



Validación convergente de la batería Condición Física Secundaria como alternativa a Eurofit en ESO

Convergent validation of the Secondary School Physical Fitness battery as an alternative to Eurofit in secondary education

Autores

Julio Martín-Ruiz ¹
 Concepción Ros-Ros ¹
 Clara Gallego-Cerveró ¹
 Ignacio Tamarit-Grancha ¹
 Laura Ruiz-Sanchis ¹

¹ Universidad Católica de Valencia (España)

Autor de correspondencia:
 Julio Martín-Ruiz
julio.martin@ucv.es

Recibido: 19-04-26
 Aceptado: 30-05-26

Cómo citar en APA

Martín-Ruiz, J., Ros-Ros, C., Gallego-Cerveró, C., Tamarit-Grancha, I., & Ruiz-Sanchis, L. (2026). Validación convergente de la batería Condición Física Secundaria como alternativa a Eurofit en ESO. *Retos*, 83, 211-228. <https://doi.org/10.47197/retos.v83.119724>

Resumen

Introducción: La evaluación de la condición física en la asignatura de Educación Física requiere pruebas válidas, aplicables y útiles para orientar la intervención docente.

Objetivo: Este estudio comparó la batería Eurofit con una propuesta alternativa a la que se denominó Condición Física Secundaria (CFS), diseñada para valorar la flexibilidad, la fuerza del tronco, el salto, el lanzamiento, la velocidad y la resistencia en alumnado de ESO.

Metodología: Participaron 148 adolescentes; 82 chicos y 66 chicas, de 1º a 4º curso. Se analizaron asociaciones entre pruebas homólogas, diferencias entre notas y esfuerzo percibido, relación con sexo, edad, práctica física semanal y variables antropométricas, así como la validez convergente de la calificación final CFS respecto a Eurofit.

Resultados: La batería CFS mostró una asociación significativa con Eurofit en la nota final ($\rho = .54$; $p < .001$; $r = .55$; $p < .001$), aunque el acuerdo absoluto fue moderado-bajo ($ICC = .41$), indicando que ambas baterías no son intercambiables en todos los aspectos. La CFS redujo el esfuerzo percibido global en un 32.4% ($p < .001$) y mantuvo asociaciones relevantes en velocidad, lanzamiento y resistencia.

Discusión: El test Thomas, Biering-Sørensen, la combinación SJ-CMJ-índice elástico, el lanzamiento rotacional, los 30 m y la prueba de Rockport, aportaron información más específica y aplicable al contexto escolar.

Conclusiones: Los resultados apoyaron la validación convergente de CFS como alternativa viable, funcional e inclusiva para evaluar la condición física en Educación Física, y futuras investigaciones deberán confirmar su fiabilidad, sensibilidad al cambio y validez externa.

Palabras clave

Condición física; educación física; evaluación; eurofit; educación secundaria; pruebas de campo.

Abstract

Introduction: The assessment of physical fitness in Physical Education requires tests that are valid, applicable, and useful for guiding teaching practice.

Objective: This study compared the Eurofit battery with an alternative test known as the Secondary School (CFS), designed to assess flexibility, core strength, jumping, throwing, speed, and endurance in pupils in secondary education.

Methodology: A total of 148 adolescents participated: 82 boys and 66 girls from years 7 to 10. This study analyzed the associations between the corresponding tests, differences in marks and perceived effort, relationships with sex, age, weekly physical activity, and anthropometric variables, as well as the convergent validity of the final CFS mark in relation to Eurofit.

Results: The CFS battery showed a significant association with Eurofit in terms of the final mark ($\rho = .54$; $p < .001$; $r = .55$; $p < .001$), although the absolute agreement was moderate-low ($ICC = .41$), indicating that the two test batteries are not interchangeable in all respects. The CFS reduced overall perceived exertion by 32.4% ($p < .001$) and maintained significant associations with speed, throwing, and endurance.

Discussion: The Thomas test, Biering-Sørensen test, SJ-CMJ-elasticity index combination, rotational throw, 30 m sprint, and Rockport test provided information that was more specific and applicable to the school context.

Conclusions: The results support the convergent validation of the CFS as a viable, functional, and inclusive alternative for assessing physical fitness in Physical Education, and future research should confirm its reliability, sensitivity to change, and external validity.

Keywords

Physical fitness; physical education; assessment; eurofit; secondary education; field tests.

Introducción

La valoración de la condición física de los escolares entre los 12 y 16 años es un elemento imprescindible en el área de Educación Física (EF) (Ortega et al., 2008). En ella se exploran las capacidades condicionales a potenciar en los jóvenes, basadas en procesos energéticos del ejercicio en su expresión de resistencia, fuerza, velocidad y flexibilidad, que deben ser valoradas en el alumnado, al tener una alta relevancia como marcador de salud (Ortega et al., 2008) y relación con la autopercepción física (Galán-Arroyo et al., 2024).

La forma más habitual de hacerlo es concentrar las pruebas en baterías. Dentro de la EF, la Eurofit destaca por su validez científica tras evaluar la aptitud física a más de 50000 escolares europeos (Tomkinson et al., 2007) y por la posibilidad de comparación de los datos obtenidos a nivel internacional (Cuadrado et al., 2009). Cada componente de la condición física dispone de su instrumento de evaluación, pero no es sencillo medir de forma precisa especialmente en niños (Cordente-Martínez, 2006), por ello las pruebas de campo, habituales en EF tienen la ventaja de la especificidad y similitud con la actividad (Sainz de Baranda & Ayala 2010). La aplicación de estas permite que el alumnado conozca sus posibilidades motrices (González et al., 2013) y facilita el proceso de evaluación al docente, con un baremo claro, validez de criterio y fácil interpretación desde el plano pedagógico (Castro-Piñero et al., 2010; Ruiz et al., 2011). La Batería Eurofit está muy arraigada en la EF, aunque la evidencia en las últimas décadas invita a plantear alternativas a las pruebas que la componen por otras con procedimientos e información más precisa tras la medición para el profesorado, y con adecuación a las necesidades del alumnado.

En el caso de la fuerza, es una variable imprescindible porque su nivel de aptitud muscular en niños y adolescentes se asocia a indicadores de salud cardiometabólica, ósea y funcional (Smith et al., 2014), por ello no debe emplearse un único resultado global de rendimiento. En Eurofit se describen tres pruebas; salto vertical, lanzamiento de balón medicinal y el test abdominal en un minuto, que presentan ciertas limitaciones. El salto vertical (Sargent, 1921), valora la fuerza explosiva ejecutada en contramovimiento (Izquierdo & Badillo, 2008). La información que se traslada se limita a la comparación de alturas obtenidas entre sujetos, y el baremo, no especifica la cantidad de fuerza explosiva generada o las directrices claras a seguir en función del resultado. En el caso del lanzamiento de balón medicinal, su mejor versión depende de variables antropométricas que sesgan el resultado; a mayor altura y/o envergadura, mayor será el ángulo y altura de salida (Borms et al., 2016), incluso cuando el sujeto dispone de una menor fuerza relativa. Por otro lado, el gesto bilateral, es inespecífico de las principales acciones de fuerza cotidianas (unilaterales). La última propuesta de esta capacidad es la del abdominal en un minuto a través de una flexión bilateral del tronco, valorada por la cantidad total de repeticiones. Se ejecuta a velocidad máxima y ese carácter balístico y de gran amplitud, incurre en gestos desaconsejados como la sobrecarga muscular de flexores de cadera, agravado si hay agarre de los pies (Miñarro 2008), con impactos que comprometen el raquis cervical.

En el caso de la velocidad, se realiza una aceleración de 50 metros de carácter anaeróbico aláctico (Inbar & Amon, 2016). La inexperiencia del alumnado provoca una pérdida de velocidad que cuestiona el término de aceleración, ya que en sprinters hombres y mujeres, la máxima velocidad se alcanza entre los 45 y 60 m (Ward-Smith 2001), evidenciando que el reto propuesto al alumnado genera una fatiga no deseada en el diseño de la prueba y debería ser más breve para cumplir de forma óptima con su propósito.

En la resistencia, el test de Cooper recoge información sobre la condición física aeróbica (Bandyopadhyay, 2015). Su duración e intensidad derivan en que la obtención de energía sea de carácter láctico (O'gorman et al., 2000), por lo que resulta excluyente para algunos perfiles (cuadros asmáticos o baja condición física con imposibilidad de alcanzar o mantener una velocidad mínima) para el objetivo que teóricamente cumple, haciendo deseable contar con una prueba asumible por todas las realidades.

En lo referente a la flexibilidad, se incluye el conocido sit and reach o banco de Wells (Wells & Dillon 1952), que calibra la flexibilidad isquiotibial. Presenta conflictos al requerir una inversión lumbar y una hipercifosis dorsal, que invalida el análisis isquiotibial al bascular la cadera. Su habitual denominación de prueba de flexibilidad global no aclara los puntos de corte y variables antropométricas como la longitud de pierna o la envergadura condicionan el resultado, por lo que la integración de una prueba que determine de forma localizada la capacidad muscular de flexar, optimizaría su objetividad.



