mediante procedimientos confirmatorios.

Método

Participantes

En este trabajo participaron 758 alumnos (347 hombres=45,8%; 411 mujeres=54,2%) correspondientes a siete centros de educación secundaria de la Región de Murcia. El rango de edad estuvo comprendido entre 12 y 18 años ($M=15,22$; $SD=1,27$), siendo la edad media de los chicos 15,2 ($SD=1,29$) y la de las chicas 15,18 ($DT=1,26$). La selección de la muestra fue de tipo no probabilístico y por conveniencia, según los estudiantes a los que se pudo acceder.

Proceso de adaptación del instrumento al español

La adaptación española del ETCS se realizó atendiendo a los estándares metodológicos internacionales recomendados por la International Test Commission para adaptar correctamente test y escalas de unas culturas a otras (Hambleton, 2005; Muñiz y Bartram, 2007). Para evitar imprecisiones, se combinaron los diseños de traducción directa e inversa de los ítems, según el procedimiento *parallel back translation* (Brislin, 1986).

La evaluación cualitativa de ítems (*validez de contenido*) se efectuó mediante el juicio de cuatro expertos (Osterlind, 1989): dos en construcción de escalas y dos conocedores del constructo a evaluar. Todos los ítems fueron analizados y revisados para que se recogiese la dimensión teórica de la forma más clara y precisa. Así, tras la adaptación de los originales, se les mostró el listado de ítems para que emitieran un juicio sobre su pertinencia y su comprensión en escala de 1 (*muy en desacuerdo*) a 4 (*muy de acuerdo*). Asimismo, dispusieron de un apartado en el que hacer las anotaciones y observaciones generales sobre cada uno de los ítems, sugiriendo una redacción alternativa de cada ítem si lo consideraban conveniente. Se revisaron los ítems con puntuaciones medias <2,5, tanto en pertinencia como en comprensión.

Se ha de destacar en este punto que todos los expertos estuvieron de acuerdo en suprimir la cuestión 1 de la escala original y, por tanto, la correspondiente a su versión al español. Al tratarse de un instrumento validado en su versión original en el contexto universitario, la redacción del ítem de *disponibilidad del profesor (Availability)* podía llevar a confusión entre el alumnado de secundaria, dado que en este nivel educativo el profesorado no dispone de horario expreso de atención individual al alumnado, como sí ocurre en la universidad.

Finalmente, indicar que la concordancia global de los cuatro expertos sobre la pertinencia y comprensión de los ítems se midió mediante el Coeficiente de Correlación Intraclass (CCI), a partir de un modelo de efectos mixtos y asumiendo una definición de acuerdo absoluto; los valores obtenidos fueron: $CCI=.84$ en pertinencia y $CCI=.90$ en comprensión.

Una vez adaptada y preparada la escala, el encabezamiento fue: *Respecto a tu profesor de Educación Física...* y esta versión fue administrada a 50 alumnos de secundaria. Los comentarios de los mismos sobre instrucciones y forma de redacción supusieron cambios menores. Tras el análisis de los resultados psicométricos obtenidos y una última revisión por parte del equipo de investigación se llegó a la versión final española del ETCS-EF (Anexo 1).

Instrumentos

Se adaptó la versión original del *Evaluation of Teaching Competencies Scale* (ETCS) de Catano y Harvey (2011). Esta adaptación al contexto español de educación secundaria consta de ocho ítems para medir la percepción de los estudiantes de la eficacia de la docencia; en este caso se adaptó para medir la competencia docente del profesor de EF. En las instrucciones se pidió al alumnado que indicara el grado de acuerdo con los ítems, recogiendo las respuestas en una escala de ítems politómicos de siete puntos que oscila entre *Bajo* (1, 2), *Medio* (3, 4, 5) y *Alto* (6, 7). Los ocho constructos se explican seguidamente:

1. *Comunicación*: capacidad de elocuencia y flexibilidad verbal y escrita para comunicarse con claridad, precisión y propósito con el alumnado, así como dedicar el tiempo necesario a escucharlos.
2. *Conciencia de trabajo*: capacidad para producir un trabajo actualizado y de alta calidad poniendo estricta atención en los detalles; percibir aspectos que otros profesores pasan por alto y rechazar cualquiera que no cumpla con el criterio de la excelencia.
3. *Creatividad*: capacidad para promover ideas innovadoras en clase y nuevos procesos metodológicos, así como ser flexible en las actitudes personales, reglas o modos de funcionamiento.
4. *Feed-back*: capacidad para proporcionar al alumnado una información específica y detallada de los puntos importantes de su conducta y actuación, tanto verbalmente como por escrito; proponer medidas de corrección para la mejora cuando sea necesario.
5. *Consideración individual al alumnado*: capacidad de mostrar sensibilidad, empatía y atender a las necesidades del alumnado; proporcionar apoyo y ayuda individual cuando sea necesario.
6. *Profesionalidad*: capacidad para demostrar honestidad e integridad siendo consecuente entre sus palabras y sus acciones; mostrar sentido de equidad y justicia y mantener la confidencialidad de la información recibida del alumnado.
7. *Resolución de problemas*: capacidad para analizar los problemas, buscar soluciones eficaces y decidir sobre las medidas correctivas que se deben adoptar.
8. *Conciencia social*: capacidad para estar en sintonía con los pensamientos y sentimientos del alumnado y ser discreto, manteniendo la disciplina de clase.

Finalmente, se utilizó la versión española del *Sport Satisfaction Instrument* (SSI) adaptada a la EF (SSI-EF) por Baena-Extremera, Granero-Gallegos, Bracho-Amador, y Pérez-Quero (2012). El SSI-EF consta de ocho ítems para medir la satisfacción en una actividad deportiva mediante dos subescalas que miden *satisfacción/diversión* (cinco ítems) y *aburrimiento* (tres ítems). En las instrucciones se pide a los sujetos que indiquen su grado de acuerdo con los ítems que reflejan criterios de diversión o aburrimiento, recogiendo las respuestas en una escala de ítems politómicos de cinco puntos que oscila desde *Muy en desacuerdo* (1) a *Muy de acuerdo* (5). La consistencia interna hallada fue: *satisfacción/diversión*, $\alpha=,77$ y *aburrimiento*, $\alpha=,71$.

Procedimiento

Se obtuvieron los permisos de los órganos competentes, tanto de los centros de secundaria como universitarios, para el de-

sarrollo de este trabajo de investigación. En primer lugar se habló con los equipos directivos de los centros y se concertó una reunión previa con los docentes de cada centro en la que se les explicó el objetivo del estudio. Asimismo, los padres/madres/tutores y estudiantes fueron pertinentemente informados acerca del protocolo y objeto del estudio. La firma del consentimiento informado fue requisito indispensable para participar. Igualmente, los estudiantes fueron informados del propósito del estudio, voluntariedad y confidencialidad de las respuestas y manejo de datos y que no había respuestas correctas o incorrectas. Cada participante dispuso de 10-20 minutos para completar el cuestionario que fue aplicado en el aula por los investigadores y sin la presencia del docente de EF.

Para evaluar la estabilidad temporal del ETCS se utilizó una segunda muestra de 150 alumnos, elegida previamente de forma aleatoria y compuesta por 73 chicos ($M=15,02$; $SD=1,37$) y 77 chicas ($M=14,95$; $SD=1,42$) que completaron nuevamente el instrumento siete semanas más tarde.

Análisis estadísticos

Los análisis de ítems, homogeneidad, correlación, consistencia interna de la escala y análisis y regresión lineal se realizaron con SPSS 20.0. Inicialmente se realizó un análisis de cada ítem de la escala para estudiar la conveniencia de conservarlo dentro del factor. Para el análisis de consistencia interna también se hallaron la *fiabilidad compuesta* y la *varianza media extraída*. Además, se calculó el coeficiente de Mardia-Based-Kappa para estimar la normalidad multivariante y determinar así el método más adecuado de estimación con LISREL. Así, se evaluó la estructura factorial del ETCS mediante análisis factorial confirmatorio (AFC) con LISREL 8.80 (Jöreskog y Sörbom, 2003), y se utilizó el método de estimación *weighted least squares* (WLS); se calcularon índices de ajuste absolutos y relativos para la evaluación del modelo.

Resultados

Análisis de ítems de la escala

Se ha seguido un procedimiento de análisis según lo establecido por Carretero-Dios y Pérez (2007). En el análisis de ítems se mantuvo la distribución ítem-factor del instrumento original (Catano y Harvey, 2011). En el estudio de los ítems se analizó si la consistencia interna de la escala aumentaba con la eliminación de algún ítem, así como los requisitos establecidos por Nunnally y Bernstein (1995) para conservar un ítem dentro de un factor: coeficiente de correlación corregido ítem-total ($CCIT-c$) $\geq ,30$, desviación típica (SD) $> 1,0$, y todas las opciones de respuesta usadas en algún momento. Como recomiendan Bollen y Long (1994), los índices de asimetría

y curtosis son próximos a 0 y <2. (Tabla 1). Los ítems del ETCS en EF presentaron valores medios entre 5,73 (ítem-1) y 4,95 (ítem-3). Las *SD* fueron >1. La consistencia interna fue

adecuada ($\alpha=.90$) y no mejora con la eliminación de ningún ítem. Todos los *CCIT-c* presentaron valores $\geq .60$.