Por tanto, las pruebas explicadas en ocasiones incurren en prácticas potencialmente lesivas, técnicamente mejorables o fisiológicamente inespecíficas, ofreciendo información limitada para proponer mejoras en cada componente de la condición física. Por estos motivos, la finalidad de este proyecto fue la de comparar la Batería de test Eurofit con una selección de pruebas que se ha denominado Condición Física Secundaria (CFS). Se comprobó su precisión en la valoración de la condición física de alumnado de la ESO, para ofrecer una alternativa al profesorado en la evaluación de este contenido, con renovada evidencia técnica y fisiológica.

Con los resultados obtenidos, se espera proponer una evaluación de baja dificultad de ejecución y sencilla toma de datos para el profesorado, adaptados al ámbito educativo en material y tiempo disponible en las aulas de EF.

La hipótesis principal es que la batería CFS aportará información más precisa sobre el nivel de condición física del alumnado de Educación Secundaria, y supondrá una alternativa viable y validada con respecto a la Batería Eurofit.

Método

Para determinar la eficacia de una batería de test a la que se denominó Condición Física Secundaria (CFS), con respecto a la Batería Eurofit, se realizó un estudio comparativo, observacional y transversal con alumnado de la ESO de la ciudad de Valencia durante el curso lectivo 2025-26.

Se recogieron las autorizaciones de tres centros y los consentimientos de los padres del alumnado, al ser menores de edad. Las sesiones se llevaron a cabo en los horarios lectivos de EF siguiendo las directrices de la Declaración de Helsinki, y el estudio tuvo la aprobación del Comité de Ética de la Universidad Católica de Valencia con código UCV/2025-26/120.

Participantes

La muestra estuvo formada por alumnado de Educación Secundaria de 1º a 4º curso (191 adolescentes). El total que terminó las 12 pruebas de que constó el estudio fueron 148 (82 chicos y 66 chicas) de tres centros educativos de la ciudad de Valencia. Los requisitos de inclusión fueron estar entre 1º y 4º de la ESO, asistir regularmente a las clases de EF y no realizar práctica de actividad física en el día de la prueba ni el anterior.

De los 191 alumnos inicialmente registrados, 43 no completaron el protocolo completo de 12 pruebas. No se dispuso de un registro desagregado por criterio en la base de datos final remitida, por lo que no fue posible construir un diagrama de flujo CONSORT completo. Los motivos previstos en el protocolo fueron lesión en el momento de la evaluación, convalecencia inferior a un mes y exención médica de Educación Física. Esta ausencia de desglose se reconoce como limitación y se propone su registro sistemático en futuros estudios para mejorar la transparencia y valorar posible sesgo de selección.

Los motivos de exclusión fueron padecer una lesión en el momento del estudio, haber estado convaleciente menos de un mes antes de las pruebas o estar exento de cursar EF por motivos médicos.

Procedimiento

Tras obtener los consentimientos de todos los participantes, se realizaron cinco sesiones de trabajo: 1.a) Toma de datos antropométricos y pruebas de flexibilidad, tronco y salto vertical de ambas baterías, 2.a) realización de las pruebas de lanzamiento, 3.a) sprint de ambas baterías, 4.a) prueba de resistencia de Eurofit y 5.a) test de Rockport de la batería CFS. Las sesiones 2 a 5 se balancearon para no condicionar los resultados. Entre cada una de ellas hubo un intervalo mínimo de 72 horas.

Las pruebas antropométricas se llevaron a cabo por el investigador principal (ISAK-I); envergadura, distancia biacromial, talla, peso y longitud de pierna. La envergadura requirió que el sujeto estuviera en posición erguida con hombros en abducción de 90º y codos extendidos, tomando la medida entre las falanges más distales. La distancia biacromial, con el sujeto en posición anatómica recogiendo la distancia entre ambos acromiones. El peso con una báscula digital tras dos segundos de estatismo (Seca, 750, Hamburgo, Alemania), la talla con un estadiómetro situando al sujeto en el ángulo de Frankfort (Seca,

213, Hamburgo, Alemania) y por último la longitud de pierna; con el sujeto en sedestación, rodilla extendida y tobillo en flexión plantar, se tomó la medida desde el trocánter mayor hasta la punta del pie.

Se diseñó un calentamiento estandarizado para todas las sesiones que incluía: ejercicios de movilización y compactación del raquis: Cat-camel (6 rep), Bird-dog (6 rep) y plancha frontal (15seg). A continuación, movilidad articular global (2 min), lanzamientos verticales de balón medicinal (10rep) y dos progresivos de carrera de 30m.

Las pruebas Eurofit, denominadas EFIT en los procedimientos del presente estudio, se gestionaron según las directrices generales del manual Eurofit del Consejo de Europa, y se calificaron mediante los baremos escolares propuestos por Blázquez para Educación Física (2006), ampliamente conocidos por el profesorado. En la batería CFS, cada prueba fue seleccionada a partir de literatura específica sobre su componente físico principal con las adaptaciones mínimas necesarias para el contexto escolar. Tras la realización de cada prueba, se requirió a cada participante que indicara su percepción del esfuerzo usando una escala específica con un pictograma, más intuitiva para el contexto escolar (Lagally & Robertson 2006).

Flexibilidad

En el caso de la EFIT (Wells & Dillon 1952), se realizó desde sedestación, con pies descalzos y rodillas extendidas, realizando una flexión máxima de tronco, deslizando con ambas manos una señal por una superficie centimetrada, diseñada ad hoc. Se midió desde el valor 0 hasta el lugar al que se llegasen ambas manos (positivo o negativo), tomando el mejor de dos intentos.

Para la CFS, se realizó el test de Thomas modificado de flexibilidad isquiosural (Peeler & Anderson 2008) orientado a identificar limitaciones en la flexibilidad de flexores de cadera, control lumbopélvico y posibles asimetrías. El sujeto se situó sobre una superficie elevada quedando en posición de tendido supino, con todo su cuerpo en apoyo salvo la mitad distal de los muslos y sus piernas, suspendidas con flexión de rodillas. Desde esa posición, realizó la flexión de una de las caderas con sus manos atrayendo el muslo hacia el tórax. Se comprobó si la cadera contraria se flexionaba de forma involuntaria y/o si existía extensión espontánea de la rodilla. En caso de no ocurrir la puntuación fue de 1, y de lo contrario 2. La descripción gráfica de ambas pruebas puede verse en la Figura 1.

Figura 1. Pruebas de flexibilidad EFIT y CFS.



Nota. Panel A: cajón de flexibilidad diseñado ad hoc; Panel B: representación de la prueba de flexibilidad EFIT.; Panel C: test de Thomas de CFS.

Tronco

En la EFIT se realizó el test abdominal 1 minuto (Cuadrado et al., 2009). Desde tendido supino con rodillas flexionadas y pies bajo un soporte (o compañero). Se realizó flexión de tronco bilateral de forma continua durante 60 segundos, anotando el número de repeticiones logrado en un solo intento.

Para la CFS, se propuso la prueba de Biering-Sørensen, que valora la resistencia isométrica de los extensores del tronco y el mantenimiento postural. Está ampliamente descrita y con evidencia suficiente de fiabilidad en comparación con otros test de extensión del tronco (Martínez-Romero et al., 2020). El sujeto se situó sobre una camilla en posición de tendido prono, con la cadera al borde sobre la cresta iliaca anterosuperior, sujetado con cinchas para asegurar al participante. A la señal de investigador, se mantuvo la posición horizontal a través de una extensión paravertebral con las manos cruzadas por delante

y el raquis cervical en posición neutra. Se anotaron los segundos que pudo mantenerse en estatismo sin que hubiera oscilaciones o temblores derivados de la fatiga. El detalle de estas pruebas puede verse en la Figura 2.

Figura 2. Prueba abdominal EFIT y de estabilización del tronco CFS.



Nota. Panel A: posición inicial y final del test abdominal EFIT.; Panel B: representación de la prueba de Biering-Sørensen en CFS.

Salto vertical

En EFIT se realizó el test de salto libre de Sargent (Sargent, 1921). Desde bipedestación bipodal con pies separados a la anchura de los hombros, flexionaron un hombro hasta alcanzar la mayor altura posible marcada en una regleta en la pared. A continuación, se realizó un salto mediante una rápida flexo-extensión de rodillas. Con la mano dominante, se trató de alcanzar la máxima altura tocando la regleta. Se contabilizó la diferencia entre la posición inicial con el hombro en flexión y la altura alcanzada en el salto.