Tabla 1. Estadísticos descriptivos, de consistencia interna y de homogeneidad ($N=758$).

Escala ETCS ($\alpha=.90$)	<i>M</i>	<i>SD</i>	<i>CCIT-c</i>	α sin ítem	Asimetría	Curtosis
1 Comunicación	5,73	1,32	,70	,88	-1,05	,57
2 Conciencia de trabajo	5,72	1,39	,64	,89	-1,39	1,76
3 Creatividad	4,95	1,84	,66	,89	-,69	-,73
4 Feed-back	5,33	1,43	,68	,89	-,75	-,02
5 Consideración individual al alumno/a	5,22	1,59	,75	,88	-,90	,12
6 Profesionalidad	5,34	1,47	,63	,89	-,90	,39
7 Resolución de problemas	5,29	1,51	,74	,88	-,85	,13
8 Conciencia social	5,33	1,50	,70	,88	-,92	,32

Nota. *N*: muestra; *M*: media; *SD*: desviación típica; *CCIT-c*: coeficiente de correlación corregido ítem-total

Análisis factorial confirmatorio

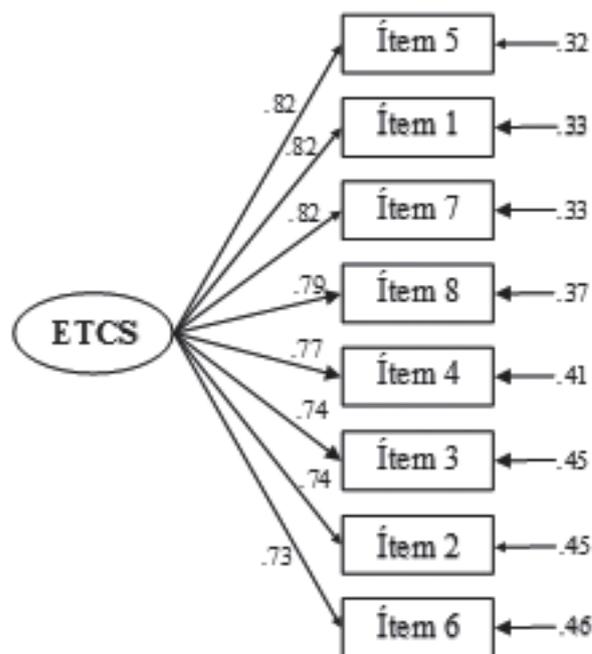
Para estudiar las propiedades psicométricas de la dimensionalización original propuesta teóricamente por Catano y Harvey (2011) se aplicaron modelos de ecuaciones estructurales. La estructura factorial del ETCS se evaluó con AFC.

Previo a este análisis factorial, se efectuó un análisis de la normalidad multivariante de la escala ETCS. Se realizó el *test de normalidad basado en la curtosis multivariante relativa* (RMK) de PRELIS, del programa LISREL 8.80. El valor de la *curtosis normalizada multivariante* del ETCS fue: 44,489 (*Mardia-Based-Kappa* =,511), siendo 1,96 (5%) el valor crítico considerado del test. Los resultados mostraron que no se puede aceptar la normalidad multivariante, lo que implica la utilización de estimadores robustos. Así, se utilizó el método de estimación *weighted least squares* (WLS) del programa LISREL 8.80 de Jöreskog y Sörbom (2003). La matriz de correlaciones policóricas y la matriz de covarianzas asintóticas fueron utilizadas como input para el análisis de los datos. Se hipotetizó un modelo de medida de un solo factor.