En la CFS, se evaluó mediante SJ, CMJ e índice elástico (Bosco & Komi 1979), porque la comparación entre un salto sin contramovimiento y otro con él, permite diferenciar la fuerza explosiva concéntrica del aprovechamiento del ciclo estiramiento-acortamiento (Markovic et al., 2004). La inclusión conjunta de ambas pruebas aporta una lectura más completa que la altura de salto aislada, y se apoya en estudios clásicos y actuales sobre la fiabilidad y validez del SJ y el CMJ (Petrigna et al., 2019). La ejecución del SJ fue desde bipedestación bipodal con pies separados a distancia biacromial y manos en la cadera. Ahí se realizó flexión de rodillas hasta 90° y tras 4 segundos estáticos, se realizó un salto vertical. Desde la misma posición inicial, el CMJ se realizó con un salto libre mediante flexo-extensión de rodillas.

Todos los test se midieron con una plataforma de contacto (Contact Platform kit, DIN A-2, Chronojump, Barcelona, España). La descripción gráfica se encuentra en la Figura 3.

Figura 3. Pruebas de salto vertical EFIT y CFS.



Nota. Panel A: plataforma de contacto.; Panel B: recreación del salto vertical Sargent en EFIT.; Panel C: saltos SJ y CMJ en CFS.

Lanzamiento de balón medicinal

Para la EFIT se empleó la prueba de lanzamiento de balón medicinal (Cuadrado et al., 2009). Desde bipedestación bipodal con pies a distancia biacromial y tras una línea, se realizó un lanzamiento de balón medicinal (3kg chicas, 5kg chicos) con ambos brazos, ayudado de una extensión de raquis. Se midió desde la línea tras la que se sitúan los pies, hasta la huella más cercana a éstos.

En la CFS, se propuso un lanzamiento rotacional unilateral con idéntico peso al anterior, con un gesto apoyado en la literatura sobre pruebas de potencia rotacional con balón medicinal (Hardy et al., 2025). Se ejecutó desde bipedestación bipodal, con pies separados a 1,5 veces la distancia biacromial, tras realizar una rotación de tronco y flexión de rodillas, para deshacer el movimiento en sentido contrario con una rápida extensión articular. Se realizó por ambos lados y se midió desde la línea tras la que se situaron los pies, hasta la huella más cercana a éstos en la que cayera el balón.

Dado que no existen baremos publicados específicos para alumnado de ESO con el protocolo aplicado en este estudio, la conversión a nota se realizó mediante un baremo propio, generado a partir de la distribución de resultados por sexo y edad de la muestra y de análisis preliminares del grupo investigador actualmente en revisión editorial a espera de validación externa. Al no estar publicada en el momento de la redacción, no se empleó como evidencia bibliográfica independiente, sino como criterio interno explícito del presente trabajo.

Un ejemplo de ambas pruebas de lanzamiento puede verse en la Figura 4.

Figura 4. Lanzamientos de las baterías EFIT y CFS.



Nota. Panel A: colocación de la cinta métrica.; Panel B: representación de lanzamiento bilateral EFIT.; Panel C: representación de lanzamiento rotacional unilateral CFS.

Sprint

En el caso de EFIT, se realizó una aceleración de 50 metros (Cuadrado et al., 2009). Desde posición de bipedestación bipodal con un pie adelantado y en posición de tres apoyos. La mano de apoyo permaneció sobre un pulsador (Floor push button Kit, Chronojump, Barcelona, España). A la señal del investigador, el participante salía, momento en el que separaba la mano del pulsador activando el tiempo y pasando por dos fotocélulas a la máxima velocidad posible.

La velocidad en CFS se evaluó mediante una carrera de 30 m con un procedimiento idéntico al anterior salvo por la menor distancia. Se seleccionó esta medida por ajustarse mejor al objetivo de valorar la aceleración en alumnado no especialista y contar con evidencia sobre fiabilidad interevaluador y validez de la medición (Vicente-Rodríguez et al., 2011).

Ambas pruebas se midieron con fotocélulas inalámbricas (WICHRO: Wireless Race Kit, Chronojump, Barcelona, España). Un ejemplo de estas pruebas se encuentra en la Figura 5.

Figura 5. Pruebas de velocidad EFIT y CFS.



Nota. Panel A: disposición de fotocélulas de las pruebas de velocidad.; Panel B: pulsador de salida.; Panel C: representación de las pruebas de sprint EFIT y CFS.

Resistencia

En el caso de la EFIT, se realizó el test de Cooper (Bandyopadhyay, 2015). Los participantes realizaron una carrera de 12 minutos ininterrumpidos a la máxima velocidad media posible, contabilizando el número de metros recorridos. Al término de la prueba, se instó al alumnado a quedarse en posición estática para poder calcular el número adicional de metros con respecto a la salida.

La resistencia cardiorrespiratoria en CFS se evaluó mediante el test de la milla/Rockport, consistente en recorrer 1609 m caminando lo más rápido posible y registrando el tiempo y la frecuencia cardiaca al finalizar. Para su cálculo se emplearon brazaletes (Moofit, Shenzhen, Guangdong, China), recogiendo la señal con el software Pulsemonitor (Pulsemonitor, Michalowice, Polonia). La prueba se seleccionó al ser submáxima, por su facilidad de aplicación en contexto escolar y su utilidad para estimar el VO_2max mediante ecuaciones de campo. La ecuación original fue desarrollada en adultos, aunque cuenta con validación en población de edad escolar (McSwegin et al., 1998), lo que respalda su aplicabilidad como alternativa más inclusiva a pruebas máximas de carrera y se ha empleado, además, para calcular el VO_2peak en pruebas de campo junto a medidas antropométricas (Cureton et al., 1995).

El detalle de ambas pruebas se refleja en la Figura 6.

Figura 6. Pruebas de resistencia EFIT y CFS.



Nota. Panel A: captura del perímetro de la prueba de Cooper (EFIT) y de su ejecución.; Panel B: entorno de la prueba de Rockport (CFS) con el detalle del software Pulsemonitor y fotograma de una de las sesiones.

En la Tabla 1, se muestra el resumen de baremos empleados en ambas baterías.

Tabla 1. . Baremos de las baterías EFIT y CFS.

Dimensión	EFIT	CFS	Referencia principal CFS
Flexibilidad	Sit-and-reach	Thomas	Peeler & Anderson, 2007
Tronco	Abdominal 1 min	Biering-Sørensen	Martínez-Romero et al., 2020
Salto	Salto vertical	SJ, CMJ e índice elástico	Bosco & Komi, 1979
Lanzamiento	Lanzamiento frontal	Lanzamiento rotacional	Hardy et al., 2025
Velocidad	50 m	30 m	Vicente-Rodríguez et al., 2011
Resistencia	Cooper	Rockport/milla	McSwegin et al., 1998

Análisis de datos

El análisis se realizó sobre 148 escolares de 1º a 4º de ESO. Las variables continuas se describieron mediante media y desviación estándar, y las categóricas mediante frecuencias. La comparación principal entre pruebas emparejadas se realizó con la prueba de rangos con signo de Wilcoxon, adecuada para puntuaciones por baremo, variables ordinales y distribuciones potencialmente no normales. La magnitud del cambio pareado se expresó mediante *d* de Cohen para muestras dependientes. Las asociaciones entre pruebas homólogas, edad, práctica semanal de actividad física y variables antropométricas se estimaron mediante correlación de Spearman.

Las diferencias por sexo se analizaron con la prueba U de Mann-Whitney. Para cuantificar el posible sesgo antropométrico de cada prueba, se calculó la media de las correlaciones absolutas entre la nota de la prueba y peso, talla, IMC, envergadura, diámetro biacromial y longitud de pierna. En velocidad se añadió una comparación específica de los tramos finales de 10 m: Eurofit 40-50 m frente a CFS 20-30 m. Se analizó mediante Wilcoxon, correlación de Spearman y proporción de alumnado cuyo tiempo final fue mayor en el tramo 40-50 m, con el objetivo de comprobar si el 50 m introduce pérdida de velocidad o fatiga terminal respecto al 30 m.

La validación de la batería CFS frente a Eurofit se abordó mediante validez convergente entre notas finales, acuerdo absoluto mediante ICC(2,1), análisis de Bland-Altman, consistencia interna mediante alfa de Cronbach, validez factorial exploratoria mediante asociación con el primer componente principal y predicción cruzada de la nota final Eurofit a partir de las seis pruebas CFS con regresión lineal y validación cruzada de cinco particiones. El nivel de significación se fijó en $p < .05$.

Resultados

Características de la muestra

La muestra fue de 82 chicos y 66 chicas, con una edad media de 13.86 ± 1.31 años. Los chicos indicaron más horas semanales de práctica física extraescolar que las chicas y mayor talla, peso, envergadura, diámetro biacromial y longitud de pierna. No se observaron diferencias por sexo en la nota final EFIT ni en la CFS, lo que sugiere que los baremos que se aplicaron compensaron de forma adecuada las diferencias biológicas y antropométricas esperables en esta etapa (Tabla 2).