Se calcularon varios índices de ajuste para la evaluación del modelo, combinando índices absolutos y relativos. Entre los absolutos, se utilizó el valor *p*, asociado con el estadístico chi-cuadrado (χ^2) que prueba el modelo nulo frente al modelo hipotetizado (Barrett, 2007); la *ratio* entre χ^2 y grados de libertad (*gl*) (χ^2/gl), considerándose las <2,0 como indicadores de muy buen ajuste del modelo (Tabachnik y Fidell, 2007). Se ha calculado el *GFI* (índice de bondad de ajuste) y autores como Hooper, Coughlan, y Mullen (2008) consideran valores $\geq .95$ para un mejor ajuste. Índices relativos: *NFI* (índice de ajuste normalizado), *NNFI* (índice de ajuste no normativo) y *CFI* (índice de ajuste comparativo); los valores $\geq .95$ indican un buen ajuste (Hu y Bentler, 1999). Estos últimos autores consideran que un valor $\leq .06$ para el *RMSEA* (*error de aproxi-*

mación cuadrático medio) indicaría un buen ajuste, así como valores $\leq .08$ para el *RMSR* (*raíz del promedio de los residuales estandarizados*). Los parámetros estimados se consideran significativos cuando el valor asociado al valor *t* >1,96 ($p < .05$).

Figura 1. Path Diagram del AFC, con pesos estandarizados y errores de medición de cada uno de los ítems del ETCS en EF de secundaria.



El modelo puesto a prueba presenta unas altas cargas factoriales en todos los reactivos ($\geq .73$). Todos los ítems presentaron valores $> .50$ en fiabilidad individual (R^2) (entre ,55 del

ítem-2 y ,68 del ítem-5). Los datos de ajuste fueron excelentes: $\chi^2=27,68$, $gl=20$, $p=,117$, $\chi^2/gl=1,38$, $GFI=,99$, $NFI=,98$, $NNFI=,99$, $CFI=,99$, $RMSEA=,023$; $RMSR=,029$). El modelo cumple los requisitos que garantizan la validez convergente del modelo (Hair, Black, Babin, y Anderson, 2009): todas las cargas factoriales $>,60$ y estadísticamente significativas ($t\text{-value}>1,96$). Incluso, el ETCS cumple con los requisitos más exigentes expuestos por Barrett (2007) y que hacen referencia a la prueba χ^2 para aceptar un modelo evaluado, dado que en este caso la prueba no resultó significativa ($p=,117$), como requiere el citado autor.

También es importante en el AFC de las escalas con naturaleza ordinal de la matriz de correlaciones de datos, ofrecer los resultados de *fiabilidad compuesta* y *varianza media extraída* (AVE - Average Variance Extracted) para cada una de las dimensiones críticas. Según Hair et al. (2009) la *fiabilidad compuesta* debe tener un valor mínimo de ,70 y la AVE de ,50. El modelo evaluado presentó una *fiabilidad compuesta* de ,93 y una AVE de ,61. En la evaluación de la estabilidad temporal los resultados del pretest fueron: $\alpha=,89$ y los del postest: $\alpha=,90$. Asimismo, en el CCI realizado a partir de un modelo de efectos mixtos y asumiendo una definición de acuerdo absoluto e intervalo de confianza del 95%, los valores obtenidos fueron: CCI=,87 en pretest y CCI=,89 en postest.

Respecto al otro instrumento utilizado (SSI-EF), se realizó un análisis factorial confirmatorio (AFC) según la estructura original (Baena-Extremera et al., 2012) utilizando el método de estimación WLS para variables ordinales del programa LISREL 8.80. El modelo presentó un aceptable ajuste: $\chi^2=49,86$, $gl=19$, $p<,001$, $\chi^2/gl=2,62$, $GFI=,95$, $NFI=,98$, $NNFI=,99$, $CFI=,99$, $RMSEA=,05$.

Validez externa

Se ha realizado un análisis de regresión para comprobar en qué medida el ETCS-EF (variable independiente) predice las dimensiones del SSI-EF (variable dependiente). Se comprueba que el ETCS presenta una potente relación de predicción con la *satisfacción/diversión* del alumno en clase ($F=195,65$; $\beta=,45$; $p<,0001$), con un 30,52% de la varianza total explicada. La relación es negativa con el *aburrimiento* ($F=78,28$; $\beta=-,31$; $p<,0001$): a mayor competencia docente del profesor menor probabilidad de que el alumnado se aburra en EF (20,30% de la varianza total explicada).

Discusión

El objetivo del presente trabajo consistió en examinar las propiedades psicométricas de la versión española del ETCS adaptado a EF en secundaria. La estructura factorial del ETCS fue investigada inicialmente por Catano y Harvey (2011), obteniendo una varianza explicada del 68%. En este trabajo, el

AFC basado en modelos de ecuaciones estructurales apoya la validez factorial y la fiabilidad del instrumento adaptado en una dimensión (percepción de los estudiantes de la eficacia de la docencia), similar al del modelo hipotetizado de origen.