Tabla 2. Características de la muestra y comparación por sexo.

Variable	Total	Chicos	Chicas	<i>p</i> sexo
Edad	13.86 ± 1.31	14.02 ± 1.28	13.67 ± 1.34	.071
Horas/semana extracurriculares	3.87 ± 3.31	4.61 ± 3.29	2.96 ± 3.13	< .001
Peso	56.77 ± 12.53	59.54 ± 13.55	53.34 ± 10.24	.002
Talla	163.62 ± 9.75	167.62 ± 9.57	158.65 ± 7.45	< .001
IMC	21.13 ± 4.02	21.05 ± 3.92	21.22 ± 4.18	.912
Envergadura	164.19 ± 13.47	168.33 ± 15.46	159.05 ± 8.0	< .001
Biacromial	37.96 ± 3.44	39.09 ± 3.54	36.56 ± 2.75	< .001
Longitud_pierna	99.29 ± 5.99	101.07 ± 6.27	97.08 ± 4.8	< .001
Nota final EFIT	4.69 ± 1.75	4.82 ± 1.81	4.54 ± 1.66	.360
Nota final CFS	5.59 ± 1.1	5.51 ± 1.11	5.7 ± 1.08	.368

Nota. Los valores se presentan como media $\pm$ desviación estándar. La comparación por sexo se realizó mediante U de Mann-Whitney.

Comparación entre pruebas homólogas Eurofit y CFS

La comparación por dimensiones mostró que CFS no replica exactamente a Eurofit, sino que mantiene una relación variable con ella y, a su vez, modifica el tipo de información que recibe el profesorado. Las pruebas con mayor convergencia fueron lanzamiento, velocidad y resistencia. En lanzamiento, el rotacional CFS se asoció de forma alta con el frontal Eurofit; en velocidad, el 30 m CFS mostró una convergencia alta con el 50 m EFIT; y en resistencia, Rockport se relacionó significativamente con Cooper, pero reduciendo de forma marcada el esfuerzo percibido.

En flexibilidad y tronco, las correlaciones con EFIT fueron bajas, algo que debe interpretarse como una diferencia conceptual, no de falta de utilidad. Thomas no mide la misma combinación de movilidad de cadera, pelvis y columna que el sit and reach, sino una función más localizada de flexores de cadera. Biering-Sørensen evalúa resistencia extensora y control postural del tronco, mientras que el abdominal EFIT cuantifica repeticiones de flexión de tronco en un minuto. En salto, aunque las notas medias de ambas pruebas fueron equivalentes, el índice CFS basado en SJ, CMJ e índice elástico aportó información neuromuscular diferente a la altura aislada del salto EFIT (Tabla 3).

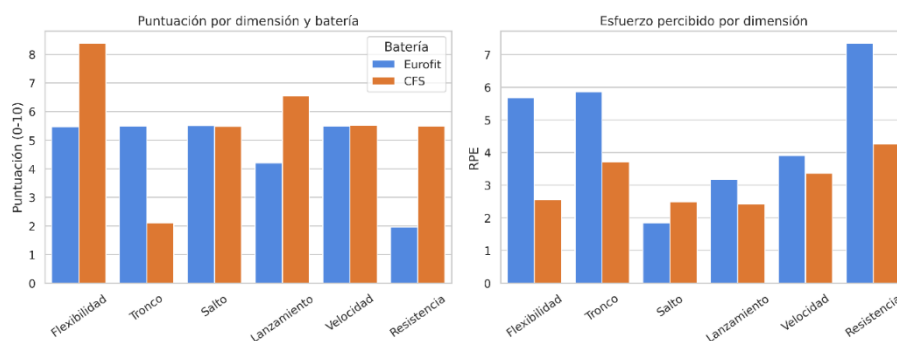
Tabla 3. Comparación estadística de las pruebas homólogas Eurofit-CFS.

Dimensión	Eurofit	CFS	Eurofit nota media ± DE	CFS nota media ± DE	Wilcoxon on p	rho not as	p rho not as	RPE Eurofit media ± DE	RPE CFS media ± DE	Wilcoxon p RPE	Bias antropométrico Eurofit	Bias antropométrico CFS	rho horas Eurofit	rho horas CFS	rho edad Eurofit	rho edad CFS	p sexo Eurofit	p sexo CFS
Flexibilidad	Eurofit sit-and-reach	CFS Thomas	5.48 ± 2.79	8.4 ± 1.96	<.001	.01	.874	5.68 ± 2.38	2.55 ± 1.8	<0.001	.05	.05	.23	-.017	-.01	-.11	0.882	.023
Tronco	Eurofit abdominal 1 min	CFS Biering-Sørensen	5.5 ± 2.77	2.12 ± 0.56	<.001	.18	.026	5.85 ± 2.13	3.72 ± 1.92	<0.001	.11	.07	.23	0.11	.0	.2	0.998	.313
Salto	Eurofit vertical jump	CFS SJ-CMJ-EI	5.51 ± 2.79	5.49 ± 2.79	.959	-.28	<.001	1.85 ± 1.36	2.49 ± 1.6	<0.001	.14	.03	.2	-.002	.01	-.01	0.965	.949
Lanzamiento	Eurofit frontal throw	CFS rotational throw	4.2 ± 2.83	6.55 ± 2.96	<.001	.64	.001	3.18 ± 1.5	2.42 ± 1.24	<0.001	.25	.27	.25	0.3	.03	.08	0.664	.484
Velocidad	Eurofit 50 m	CFS 30 m	5.5 ± 2.78	5.52 ± 2.79	.736	.59	<.001	3.91 ± 1.9	3.36 ± 1.81	<0.001	.1	.14	.31	0.21	.0	.01	0.986	.912
Resistencia	Eurofit Cooper	CFS Rockport mile	1.96 ± 2.73	5.49 ± 2.78	<.001	.43	.001	7.35 ± 1.98	4.26 ± 2.11	<0.001	.14	.2	.36	0.37	.09	0	<0.001	.96

Nota. rho = correlación de Spearman. RPE = esfuerzo percibido. Bias antropométrico = media de correlaciones absolutas con peso, talla, IMC, envergadura, biacromial y longitud de pierna. Las notas se expresan en escala 0-10.

La Figura 7 muestra que CFS desplaza la evaluación hacia pruebas con menor carga percibida en flexibilidad, tronco, lanzamiento, velocidad y resistencia, y sitúa las notas en un rango operativo más equilibrado para la evaluación escolar.

Figura 7. Puntuación por dimensión y batería.



Nota. Las barras representan la media de cada dimensión. RPE = percepción subjetiva del esfuerzo.

Análisis específico de velocidad: tramo final 40-50 m frente a 20-30 m

La comparación específica de los últimos 10 m permitió comprobar si la prueba EFIT de 50 m introduce una fase final más lenta que pueda alejarla de valorar principalmente la aceleración. El tiempo medio en el tramo 40-50 m de EFIT fue de $1,66 \pm 0,32$ s, mientras que el de 20-30 m de CFS fue de $1,65 \pm 0,28$ s. La diferencia media muy reducida (0,01 s), y no alcanzó significación estadística aplicando la prueba de Wilcoxon ($W = 4239,0$; $p = 0,148$; $d z = 0,05$).

A pesar de que la diferencia media no fue significativa, el 56,9% del alumnado registró un tiempo más lento en el tramo 40-50 m que en el 20-30 m, y ambos se asociaron de forma alta ($\rho = .78$; $p < .001$).



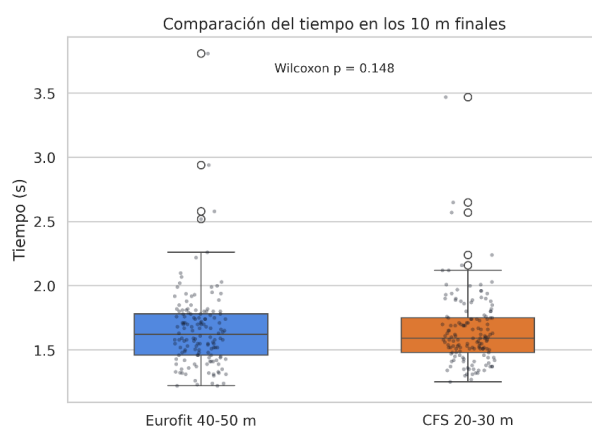
En términos grupales no se confirma una pérdida terminal estadísticamente significativa, pero la tendencia individual mayoritaria apunta a que una proporción relevante del alumnado empieza a perder eficiencia al final del 50 m. Desde un punto de vista pedagógico y fisiológico, 30 m son más generalizables para la Educación Física cuando el objetivo es valorar aceleración, al no añadir distancia innecesaria ni un posible componente de fatiga en alumnado inexperto (Tabla 4, Figura 8).

Tabla 4. Comparación de los últimos 10 m en la prueba de velocidad.

Comparación	Eurofit 40-50 m media±DE	CFS 20-30 m media±DE	Diferencia media EFIT-CFS	IC95% diferencia	Wilcoxon W	p	d z	rho Spearman	p rho	% más lentos en 40-50 m
Últimos 10 m EFIT 50 m vs CFS 30 m	1.66 ± 0.32	1.65 ± 0.28	.01	[-.02, .04]	4239.0	.148	.05	.78	< .001	56.9

Nota. EFIT 40-50 m representa el tramo final de la prueba de 50 m. CFS 20-30 m representa el tramo final de la prueba de 30 m. La comparación pareada se realizó mediante Wilcoxon

Figura 8. Comparación del tiempo en los 10 m finales.



Nota. EFIT 40-50 m corresponde al tramo final de la prueba de 50 m. CFS 20-30 m corresponde al tramo final de la prueba de 30 m. Valores más altos indican menor velocidad en el tramo.

Prueba que define mejor el perfil de condición física en cada dimensión

La elección de la prueba para definir el perfil de condición física debía priorizar que fuera interpretable para Educación Física, menos dependiente de la antropometría, con menor esfuerzo percibido y más específica del componente físico evaluado. Con esas premisas, las seis dimensiones favorecieron la alternativa CFS, por motivos distintos (Tabla 5).

Tabla 5. Justificación de la prueba que define mejor cada dimensión del perfil físico.

Dimensión	Prueba que define mejor el perfil	Evidencia estadística clave	Justificación aplicada
Flexibilidad	CFS Thomas	rho notas = .01, $p = .874$; RPE CFS = 2.55 ± 1.8 vs EFIT = 5.68 ± 2.38	Thomas localiza el acortamiento flexor de cadera, evita la compensación lumbar del <i>Sit-and-reach</i> y reduce claramente el esfuerzo percibido.
Tronco	CFS Biering-Sørensen	rho notas = .18, $p = .026$; RPE CFS = 3.72 ± 1.92 vs EFIT = 5.85 ± 2.13	Biering-Sørensen informa sobre resistencia extensora del tronco y evita la flexión balística repetida del test abdominal.
Salto	CFS SJ-CMJ-EI	rho notas = -.28, $p = < .001$; RPE CFS = 2.49 ± 1.6 vs EFIT = 1.85 ± 1.36	SJ-CMJ-EI separa fuerza explosiva, utilización del ciclo estiramiento-acortamiento e índice elástico; reduce el sesgo antropométrico.
Lanzamiento	CFS rotational throw	rho notas = .64, $p = < .001$; RPE CFS = 2.42 ± 1.24 vs EFIT = 3.18 ± 1.5	El lanzamiento rotacional mantiene convergencia alta con Eurofit y añade información unilateral/asimetría más transferible a acciones motrices reales.
Velocidad	CFS 30 m	rho notas = .59, $p = < .001$; RPE CFS = 3.36 ± 1.81 vs EFIT = 3.91 ± 1.9	El 30 m mantiene convergencia alta con 50 m, pero disminuye la componente de fatiga y se ajusta mejor al objetivo de aceleración.



Resistencia

CFS Rockport

rho notas= .43, $p < .001$; RPE
CFS=4.26 $\pm$ 2.11 vs EFIT=7.35 $\pm$ 1.98Rockport mantiene convergencia con Cooper,
reduce RPE y permite valorar a perfiles con
menor condición física o restricciones de
intensidad.

Nota. La decisión combina evidencia estadística, esfuerzo percibido, sesgo antropométrico e interpretación pedagógica de la prueba.

Validación de la batería CFS frente a EFIT

La nota final CFS presentó una asociación moderada-alta con la nota final EFIT ($\rho = .54$; $p < .001$; r de Pearson = .55; $p < .001$). La regresión con validación cruzada mostró que las seis pruebas CFS predijeron la nota final EFIT con un R^2 medio de 0.4 y un error absoluto medio de 1.0 puntos en escala 0-10. Esto apoya la validez convergente de CFS como alternativa de campo. No obstante, el ICC fue moderado-bajo (.41), y el análisis de Bland-Altman mostró una diferencia media positiva de 0.9 puntos a favor de CFS, factor que señala que no debe interpretarse como una batería intercambiable, sino como una alternativa válida con criterio propio.

La consistencia interna fue mayor en EFIT (Alfa = .69) que en CFS (Alfa = .2). Este resultado es coherente con la lógica de una batería multidimensional: CFS no pretende maximizar la homogeneidad entre pruebas, sino cubrir componentes diferentes de la condición física sin redundancias y con mayor especificidad funcional. La asociación de la media CFS con el primer componente principal fue significativa ($\rho = .66$; $p < .001$), lo que confirma que CFS conserva un componente común de condición física general, con una estructura menos redundante que EFIT (Tabla 6, Figura 9).

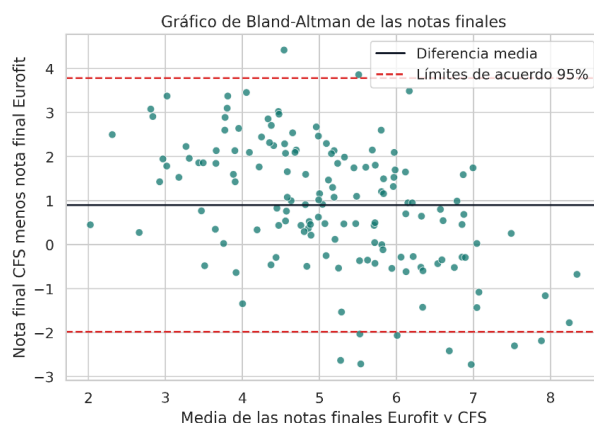
Dado que la batería integra dimensiones heterogéneas, el alfa de Cronbach no debe interpretarse como único indicador de calidad, pues puede subestimar la fiabilidad en estructuras multidimensionales y penalizar la especificidad funcional (Sijtsma, 2009; McNeish, 2018). Por ello se proponen en futuros trabajos índices de fiabilidad compuesta por dimensión o modelos bifactor.

Tabla 6. Evidencias de validación de la batería CFS.

Criterio	Prueba estadística	Estadístico	p	Interpretación
Convergencia de notas finales	Spearman rho	.54	< .001	Asociación directa entre CFS y EFIT
Convergencia lineal	Pearson r	.55	< .001	Relación lineal entre promedios
Acuerdo absoluto	ICC(2,1)	.41		Acuerdo bajo por diferencia de escala/criterio
Diferencia pareada	Wilcoxon	.61	< .001	CFS puntúa más alto que EFIT
Consistencia interna Eurofit	Alfa de Cronbach	.69		Homogeneidad de la batería original
Consistencia interna CFS	Alfa de Cronbach	.2		Homogeneidad de la alternativa
Validez factorial Eurofit	rho con PC1	.93	< .001	Carga sobre factor general de condición física
Validez factorial CFS	rho con PC1	.66	< .001	Carga sobre factor general de condición física
Predicción cruzada	R2 CV 5-fold	.4		CFS predice la nota Eurofit
Error de predicción	MAE	1.0		Error medio absoluto en escala 0-10

Nota. El ICC evalúa acuerdo absoluto; Spearman y Pearson evalúan convergencia. La validación cruzada utilizó cinco particiones.

Figura 9. Gráfico de Bland-Altman de las notas finales.



Nota. La línea central representa la diferencia media entre la nota final CFS y la nota final Eurofit. Las líneas discontinuas representan los límites de acuerdo del 95%.

Batería ideal propuesta

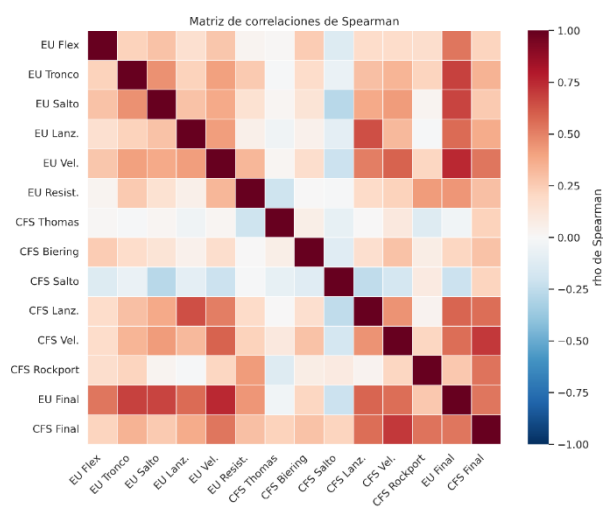
La batería ideal resultante coincide con la batería CFS completa: Thomas, Biering-Sørensen, SJ-CMJ-índice elástico, lanzamiento rotacional, 30 m y Rockport. La selección maximiza la utilidad escolar de la evaluación porque reduce esfuerzos percibidos excesivos, evita pruebas con gestos potencialmente menos recomendables, disminuye la dependencia de la antropometría en dimensiones clave como el salto y ofrece indicadores más accionables para el profesorado. En velocidad, la comparación adicional de los tramos finales no confirmó una pérdida media significativa en 50 m, pero sí mostró que más de la mitad del alumnado fue más lento en el tramo 40-50 m, reforzando la conveniencia práctica del 30 m para valorar aceleración en el contexto escolar (Tabla 7, Figura 10).