Al igual que el ETCS en su versión original, en esta investigación la consistencia interna es adecuada, no mejorando con la eliminación de ningún ítem. Además, el resto de indicadores proponen dejar todos los ítems analizados, como el *CCIT-c*, con valores ,64 a ,75, aceptable para Nunnally y Bernstein (1995). Otro aspecto a destacar son las altas cargas factoriales, lo que indican la pertenencia de cada ítem al factor, así como los valores de la fiabilidad individual, todos ellos por encima del mínimo exigido.

Los datos de ajuste obtenidos en el AFC son excelentes; pero además, es destacable el cumplimiento de esta escala, incluso, de los indicadores más exigentes según Barret (2007), al no resultar significativo la prueba de χ^2 . Además, se han utilizado otros indicadores, como el *coeficiente de fiabilidad compuesta*, la *AVE* o la *estabilidad temporal*; todos ellos dan muestra de la fiabilidad y validez de esta escala con ocho ítems. Hay que destacar que Catano y Harvey (2011) no aportaron en su momento estos estadísticos en la versión original.

En el análisis de validez externa se utilizó la escala SSI-EF, debido a las posibles relaciones entre ambos constructos y a las aplicaciones prácticas que puede surgir de tal relación, aspectos que ya han manifestado otros investigadores, como se comprobará más adelante. En los resultados se observa que la ETCS-EF se establece como un buen predictor de la *satisfacción/diversión* del alumno en EF, y negativa con el *aburrimiento*. Esto hace entender que cuanto mayor es la competencia del profesor en su enseñanza, mayor será la *satisfacción/diversión* y menor el *aburrimiento* entre el alumnado. Esto pone de relevancia el estudio del papel y competencia del docente de EF de cara al fomento de hábitos de actividad física fuera de las aulas, en el que el disfrute del alumnado tiene un papel importante, junto al feed-back, contenidos, etc., (Wallhead y Buckworth, 2004), variables que incluye el ETCS-EF.

En relación a estos datos, Elliot y Shin (2002) identificaron algunos factores que predijeron la satisfacción del estudiante relacionada con el curso y la calidad de la enseñanza, entre los que destaca la excelencia y la calidad de la instrucción, el tener un docente profesional, que el profesor sea justo e imparcial, entre otras. Similares resultados se han puesto de manifiesto en el trabajo de Ginns, Prosser, y Barrie (2007), quienes hablan de buena enseñanza, tener metas claras, realizar una evaluación apropiada, plantear una carga de trabajo adecuada, y tener habilidades generales para la docencia. Kuzmanovic et al. (2013) hallaron en su estudio que los alumnos se encuentran satisfechos cuando el profesor está disponible para resolver sus problemas, ofreciendo por tanto cierta consideración individual hacia los estudiantes, dándoles feed-back, etc. Como se observa, muchos de ellos son indicadores utilizados

en la ETCS-EF. A pesar de todo, no existe una regla general aceptada por la comunidad científica acerca de las competencias que debe reunir un buen docente (Clayson, 2009), pero sí es cierto que este instrumento contribuye a evaluar algunas de ellas, reconocidas en las investigaciones mencionadas.

Se podría considerar como una limitación de este estudio el resto de competencias que el instrumento estudiado no permite evaluar, si bien, este trabajo se limita a la adaptación del instrumento original y no son, por tanto, achacables a este estudio; bien es cierto que en futuros trabajos se podría complementar dicho instrumento. Por contra, se aporta una nueva escala, a las pocas existentes en este ámbito, al contexto educativo español, válida, fácil de utilizar y que puede abrir nuevas vías de investigación en EF sobre la evaluación de la enseñanza.

Para concluir, se puede afirmar que la escala ETCS-EF cumple con los requisitos exigidos para su validación, siguiendo los factores hipotetizados según la teoría de origen. Los resultados obtenidos avalan la utilización del ETCS-EF, ya que resultan coherentes con las escasas investigaciones existentes tanto en el ámbito académico universitario como de secundaria y con las relaciones lógicas de otras escalas como la SSI-EF.

Aplicaciones prácticas

Es de destacar la importancia de esta escala y su utilidad para los docentes, tanto en su formación inicial, como en su práctica docente diaria, así como la repercusión de éstos en

los alumnos. La figura del profesorado está siendo revisada continuamente, a la vista de ciertos resultados de evaluaciones externas, tales como el informe PISA, y se pide, a su vez, que sea también una figura importante en temáticas como la mediación, la resolución de conflictos, etc. En la actualidad se pretende que la figura del docente sea dinámica, eficaz, competente, capaz de aunar e integrar una gran capacidad de comunicación personal con el uso adecuado de las nuevas tecnologías en el aula, que se actualice tanto en el campo científico como metodológico para hacer frente a los retos de la escuela del siglo XXI, capaz de pasar de un modelo transmisivo a un modelo más actual y colaborativo. Asimismo, debe ser una figura sensible y con gran capacidad empática, no directivo pero sí firme en sus decisiones, con gran capacidad de iniciativa y de resolución de problemas.

Uno de los aspectos que se destacan en la actualidad es la ausencia de evaluaciones en bastantes comunidades autónomas, sobre la actuación del profesorado tanto de primaria como de secundaria. Evaluaciones que puedan servir para tomar decisiones a la hora de la formación permanente de los docentes. La escala ETCS-EF supone un instrumento contrastado y que sirve para evaluar la competencia docente del profesorado actual de EF, contribuyendo así a solventar un importante vacío actual. La validación y uso de esta escala ETCS-EF en el contexto español debe constituir la base para su adaptación a otras áreas y niveles, y así poder evaluar la competencia docente del profesorado de distintas materias.

Referencias

- Baena-Extremera, A., Granero-Gallegos, A., Bracho-Amador, C., y Pérez-Quero, F. J. (2012). Versión española del Sport Satisfaction Instrument (SSI) adaptado a la Educación Física. *Revista de Psicodidáctica*, 17(2), 377-395.
- Barret, P. (2007). Structural equation modelling: Adjudging model fit. *Personality and Individual Differences*, 42, 815-824.
- Blázquez, D. (2013). *Diez competencias docentes para ser mayor profesor de Educación Física*. Barcelona: Inde.
- Bollen, K. A. y Long, J. (1994). *Testing structural equation models*. Newbury Park, CA: Sage.
- Brislin, R. W. (1986). The wording and translation of research instruments. En W. Lonner y J. Berry (Eds.), *Field methods in cross-cultural research* (pp. 137-164). Beverly Hills, CA: Sage.
- Carretero-Dios, H. y Pérez, C. (2007). Standards for the development and the review of instrumental studies: Considerations about test selection in psychological research. *International Journal of Clinical and Health Psychology*, 7, 863-882.
- Catano, V. M. y Harvey, S. (2011). Student perception of teaching effectiveness: development and validation of the Evaluation of Teaching Competencies Scale (ETCS). *Assessment y Evaluation in Higher Education*, 36(6), 701-717.
- Clayson, D. E. (2009). Student evaluations of teaching: Are they related to what students learn? A metaanalysis and review of the literature. *Journal of Marketing Education*, 31, 16-30.
- Denyer, M., Furnemont, J., Poulain, R. y Vanloubbeek, G. (2007). *Las competencias en la educación, un balance*. México: Fondo de Cultura Económica.
- Elliot, K. y Shin, D. (2002). Student satisfaction: An alternative approach to assessing this important concept. *Journal of Higher Education Policy and Management*, 24(2), 197-209.
- Ferguson, P. y Womack, S. T. (1993). The impact of subject matter and education coursework on teaching performance. *Journal of Teacher Education*, 44(1), 55-63.
- García-Cabrero, B., Loredó, J., y Carranza, G. (2008). Análisis de la práctica educativa de los docentes: pensamiento, interacción y reflexión. *Revista Electrónica de Investigación Educativa*, 10, 1-15.
- Ginns, P., Prosser, M., y Barrie, S. (2007). Students' perceptions of teaching quality in higher education: The perspective of currently enrolled students. *Studies in Higher Education*, 32(5), 603-615.
- Hair, J. F., Black, W. C., Babin, B. J., y Anderson, R. E. (2009). *Multivariate Data Analysis* (7th ed.). New York: Pearson Prentice Hall.
- Hambleton, R. K. (2005). Issues, designs and technical guidelines for adapting tests into multiple languages and cultures. En R. K. Hambleton, P. F. Merenday, S. D. Spielberger (Eds.), *Adapting educational and psychological tests for cross-cultural assessment* (pp. 3-38). Mahwah, NJ: Lawrence Erlbaum Associates.
- Hooper, D., Coughlan, J., y Mullen, M. (2008). Structural Equation Modelling: Guidelines for Determining Model Fit. *Electronic Journal of Business Research Methods*, 6(1), 53-60.
- Hu, L. y Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modelling*, 6, 1-55.
- Jöreskog, K. G., y Sörbom, D. (2003). *Structural equation modeling with the SIMPLIS command language*. Chicago: Scientific Software International.