Tabla 7. Composición propuesta de la batería ideal y criterios de decisión.

Dimensión	Prueba ideal	RPE EFIT	RPE CFS	Bias antropométrico EFIT	Bias antropométrico CFS	rho horas EFIT	rho horas CFS	Decisión
Flexibilidad	CFS Thomas	5.68	2.55	.05	.05	.23	.17	CFS
Tronco	CFS Biering-Sørensen	5.85	3.72	.11	.07	.23	.11	CFS
Salto	CFS SJ-CMJ-EI	1.85	2.49	.14	.03	.20	.02	CFS
Lanzamiento	CFS lanz. rotacional	3.18	2.42	.25	.27	.25	.30	CFS
Velocidad	CFS 30 m	3.91	3.36	.10	.14	.31	.21	CFS
Resistencia	CFS Rockport	7.35	4.26	.14	.20	.36	.37	CFS

Nota. La decisión se expresa en función de los resultados estadísticos y de la adecuación pedagógica/fisiológica de la prueba.

Figura 10. Matriz de correlaciones de Spearman.



Nota. Los tonos azules indican asociaciones positivas y los rojos negativos entre pruebas y notas finales. rho = coeficiente de correlación de Spearman.

Las horas semanales de práctica de actividad física se asociaron de forma más clara con las pruebas de rendimiento global que con las pruebas de movilidad o tronco. En EFIT, las correlaciones con práctica semanal fueron más evidentes en resistencia, velocidad y lanzamiento; en CFS, destacaron resistencia y lanzamiento. Este aspecto refuerza que las pruebas de campo captan el efecto de la práctica extraescolar, aunque no de manera uniforme. La ausencia de diferencias por sexo en las notas finales, pese a diferencias antropométricas significativas, sugiere que la evaluación por baremos evita que la batería se convierta en una simple medida del tamaño corporal.

En conjunto, los datos permiten validar CFS como alternativa viable a EFIT en el contexto de Educación Física en ESO. La validación es convergente y funcional, no de equivalencia absoluta. La batería CFS conserva relación significativa con EFIT, predice una parte relevante de su nota final y ofrece ventajas en esfuerzo percibido, especificidad de la medición e interpretabilidad pedagógica.

Por ello, la recomendación derivada de estos resultados es sustituir la batería Eurofit por la batería CFS cuando el objetivo sea una evaluación escolar más segura, específica y útil para orientar la mejora individual del alumnado.

Discusión

El estudio ofrece una alternativa viable en la evaluación de la condición física, a través de una batería específica, segura, inclusiva y útil desde una perspectiva pedagógica (Ruiz et al., 2011). Sin pretender sustituir o anular EFIT; aporta una lectura funcional diferente y más ajustada al contexto escolar actual y coherente con la consideración de la condición física en la infancia y adolescencia como un marcador relevante de salud (Ortega et al., 2008). La batería CFS se asocia en la nota final de forma significativa con EFIT, presentando una validación convergente. A pesar de que la prueba Bland-Altman y el ICC resultaran bajos, CFS no fue diseñada para ser una réplica, sino para constituir una prueba propia, por lo que el resultado es lógico desde ese enfoque; no son baterías intercambiables.

Por otro lado, la batería CFS reduce el esfuerzo percibido a nivel global 32.4% ($p < .001$; $d z = -1.73$) indicando un descenso de la carga interna (Borg 1982), sobre todo en flexibilidad, tronco, lanzamiento, velocidad y resistencia, algo importante y aplicable a grupos escolares heterogéneos con diferente motivación (EFIT 4.64 ± 1.11 ; CFS 3.13 ± 1.13). Las pruebas muy exigentes miden no solo la capacidad física, sino la tolerancia al esfuerzo, experiencia competitiva, o capacidad de gestionar la fatiga, algo muy resaltable en la prueba de Rockport de la batería CFS, por tanto, se promueve la evaluación sin exclusión.

Considerando la flexibilidad, no hay correlación entre pruebas, al valorar componentes distintos. El test sit and reach está influido por la posición de la cadera y la flexión lumbar, y el de Thomas se centra en los acortamientos de la zona lumbar-crural, en consonancia con las premisas de ACSM, con información más concreta y accionable para la intervención docente por su facilidad de puesta en marcha (Peeler & Anderson 2008), además de menor percepción de esfuerzo (Thomas 2.55 ± 1.80 ; sit and reach 5.68 ± 2.38 ; $p < .001$; $d z = -1.24$).

En referencia al tronco, la prueba Biering-Sørensen obtuvo una asociación baja pero significativa con el test abdominal ($\rho = .18$; $p = .026$), compartiendo parte de la información, aunque sin evaluar el mismo comportamiento motor. Este aspecto aporta una alternativa funcional y de control postural además de un esfuerzo menor (CFS 3.72 ± 1.92 ; EFIT 5.85 ± 2.13 ; $p < .001$; $d z = -.81$). Este hecho la posiciona como la más elegible para el contexto escolar al contar con un gran respaldo científico (Martínez-Romero et al., 2020) y evitar repeticiones balísticas que pueden provocar sobrecargas en flexores de cadera.

En el plano de los saltos, ambas baterías comparten prácticamente la misma nota. A pesar de esto los valores directos mostraron una asociación alta ($\rho = .73$; $p < .001$), y las notas por baremo se asociaron de forma negativa ($\rho = -.28$; $p < .001$). La combinación de SJ y CMJ permite diferenciar el componente concéntrico de la fuerza explosiva y el aprovechamiento del ciclo estiramiento-acortamiento, por lo que ofrece una lectura más objetiva que una única prueba (Bosco & Komi 1979; Markovic et al., 2004).

El cambio a un lanzamiento rotacional con ambos lados aporta gran convergencia con EFIT, tanto en la nota ($\rho = .64$; $p < .001$) como en el valor directo de la prueba ($\rho = .72$; $p < .001$). Además, la nota CFS fue superior a la EFIT (6.55 ± 2.96 frente a 4.20 ± 2.83 ; $p < .001$) y el esfuerzo percibido fue menor (2.42 ± 1.24 frente a 3.18 ± 1.50 ; $p < .001$). Esto aporta una información novedosa muy relevante, como la lateralidad y la coordinación intersegmentaria apoyada por la evidencia reciente sobre pruebas de potencia rotacional con balón medicinal (Hardy et al., 2025). Aunque permite detectar asimetrías (inviabile en el lanzamiento bilateral), la baremación escolar empleada en este estudio debe interpretarse como una propuesta propia y provisional pendiente de validación externa publicada.

En velocidad, los 30 m CFS mostraron una elevada convergencia con los 50 m y el esfuerzo percibido fue menor (3.36 ± 1.81) que en EFIT (3.91 ± 1.90 ; $p < .001$). La comparación específica de los últimos 10 m obtuvo tiempos similares entre el tramo 40-50 m de Eurofit y el de 20-30 m de CFS, sin diferencias



significativas ($p = .148$). No obstante, el 56.9% del alumnado fue más lento en el tramo 40-50 m que en el 20-30 m, y en ambos casos se asociaron de forma alta ($\rho = .78$; $p < .001$). Por esta razón, no se confirma una pérdida de velocidad a nivel grupal, pero sí una tendencia individual que refuerza la idoneidad del 30 m como prueba a emplear, menos fatigante y ajustada a valorar la aceleración en alumnado inexperto y que cuenta con evidencia previa de fiabilidad interevaluador y validez, lo que respalda su aplicabilidad en este contexto (Vicente-Rodríguez et al., 2011).

Por último, en el caso de la resistencia, ambos test mantuvieron una asociación significativa ($\rho = .43$; $p < .001$) como en el VO_2 estimado ($\rho = .51$; $p < .001$). El test de Cooper es más exigente con el perfil del alumnado, mientras que Rockport permite autorregular intensidad y reduce la percepción del esfuerzo (RPE Cooper fue de 7.35 ± 1.98 , frente a 4.26 ± 2.11 en Rockport ($p < .001$; $d z = -1.13$)) aumentando el componente inclusivo, visible en las notas (Cooper 1.96 ± 2.73 ; Rockport 5.49 ± 2.78 ; $p < .001$) destacando su utilidad como alternativa de campo para estimar la capacidad cardiorrespiratoria en esta población (McSwegin et al., 1998).