19. Keeley, J., Christopher, A. N. y Buskist, W. (2012). Emerging Evidence for Excellent Teaching Across Borders. In J. Groccia, M. Alsudairi, y W. Buskist (Eds.), *Handbook of College and University Teaching: A Global Perspective*. (pp. 374-391). Thousand Oaks, CA: SAGE Publications.
20. Kuzmanovic, M., Savic, G., Popovic, M., y Martic, M. (2012). A New Approach to Evaluation of University Teaching Considering Heterogeneity of Students' Preferences. *Procedia-Social and Behavioral Sciences*, 64, 402-411.
21. Kuzmanovic, M., Savic, G., Popovic, M., y Martic, M. (2013). A New Approach to evaluation of university teaching considering heterogeneity of students' preferences. *High Education*, 66, 153-171.
22. Ley Orgánica 2/2006, de 3 de mayo de Educación (BOE, nº 106, del 4 de Mayo de 2006).
23. Lidice, A. y Saglam, G. (2012). Using student evaluations to measure educational quality. *Procedia-Social and Behavioral Sciences*, 70, 1009-1015.
24. Luna, E., Calderón, N., Caso, J. y Cordero, G. (2012). Desarrollo y validación de un cuestionario de evaluación de la competencia docente con base en la opinión de los estudiantes. En E. J. Cisneros-Cohernour, B. García-Cabrero, E. Luna y R. Marín (coords.), *Evaluación de Competencias Docentes en la Educación Superior* (pp.119-158). México: Juan Pablos Editor.
25. Marsh, H. W. (2007). Students' evaluations of university teaching: Dimensionality, reliability, validity, potential biases and usefulness. In R. P. Perry y J. C. Smart (Eds.), *The scholarship of teaching and learning in higher education: An evidence-based perspective* (pp. 319-384). New York: Springer.
26. Marsh, H. W., Ginns, P., Morin, A. J. S., Nagengast, B., y Martin, A. J. (2011). Use of Student Ratings to Benchmark Universities: Multilevel Modeling of Responses to the Australian Course Experience Questionnaire (CEQ). *Journal of Educational Psychology*, 103(3), 733-748.
27. Moreno, C. (2009). Effective teachers-Professional and personal skills. *ENSAYOS, Revista de la Facultad de Educación de Albacete*, 24. Recuperado de: <http://www.uclm.es/ab/educacion/ensayos>.
28. Muñiz, J. y Bartram, D. (2007). Improving international tests and testing. *European Psychologist*, 12, 206-219.
29. Nunnally, J. C. y Bernstein, I. J. (1995). *Teoría psicométrica*. Madrid: Mcgraw-Hill.
30. Osterlind, S. J. (1989). *Constructing Test Items*. Londres: Kluwer Academic Publishers.
31. Orden de 23 de marzo de 2015, por la que se efectúa convocatoria de procedimiento selectivo para el ingreso en el Cuerpo de Maestros (B.O.J.A. nº 62 de 31 de marzo de 2015).
32. Reyes, E. P., Serrano, E. L., y Caso, J. (2014). Propiedades psicométricas del Cuestionario de Evaluación de la Competencia Docente. *Revista Iberoamericana de Evaluación Educativa*, 7(2), 117-130.
33. Simpson, P. M. y Siguaw, J. A. (2000). Student evaluations of teaching: An exploratory study of the faculty response. *Journal of Marketing Education*, 22, 199-214.
34. Stratton, G., Fairclough, S. J., y Ridgers, N. D. (2008). Physical activity levels during the school day. In, A. L. Smith, y S. J. H. Biddle (eds), *Youth Physical Activity and Sedentary Behavior. Challenges and Solutions* (pp. 321-350). Champaign, IL: Human Kinetics.
35. Tabachnick, B. G. y Fidell, L. S. (2007). *Using Multivariate Statistics* (5th ed.). New York: Allyn and Bacon.
36. Wallhead, T. y Buckworth, J. (2004). The role of physical education in the promotion of youth physical activity. *Quest*, 56, 285-301.

Anexo. Versión española de la adaptación a la EF de secundaria de la escala ETCS

1 - COMUNICACIÓN

BAJO		MEDIO			ALTO	
1	2	3	4	5	6	7
El profesor/a de EF tiene un tono monótono de voz, cuando explica depende de las notas escritas y tiene dificultades para responder preguntas de los alumnos.		A veces, las instrucciones, las explicaciones y la asignación de tareas son más o menos claras.			El profesor/a de EF habla claro, escucha a los estudiantes atentamente y responde con precisión las preguntas de los alumnos.	

2 - CONCIENCIA DE TRABAJO

BAJO		MEDIO			ALTO	
1	2	3	4	5	6	7
El profesor/a de EF a menudo llega sin prepararse las clases y/o sin los materiales apropiados. Durante el curso los materiales no están bien preparados y/o desactualizados.		El profesor/a de EF no imparte todos los contenidos que se indican en el programa de la asignatura y a veces parecen un poco improvisados.			El profesor/a de EF proporciona un detallado y completo programa del curso, prepara e imparte muy bien todos los contenidos.	

3 - CREATIVIDAD

BAJO		MEDIO			ALTO	
1	2	3	4	5	6	7
El profesor/a de EF utiliza la misma forma de enseñar (método) en todas las clases y no cambia nunca.		A veces, el profesor/a de EF varía la forma de enseñar (método) en algunas clases, por ejemplo mostrando un vídeo.			El profesor/a de EF utiliza múltiples formas de enseñar las clases (métodos de enseñanza), tales como tareas individuales y debates en grupo.	

4 – FEED-BACK

BAJO		MEDIO			ALTO	
1	2	3	4	5	6	7
El profesor/a de EF asigna tareas en clase sin proporcionar una información específica y detallada de los puntos importantes.		El profesor/a de EF, a veces, proporciona información detallada a los alumnos durante las actividades (retroalimentación), para la mejor realización de las tareas.			El profesor/a de EF proporciona siempre información detallada, durante las actividades (retroalimentación), lo que ayuda a los alumnos a progresar más.	

5 – CONSIDERACIÓN INDIVIDUAL AL ALUMNADO

BAJO		MEDIO			ALTO	
1	2	3	4	5	6	7
El profesor/a de EF no proporciona atención personal a los alumnos, y parece tener prisa y poco interés cuando se le solicita ayuda.		El profesor/a de EF suele estar disponible para ayudar, pero a veces es reacio a proporcionar ayuda individual o atención a los alumnos.			El profesor/a de EF proporciona ayuda individual a los alumnos siempre que se le requiere, dedicando tiempo extra para motivar a los estudiantes.	

6 – PROFESIONALIDAD

BAJO		MEDIO			ALTO	
1	2	3	4	5	6	7
El profesor/a de EF cambia los contenidos del curso sin consultarlo a los alumnos y trabaja con métodos totalmente injustos.		El profesor/a de EF es generalmente justo y equitativo, pero puede ser incongruente en momentos de la evaluación y calificación o cuando habla de los objetivos del curso.			El profesor/a de EF es totalmente justo y equitativo y pide el consentimiento completo de los alumnos cuando cambia alguno de los contenidos didácticos del curso y aclara plenamente el examen y los criterios de evaluación.	

7 – RESOLUCIÓN DE PROBLEMAS

BAJO		MEDIO			ALTO	
1	2	3	4	5	6	7
El profesor/a de EF se niega a llegar a un consenso con la clase cuando se plantean problemas, tales como la adecuación de contenidos del programa o hacer nuevas disposiciones para el trabajo asignado.		El profesor/a de EF se interesa, generalmente, por las preocupaciones de los estudiantes de forma eficaz, sin embargo, las soluciones no son siempre aceptadas por todos.			El profesor/a de EF se ocupa de cuestiones que impiden el aprendizaje del alumno, facilitando las explicaciones necesarias en clase cuando algún contenido no está claro.	

8 – CONCIENCIA SOCIAL

BAJO		MEDIO			ALTO	
1	2	3	4	5	6	7
El profesor/a de EF no consigue mantener el orden en clase y no acepta las críticas que se le hacen.		Generalmente el profesor/a de EF mantiene el orden en clase, pero a veces señala individualmente a los alumnos por sus errores y trata algunos temas delicados torpemente.			El profesor/a de EF parece entender completamente los sentimientos de los alumnos, tratando temas sensibles con mucho tacto y no señala individualmente a los alumnos por sus errores.	