Centrando la atención a la relación entre antropometría y rendimiento, se observaron diferencias significativas por sexo en peso ($p = .002$), talla, envergadura, diámetro biacromial y longitud de pierna ($p < .001$), aunque no en la nota final (EFIT $p = .360$; CFS $p = .368$). Esto indica que los baremos empleados contribuyen a equilibrar la interpretación del rendimiento, evitando que la calificación sea expresión única del tamaño corporal. En cualquier caso, algunas pruebas mantuvieron cierta relación con variables antropométricas (lanzamiento con sesgo de .25 en EFIT y .27 en CFS; y en resistencia CFS, con un valor medio de .20). Por estas razones siempre ha de considerarse en la interpretación la edad, sexo y características corporales del alumnado, algo en línea con Horta-Gim et al. (2025), quienes muestran relaciones entre composición corporal y condición física, sobre todo en adiposidad, IMC, variables cardiorrespiratorias y de fuerza.

En la baja consistencia de CFS con respecto a EFIT (.20 frente a .69), ha de tenerse en cuenta que, a diferencia de una escala psicométrica unidimensional, una batería de condición física integra capacidades heterogéneas, como flexibilidad, fuerza, velocidad y resistencia. Por lo que un Alfa elevado no es necesariamente el criterio principal de calidad, al poder indicar redundancia entre pruebas. A pesar de que en esta investigación CFS mostró una estructura menos homogénea, fue más específica y coherente con su propósito de valorar dimensiones diferenciadas de la condición física escolar.

Las posibles aplicaciones de los hallazgos pueden ser múltiples; como proporcionar coherencia con los modelos de desarrollo motor que relacionan competencia motriz, condición física, percepción de competencia (Stodden et al., 2008) y estar adaptadas al nivel madurativo al dar información del perfil neuromuscular (Faigenbaum et al., 2016). Una evaluación útil siempre debe identificar patrones de mejora y no solo clasificar al alumnado con ello se favorece la inclusión y una valoración más equitativa.

Por otro lado, la batería CFS permite una intervención más específica y enfocada (baremar por índice elástico, acortamiento de flexores de cadera, encontrar soluciones a la baja tolerancia al trabajo de alta intensidad, etc.). Esto es posible por las evidencias iniciales de validez convergente y aplicabilidad, aspectos que permiten emplear esta batería como alternativa a EFIT. Esto cumple con la necesidad de revisar/actualizar las pruebas escolares por la diversidad existente, la falta de consenso en algunos componentes o la importancia de seleccionar instrumentos factibles para el contexto escolar (Brazo-Sayavera et al., 2024).

Con independencia de los resultados, el estudio presenta algunas limitaciones. La batería Eurofit se compone de un mayor número de pruebas elegibles por el profesorado. En esta investigación se han seleccionado seis de ellas, por lo que no se ha establecido comparación o alternativa con todas. En esta línea, el Alfa reducido de CFS es necesario estudiarlo con análisis factoriales o modelos multidimensionales.

Además, solo se ha contado con Eurofit como referencia comparativa, sin entrar en otras baterías existentes que podrían dar una visión más amplia. Por otro lado, incluir otras pruebas externas a las incluidas en estas baterías, podría dar una visión más global y trabajar todas las facetas implícitas de la condición física que gozan de mayor visibilidad actual (agilidad, propiocepción, etc.).

La baremación del lanzamiento rotacional procede de una propuesta propia del grupo investigador y se encuentra pendiente de publicación por parte de los autores. Aunque el protocolo se apoya en la literatura sobre potencia rotacional con balón medicinal, se deberá confirmar la estabilidad, validez externa y puntos de corte.

Para evitar circularidad, los análisis de validez se presentan considerando el lanzamiento rotacional como prueba exploratoria con baremo interno provisional. Los resultados principales de validez convergente se mantienen sin cambios relevantes al excluir esta prueba del cálculo de la nota final CFS ($\rho = ,52$; $p < ,001$), hecho que refuerza que la conclusión no depende exclusivamente de este baremo.

Por último, el diseño de esta investigación fue de carácter transversal. La comparación durante un curso escolar completo y entre cursos de esta etapa educativa podría ayudar a dar solidez a los resultados para comprobar su estabilidad a lo largo del tiempo.

Tampoco se realizó un estudio intraestudio test-retest para pruebas clave como Biering-Sørensen, Thomas y SJ/CMJ bajo el protocolo específico aplicado, lo que limita la estimación de fiabilidad intra e interexaminador en esta muestra. Se reconoce como limitación y se propone incluir doble medición en un subgrupo del 10-15% en futuros trabajos.

Para futuras investigaciones, sería muy conveniente contar con un mayor número de centros educativos y contextos socioeducativos, favoreciendo la diversidad y detectando mejor la sensibilidad al cambio. Además, identificar otros factores actualmente relevantes en salud escolar como el sueño, dolor musculoesquelético, adherencia y otros factores motivacionales.

Conclusiones

Los resultados aportan evidencias iniciales de que la batería CFS constituye una alternativa viable a Eurofit para la evaluación de la condición física en alumnado de Educación Secundaria. Su relación significativa con la nota final Eurofit, respalda su validez convergente, aunque el grado de acuerdo entre baterías indica que no deben interpretarse como instrumentos intercambiables punto por punto, aspecto que ofrece una lectura funcional propia y más ajustada al contexto educativo.

La batería CFS redujo el esfuerzo percibido global, manteniendo asociaciones relevantes con las pruebas tradicionales en dimensiones como velocidad, lanzamiento y resistencia. Este hallazgo resulta especialmente importante en Educación Física, donde la evaluación debe ser aplicable a grupos heterogéneos, inclusiva y útil para orientar la intervención docente, no solo para clasificar el rendimiento.

Por dimensiones, los 30 m se mostraron como una prueba más breve y adecuada para valorar la aceleración en alumnado no especialista; Rockport ofreció una alternativa menos exigente y más inclusiva que Cooper; Biering-Sørensen permitió valorar el tronco desde una perspectiva funcional e isométrica; la combinación SJ-CMJ-índice elástico aportó una interpretación más diagnóstica del salto; el lanzamiento rotacional añadió información sobre lateralidad y coordinación intersegmentaria; y el test de Thomas proporcionó una valoración más localizada y accionable de la flexibilidad.

En conjunto, la CFS permite obtener, aparte de la calificación final, un perfil más específico de condición física, facilitando la identificación de necesidades individuales y la programación de tareas de mejora. Por ello, puede considerarse una propuesta útil, segura y pedagógicamente pertinente para actualizar la evaluación de la condición física en Educación Física, si bien futuras investigaciones deberán confirmar su fiabilidad, sensibilidad al cambio y validez en muestras más amplias y diversos contextos escolares.

Agradecimientos

Los autores desean expresar su agradecimiento a los centros y al alumnado participante en el presente estudio.



Referencias

- Bandyopadhyay, A. (2015). Validity of Cooper's 12-minute run test for estimation of maximum oxygen uptake in male university students. *Biology of Sport*, 32(1), 59–63. <https://doi.org/10.5604/20831862.1127283>
- Blázquez Sánchez, D. (2006). *Evaluar en educación física*. Inde.
- Borg, G. A. (1982). Psychophysical bases of perceived exertion. *Medicine and Science in Sports and Exercise*, 14(5), 377–381.
- Borms, D., Maenhout, A., & Cools, A. M. (2016). Upper quadrant field tests and isokinetic upper limb strength in overhead athletes. *Journal of Athletic Training*, 51(10), 789–796. <https://doi.org/10.4085/1062-6050-51.12.06>
- Bosco, C., & Komi, P. V. (1979). Mechanical characteristics and fiber composition of human leg extensor muscles. *European Journal of Applied Physiology and Occupational Physiology*, 41(4), 275–284. <https://doi.org/10.1007/BF00429744>
- Brazo-Sayavera, J., Silva, D. R., Lang, J. J., Tomkinson, G. R., Agostinis-Sobrinho, C., Andersen, L. B., García-Hermoso, A., Gaya, A. R., Jurak, G., Lee, E. Y., Liu, Y., Lubans, D. R., Okely, A. D., Ortega, F. B., Ruiz, J. R., Tremblay, M. S., & Dos Santos, L. (2024). Physical fitness surveillance and monitoring systems inventory for children and adolescents: A scoping review with a global perspective. *Sports Medicine*, 54(7), 1755–1769. <https://doi.org/10.1007/s40279-024-02038-9>
- Castro-Piñero, J., Artero, E. G., España-Romero, V., Ortega, F. B., Sjöström, M., Suni, J., & Ruiz, J. R. (2010). Criterion-related validity of field-based fitness tests in youth: A systematic review. *British Journal of Sports Medicine*, 44(13), 934–943. <https://doi.org/10.1136/bjsm.2009.058321>
- Cordente-Martínez, C. (2006). Estudio epidemiológico del nivel de actividad física y de otros parámetros de interés relacionados con la salud bio-psico-social en los alumnos de educación secundaria obligatoria del municipio de Madrid [Tesis doctoral, Universidad Politécnica de Madrid].
- Cuadrado, G., Morante, J. C., Redondo, J. C., & Zarzuela, R. (2009). *Valoración de la condición física de la población escolar mediante la batería Eurofit*. Wanceulen.
- Cureton, K. J., Sloniger, M. A., O'Bannon, J. P., Black, D. M., & McCormack, W. P. (1995). A generalized equation for prediction of VO₂peak from 1-mile run/walk performance. *Medicine and Science in Sports and Exercise*, 27(3), 445–451.
- Faigenbaum, A. D., Lloyd, R. S., MacDonald, J., & Myer, G. D. (2016). Citius, Altius, Fortius: Beneficial effects of resistance training for young athletes: Narrative review. *British Journal of Sports Medicine*, 50(1), 3–7. <https://doi.org/10.1136/bjsports-2015-094621>
- Galán-Arroyo, C., Mendoza-Muñoz, D. M., Mañana-Iglesias, C., & Rojo-Ramos, J. (2024). La aptitud física, indicador de un desarrollo saludable en el preadolescente. *Retos*, 52, 447–456. <https://doi.org/10.47197/retos.v52.99772>
- González, Y., Jiménez Garzón, L. E., Díaz Marín, J. M., & Díaz, H. (2013). Valoración de las capacidades físicas condicionales en escolares de básica secundaria y media del colegio distrital Gerardo Paredes de la localidad de Suba. *Movimiento Científico*, 7(1), 93–104. <http://dialnet.unirioja.es/servlet/oaiart?codigo=4781967>
- Hardy, S. G. J., Stelzer-Hiller, O. W., Edwards, K. M., & Freeston, J. (2025). Criterion validity and reliability of a new medicine ball rotational power test. *Journal of Strength and Conditioning Research*, 39(3), e429–e435. <https://doi.org/10.1519/JSC.00000000000005001>
- Horta-Gim, M. A., Rendon-Delcid, P. A., Vega-Orozco, S. I., Horta Gim, V. H., & Romero-Pérez, E. M. (2025). Relación entre composición corporal y condición física en niños y adolescentes del noroeste de México. *Retos*, 66, 1085–1093. <https://doi.org/10.47197/retos.v66.113985>
- Inbar, O., & Amon, E. (2016). A new field-test for the assessment of the human anaerobic and repeated sprints capability reliability and validity. *Harefuah*, 155(6), 335–339.
- Izquierdo, M., & González-Badillo, J. J. (2008). Propiedades biomecánicas del músculo. En M. Izquierdo (Ed.), *Biomecánica y bases neuromusculares de la actividad física y el deporte* (pp. 173–201). Médica Panamericana.
- Lagally, K. M., & Robertson, R. J. (2006). Construct validity of the OMNI resistance exercise scale. *Journal of Strength and Conditioning Research*, 20(1), 252–256. <https://doi.org/10.1519/r-17224.1>
- Markovic, G., Dizdar, D., Jukic, I., & Cardinale, M. (2004). Reliability and factorial validity of squat and countermovement jump tests. *Journal of Strength and Conditioning Research*, 18(3), 551–555. [https://doi.org/10.1519/1533-4287\(2004\)18%3C551:rafvos%3E2.0.co;2](https://doi.org/10.1519/1533-4287(2004)18%3C551:rafvos%3E2.0.co;2)



- Martínez-Romero, M. T., Ayala, F., De Ste Croix, M., Vera-García, F. J., Sainz de Baranda, P., Santonja-Medina, F., & Sánchez-Meca, J. (2020). A meta-analysis of the reliability of four field-based trunk extension endurance tests. *International Journal of Environmental Research and Public Health*, 17(9), 3088. <https://doi.org/10.3390/ijerph17093088>
- McNeish, D. (2018). Thanks coefficient alpha, we'll take it from here. *Psychological Methods*, 23(3), 412–433. <https://doi.org/10.1037/met0000144>
- McSwegin, P. J., Plowman, S. A., Wolff, G. M., & Guttenberg, G. L. (1998). The validity of a one-mile walk test for high school age individuals. *Measurement in Physical Education and Exercise Science*, 2(1), 47–63. https://doi.org/10.1207/s15327841mpee0201_4
- Miñarro, P. (2008). *Ejercicios desaconsejados en la actividad física*. Inde.
- O'Gorman, D. J., Hunter, A., McDonnacha, C., & Kirwan, J. P. (2000). Validity of field tests for evaluating endurance capacity in competitive and international-level sports participants. *Journal of Strength and Conditioning Research*, 14(1), 62–67.
- Ortega, F. B., Ruiz, J. R., Castillo, M. J., & Sjöström, M. (2008). Physical fitness in childhood and adolescence: A powerful marker of health. *International Journal of Obesity*, 32(1), 1–11. <https://doi.org/10.1038/sj.ijo.0803774>
- Peeler, J. D., & Anderson, J. E. (2008). Reliability limits of the modified Thomas test for assessing rectus femoris muscle flexibility about the knee joint. *Journal of Athletic Training*, 43(5), 470–476. <https://doi.org/10.4085/1062-6050-43.5.470>
- Petrigna, L., Karsten, B., Marcolin, G., Paoli, A., D'Antona, G., Palma, A., & Bianco, A. (2019). A review of countermovement and squat jump testing methods in the context of public health examination in adolescence: Reliability and feasibility of current testing procedures. *Frontiers in Physiology*, 10, 1384. <https://doi.org/10.3389/fphys.2019.01384>
- Ruiz, J. R., Castro-Piñero, J., España-Romero, V., Artero, E. G., Ortega, F. B., Cuenca, M. M., Jiménez-Pavón, D., Chillón, P., Girela-Rejón, M. J., Mora, J., Gutiérrez, A., Suni, J., Sjöström, M., & Castillo, M. J. (2011). Field-based fitness assessment in young people: The ALPHA health-related fitness test battery for children and adolescents. *British Journal of Sports Medicine*, 45(6), 518–524. <https://doi.org/10.1136/bjsm.2010.075341>
- Sainz de Baranda, P., & Ayala, F. (2010). Chronic flexibility improvement after 12 weeks of stretching program utilizing the ACSM recommendations: Hamstring flexibility. *International Journal of Sports Medicine*, 31(6), 389–396. <https://doi.org/10.1055/s-0030-1249082>
- Sargent, D. A. (1921). The physical test of a man. *American Physical Education Review*, 26(4), 188–194. <https://doi.org/10.1080/23267224.1921.10650486>
- Sijtsma, K. (2009). On the use, the misuse, and the very limited usefulness of Cronbach's alpha. *Psychometrika*, 74(1), 107–120. <https://doi.org/10.1007/s11336-008-9101-0>
- Smith, J. J., Eather, N., Morgan, P. J., Plotnikoff, R. C., Faigenbaum, A. D., & Lubans, D. R. (2014). The health benefits of muscular fitness for children and adolescents: A systematic review and meta-analysis. *Sports Medicine*, 44(9), 1209–1223. <https://doi.org/10.1007/s40279-014-0196-4>
- Stodden, D. F., Goodway, J. D., Langendorfer, S. J., Roberton, M. A., Rudisill, M. E., Garcia, C., & Garcia, L. E. (2008). A developmental perspective on the role of motor skill competence in physical activity: An emergent relationship. *Quest*, 60(2), 290–306. <https://doi.org/10.1080/00336297.2008.10483582>
- Tomkinson, G. R., Olds, T. S., & Borms, J. (2007). Who are the Eurofittest? *Medicine and Sport Science*, 50, 104–128. <https://doi.org/10.1159/000101355>
- Vicente-Rodríguez, G., Rey-López, J. P., Ruiz, J. R., Jiménez-Pavón, D., Bergman, P., Ciarapica, D., Heredia, J. M., Molnar, D., Gutierrez, A., Moreno, L. A., & Ortega, F. B. (2011). Interrater reliability and time measurement validity of speed-agility field tests in adolescents. *Journal of Strength and Conditioning Research*, 25(7), 2059–2063. <https://doi.org/10.1519/JSC.0b013e3181e742fe>
- Ward-Smith, A. J. (2001). Energy conversion strategies during 100-m sprinting. *Journal of Sports Sciences*, 19(9), 701–710. <https://doi.org/10.1080/02640410152475838>
- Wells, K. F., & Dillon, E. K. (1952). The sit and reach: A test of back and leg flexibility. Research Quarterly. American Association for Health, *Physical Education and Recreation*, 23(1), 115–118. <https://doi.org/10.1080/10671188.1952.10761965>



Datos de los/as autores/as y traductor/a:

Julio Martín Ruiz
Concepción Ros Ros
Clara Gallego Cerveró
Ignacio Tamarit Grancha
Laura Ruiz Sanchis

julio.martin@ucv.es
concepcion.ros@ucv.es
clara.galego@ucv.es
ignacio.tamarit@ucv.es
laura.ruiz@ucv.es

Autor/Responsable traducción
Autora
Autora
Autor
Autora